John Angle and Paul Siegel

Department of Sociology

A THEORY OF THE DIFFUSION OF BILINGUALISM IN POPULATIONS:
AN APPLICATION OF THE LOG-LINEAR ANALOGUE
OF FIRST DIFFERENCES

Working Paper #127
of the Center for Research on
Social Organization

February 1976

Copies Available Through:
The Center for Research on
Social Organization
The University of Michigan
330 Packard Street #214
Ann Arbor, Michigan 48109
A THEORY OF THE DIFFUSION OF BILINGUALISM IN POPULATIONS:
AN APPLICATION OF THE LOG-LINEAR ANALOGUE
OF FIRST DIFFERENCES

John Angle and Paul Siegel

February 17, 1976

University of Michigan

Profs. Albert Hermalin and Robbins Burling have commented on many previous drafts of this paper. Prof. Leslie Kish suggested a particularly elegant and painless way to sample the 1935 Special Census of Puerto Rico. The people of National Archives were very helpful. Joseph Howerton suggested a strategy for obtaining permission to access the schedules of the 1935 Special Census. Frederick Pernell gave cheerful assistance to the sampling. Louis Smith and others did a lot of hard work pulling out the folios with the schedules. Robert Schweitzer and Arthur Lerner gave advice on document photography. This research was funded in part by a dissertation grant from the Horace Rackham School of Graduate Studies of The University of Michigan and a grant from N.I.M.H. training grant #5T01 MH10244 administered by the Department of Sociology of The University of Michigan.
ABSTRACT

Language diffusion has two phases, the diffusion of bilingualism and language shift. There is a literature which indicates that the rate of language diffusion in a population is determined by the level of economic development in a country. Lieberson et al. (1974, 1975) have challenged this generalization for language shift. Using a pooled sample of the censuses of population in Puerto Rico of 1935 and 1970, this paper finds that the diffusion of bilingualism is related to economic development, specifically to a change in the occupational structure.

This paper tests a theory of why a population in a country becomes bilingual in a language at a particular rate. The data are from samples of individual records of the 1935 Special Census of Population and the 1970 Census of Population and Housing in Puerto Rico. The paper uses a specification of log-linear models which is analogous to first differencing. Such a specification is used because the theory calls for a test that holds cross-sectional relationships invariant.

There is an established literature which argues that the direction of language diffusion is determined by social power but that its rate is related to the level of economic development in the country. The higher the level of economic development the faster would be the rate of language diffusion. This paper tests whether one facet of economic development, change in the occupational structure, has an effect on the spread of bilingualism in English in Puerto Rico.

The theory is one of a market. Language learning is hypothesized to occur in proportion to the price it brings in the labor market. Its price is the advantage bilinguals have in occupational placement. The expansion of employment, particularly white collar employment, generates opportunities for bilinguals to reap this reward. Over time the rate of diffusion of bilingualism should closely lag fluctuations in the advantage of bilinguals in the labor market.
and the expansion of the labor force in the direction of greater white collar employment, if the economic reward theory of the diffusion of bilingualism is true.

Although the theory that economic development drives the process of language diffusion is well known. It has recently been challenged (cf. Lieberson and Hansen, 1974: Lieberson, Dalto, and Johnston, 1975). Lieberson et al (1974, 1975) argue that while cross-sectional regression of an index of mother-tongue heterogeneity on indicators of economic development over countries shows that the economically more developed are more homogeneous, it is incorrect to assume that economic development itself speeds language diffusion. Over-time data presented by Lieberson et al. shows little relationship between indicators of economic development and mother-tongue heterogeneity, when cross-sectional effects are held invariant. It is Lieberson et al.'s discounting of the value of cross-sectional data that necessitates use of the log-linear analogue of first differences, a model which holds cross-sectional effects invariant.

Lieberson et al.'s challenge to a widely accepted theory forces a re-examination of concepts. Language diffusion has two phases, not carefully distinguished by most of the historians and linguists who have written on the subject. Thus Brun (1923a, 1923b) and Brunot (1967a, 1967b), historians of the diffusion of the French language in France, one of the better documented cases, hardly ever distinguish between the spread of bilingualism and the later process of bilinguals switching their better mastered
language from the original language to French, the process of language shift. Language shift is usually an inter-generational process. The second language of the parent simply becomes the first language of the children. Perhaps over the centuries the latter process took the same route through the social structure as the former in France. However, when time is measured in decades rather than centuries, the two processes are not necessarily correlated over time. Lieberson (1970:13,95) documents that the spread of bilingualism in English among French Canadians is not correlated in recent census years with shift from French to English mother-tongue among French Canadian bilinguals. Angle (1976a:chap. 3) documents the fact that bilingualism in English has become widespread in Puerto Rico, while very little if any shift to English has occurred among ethnic Puerto Ricans who are not return migrants from the mainland.

Since Lieberson et al.'s data are closely related to the second phase of language diffusion, language shift, it may be that the theory that economic development in general, or change in the occupational structure in particular, drives language diffusion is valid for its first phase, the spread of bilingualism, but not its second, language shift. If such is the case, there is hope that more of the world's people will not have to undergo the uprooting experience of having to give up their mother-tongue in order to enjoy a better standard of living.

The Qualitative Literature on Language Diffusion

Haugen (1972:258) states the most widely accepted
proposition in the literature on language diffusion, when he writes in a review article that "[the spread of languages] ... is everywhere the result of a concentration of political power ..." which creates incentives for people to learn the language of those with that power. Anthropologists (cf. Leach, 1954:50) have found the same process at work in small, primitive societies. Language diffusion often happens as a form of "passing," that is, presenting oneself as a member of a more privileged group in a linguistically stratified society (cf. Greenberg, 1971:206). The question of which language is learned is relatively uninteresting, since it is so much a matter of relative numbers, history, and relative social power, contextual factors which are hard to quantify. Far more interesting is the question of why languages spread at varying rates.

Hertzler (1965: chapter 7; 1966) has abstracted propositions on what affects the rate of language diffusion from the literature on the subject among linguists. His chief source is Jespersen (1946), who himself reviews a number of sources. Hertzler's (1966:178-179) is a very clear statement of the accepted view that economic development drives the process of language diffusion by increasing the range over which people communicate and the importance of communication. Hertzler (1966) sees language as part of the "unification" of the mass society of the industrial state. Deutsch (1966: chap. 6,7) argues in a similar but somewhat more complex vein. He notes that in pre-industrial societies, languages take centuries to diffuse and that the
Process of diffusion occurs much more rapidly after industrialization. However, speakers of minority languages in industrial societies are more likely to actively and consciously resist the process, the phenomenon of nationalism in defense of a minority language. In Etzioni's (1967) terms language is less likely to be an "on-going," or un-guided, process after industrialization, or as Peyre (1933) argues, after the French Revolution, which established nationalism, defined as it is usually in linguistic terms (cf. Fishman, 1973), as a social movement in the industrializing world. If Deutsch is correct, as long as language diffusion remains an "on-going" social process, its rate will be directly related to economic development, but the likelihood of its remaining "on-going" decreases with economic development.

The deduction can be made from Deutsch's theory that countries which developed first economically and most slowly would be linguistically more homogeneous than countries which began developing rapidly only after the spread of the idea of linguistic nationalism in the 19th century. The former countries would have a longer time for the "on-going" process of language diffusion to have its effect, without opposition, than the latter. Such an historical sequence could account both for a cross-sectional relationship between mother-tongue homogeneity across countries and indicators of economic development and the findings of Lieberson et al. (1974, 1975) that change in an index of mother-tongue homogeneity is not related to changes in
indicators of economic development in comparatively recent years.

Deutsch (1966) does not specify what it is in economic development which either speeds language diffusion or turns the defense of minority languages into a political issue that can tear the state apart or check the diffusion process. Three mechanisms by which economic development can affect language diffusion are discussed in the qualitative literature. These are: 1) the mixing of people of different mother-tongues and the resulting need for a lingua franca, 2) the creation of more jobs with higher pay, prestige, and security and for which a price in linguistic conformity can be exacted from those who would be placed into these jobs, and 3) the related phenomenon of the expansion of white collar employment with the growth of organizations which, because of their need for communication, demand linguistic conformity from people with whom they deal as well as employees.

Brunot (1967:187-191,193) found that in France by the late 18th and early 19th centuries the French spoken by workers in cities was closer to standard French than that spoken by peasants in the countryside. Workers in cities chose colloquial standard French as a lingua franca. Brunot also found that even migrant agricultural workers in the same period were more likely to speak French than peasants who stayed near their village. Meillet (1928:103-108) accounts for the rapid diffusion of French in France during the 19th century by the same mechanism. Lieberson and
Curry (1971) found that the mixing of people of different mother-tongues in immigrant neighborhoods in American cities fostered the use of English as a lingua franca.

Novicow (1903:24) notes that immigrants shift more quickly to the language of the host country where there are occupational opportunities open to them. Meillet (1928:108) sees upward social mobility as a factor in the diffusion of French in France. Lieberson (1970:84,85) infers that the ownership and management of many large corporations in Quebec Province by speakers of English provides advantages for native speakers of English over those with French mother-tongue, and probably for French-Canadian bilinguals over French monolinguals. If a language group controls the government, it can require members of other mother-tongue groups to learn its language as a prerequisite for the certificates of educational achievement typically required for placement into the better paying, more prestigious occupations in industrial societies. Political control of the government by a language group usually enables it to make its language the language of the civil service. To the extent that the major business organizations too could create rewards for learning the language of the elite and penalties for not learning it, they too could influence the process. Presumably, the elite would want the other group to learn its language so as not to have to experience the marginality of attempting to work and live using a language that is less than fully mastered.

Meillet (1928:151,152) thought that the growth of
organizations was a particularly important way in which economic development put pressure on people to learn the language of a country's elite. He notes that a language barrier in dealing with these organizations would disadvantage a language group. Organizations are also a source of well paid, secure, language-sensitive jobs and their growth would create incentives to conform to their working language.

The Quantitative Literature on Language Diffusion

Much of the quantitative literature on language diffusion has a measure of mother-tongue diversity as the dependent variable. Where a plurality of a population has one mother-tongue, a decrease in a measure of mother-tongue diversity indicates shift toward the language spoken by the plurality. There are two studies (Pool, 1971; Fishman, 1968), reviewed carefully by Lieberson and Hansen (1974), which regress a measure of mother-tongue diversity on various measures of economic development. Almost any measure of economic development is highly correlated with mother-tongue homogeneity over countries. These findings confirm the widespread, traditional view that language diversity is a result of transportation and communication barriers which technology and economic development remove, within states if not necessarily between them. However, cross-sectional studies can mislead an investigator into making incorrect inferences about the over-time process which may have created the observed relationships (Carlsson, 1972).

Lieberson and Hansen (1974) and Lieberson, Dalto, and Johnston (1975) have found very little relationship
between measures of economic development and measures of linguistic diversity. Lieberson and Hansen (1974) take 23 European countries in 1930 and 1960 and attempt to correlate the variance in 1960 unexplained by 1930 diversity with measures of economic development at the two time points, literacy and urbanization. Hardly any relationship between economic development and changes in mother-tongue diversity was found. Lieberson and Hansen (1974:526) note however that 30 years may be too short a time to measure causal lags in language shift, usually an intergenerational process. They then examine longer time-series for eight countries and find no obvious pattern relating urbanization and literacy with change in mother-tongue diversity in most of these cases. Lieberson, Dalto, and Johnston (1975) extend the time range and the number of cases of Lieberson and Hansen (1974). The former paper confirms the findings of the latter paper and suggests that the primary explanatory variables of change in mother-tongue diversity are the relative sizes of the language groups within a country and the political control of the country by a language group, in short, contextual factors rather than the uniform homogenizing effect of economic development.

**Testing the Economic Reward Theory of the Diffusion of Biliteracy**

Lieberson et al. (1974, 1975) worked with aggregated data. If economic development influences the rate of language diffusion in countries, it is because it affects individual language learning behavior. However, few sources of
data on individual language learning or language shift in large populations at two or more points in time exist. The pooled sample of the 1935 and 1970 censuses in Puerto Rico is one of the very few sources on second language abilities at two points in time.

The economic reward theory asserts that the principal reason people learn a new language is that they are rewarded for it in the labor market. Reward could be in terms of an advantage bilinguals have in occupational placement, or among those in similar occupations, higher earnings for bilinguals, net of other factors which affect earnings. An advantage in occupational placement would be selective recruitment of bilinguals to occupations with higher earnings and prestige than otherwise similar monolinguals would, on the average, be recruited to.

It is assumed that if a tangible reward for bilingualism per se exists, word of it will get around and people will modify their behavior in order to take advantage of it. The validity of this assumption could, of course, be investigated. The model is that of a market. There are two aspects to the size of the reward for bilingualism. First is the degree of discrimination in favor of bilinguals, against monolinguals. The degree of discrimination is a matter of social power. In Puerto Rico it is likely that the degree of discrimination in favor of bilingual ethnic Puerto Ricans was greater during the American colonial period. A second aspect of the reward for bilingualism is the number of jobs that open up in a period of time. This second factor, is
what the economic reward theory takes as the driving force of the diffusion of bilingualism. The advantage of bilinguals at one point in time is a given, a parameter, of the process.

Data on individual earnings are not available in the 1935 census of Puerto Rico. Occupation, whether a person has access to a flush toilet at home, and the number of people per sleeping room are the only measures of socioeconomic status common to both censuses. Occupation is a useful control variable, when the question is raised of whether bilingualism has a net effect on standard of living. Access to a flush toilet at home and people per sleeping room are measures of standard of living. Access to a flush toilet has fewer problems as an indicator of standard of living than people per sleeping room. It is taken as the indicator of standard of living. The net reward for bilingualism is measured in probabilities of having access to a flush toilet at home. The indicator may seem a bit odd but it has face validity since the alternatives are less comfortable in daily use. It is closely associated with income in 1970 and there is no reason to think that would not be just as closely associated with income in 1935.

There are other sources of reinforcement for language learning besides attempting to reap a reward in the labor market: getting along with people who speak a different language, learning a new language in connection with a job, doing well in school where emphasis is put on learning a new language, being able to transact business with a
bureaucracy with a working language different from one's own. To show that it is the economic reward in standard of living rather than other sources of reinforcement that explains language learning, these other factors or their indicators have to be controlled for.

A number of hypotheses can be deduced from the economic reward theory of the diffusion of bilingualism and tested. The first hypothesis is that there should be a net reward for bilingualism, since it has been on the increase in Puerto Rico. Angle (1976b) has shown that there is this kind of reward in 1970, but it remains to be seen whether it exists in the pooled 1935-1970 data. A second deduction from the economic reward theory of the diffusion of bilingualism is that if there is change in the occupational structure in the direction of creating more higher paying, higher prestige jobs, and increasing the size of the labor force, this change will explain all or a good part of the change in bilingualism, net of the effect of changes in social background variables which affect both bilingualism and occupational placement. An expansion of the labor force, particularly in white collar jobs, is likely to draw more people into considering upward social mobility as a real possibility and what they might do to give themselves a competitive advantage.

A third deduction from the economic reward theory is that these predicted relationships will be larger in magnitude for the younger half of the population under study. The reasoning for this hypothesis is that young
people have more of an opportunity to become bilingual, through schooling, travel, or service in the armed forces, than older people. Angle (1975) shows that the level of bilingualism of most cohorts becomes stationary after the cohort becomes older than 25. Younger adults are freer of family obligations, habits, current occupations, etc., and can make use of an advantage in occupational placement. The occupational distribution of the old is to some extent a vestige of conditions extant when they were younger and entered the labor market (cf. Carlsson and Karlsson, 1970). The occupational distribution of the young is much more likely to be a response to relatively current conditions in the labor market than is the occupational distribution of the old.
Puerto Rico as a Case Study

Puerto Rico could be an independent country if its electorate so chose. Similar societies such as Cuba or the Dominican Republic are. It presently is a United States "Commonwealth," which means in law and practice that it is considerably more autonomous than a U.S. state but is still dependent on the U.S. federal government for financial aid and many governmental services. It is subject to federal law as any of the states are, but does not elect representatives to the U.S. Congress, the legislature that makes those laws. It has a national identity (cf. Maldonado-Denis, 1972; Lewis, 1974). It is ethnically fairly homogeneous. Mainlanders represent only several percentage points of the population (U.S. Bureau of the Census, 1963:33) and many of these are transient (Hirsbrunner, 1971:26). Nearly every ethnic Puerto Rican who is not a return migrant from the mainland, and many who are, have Spanish as a mother-tongue. Several studies done for the San Juan Star, the English language daily newspaper, show that in 1970 and 1973 well over four-fifths of subscribers spoke Spanish at home (Stanford Klapper Associates, 1970, 1973). What is remarkable about this figure is that the readership of the San Juan Star includes mainlanders and ethnic Puerto Ricans who would be the most likely to shift to English. If four-fifths of this group speak Spanish at home, it is a virtual certainty that shift to English is negligible in the population as a whole. Puerto Rico can be considered a country and the diffusion of bilingualism in English there can be compared
to other cases of language diffusion in national societies discussed by Meillet (1928). Deutsch (1966), or Lieberson et al. (1974, 1975).

Like so many cases of language diffusion, the spread of English in Puerto Rico began with conquest. American troops occupied the island in 1898. After two years of military government, a colonial government was organized. Its head was a governor appointed by the U.S. President. A legislature with some powers was also organized. The colonial government moved quickly to set up a system of mass public education. One of the main goals of the public school curriculum was the teaching of English. This goal was promoted over others to an absurd degree (cf. Osuna, 1923:222). Thus, in rural areas where children went to school for a year or two, if at all, Puerto Rican teachers whose command of English was weak were required to use English as the language of instruction to young children who knew no English at all (cf. Muñiz Souffrant, 1950). Literacy in Spanish and the fundamentals of arithmetic could be achieved only by the subverting of regulations. As Puerto Ricans came to have more of a say in the public school curriculum, the use of English as the language of instruction was progressively de-emphasized until in 1948, the first Commissioner of Education appointed by a governor elected by the people of Puerto Rico abolished English as a language of instruction in public schools (Wagenheim, 1970:103).

American colonial officials maintained on the whole an aloof attitude toward Puerto Ricans. They usually
refused to speak Spanish and usually insisted on being spoken to in English (Lewis, 1963:121). There was little tolerance for excusing any lack of enthusiasm for learning or speaking English on the part of Puerto Ricans (Lewis, 1963:443). The first Puerto Rican Commissioner of Education, Jose Padin, was fired by Franklin Roosevelt for raising the grade level at which English was required as the language of instruction. His successor, Jose Gallardo, was sent a letter of appointment which left no room for doubt that the President expected him to strengthen the English language program in the public schools (Rodriguez Bou, 1966:162,163). Several years later a letter from Interior Secretary Harold Ickes, who functioned as the American "colonial secretary," recalled this expectation so crudely that Gallardo threatened to resign (Epstein, 1970:26). A visiting Senate sub-committee in February, 1943, engaged in abrasive exchanges with leading members of the island's educational establishment on the pace at which English was being learned in the public schools (U.S. Congress. Senate, 1943). The English program of the public schools was quite possibly the most unpopular feature of direct American colonial rule. One witness at the February 1943 hearings called the attempt to teach young children in a foreign language "the crime of America" (U.S. Congress. Senate, 1943:291).

The ratification of the current Puerto Rican constitution, which made Puerto Rico a "commonwealth," did not end the role of government in encouraging bilingualism. The
Commonwealth government is officially bilingual and is dependent on the U.S. federal government for funds and on the investments of American businesses to maintain the prosperity to keep the electorate content. The U.S. Congress has held aid to education funds ransom to a continuing English language program in the public schools (Epstein, 1970:144-146). The efforts of a Commissioner of Education in the 1960's to prevent private schools from teaching English much more effectively than the public schools were ended by threats from U.S. Congressmen (Lewis, 1963:221). Puerto Ricans are subject to the military draft. The U.S. army operates in English. The Popular Democratic Party, in the governorship for most of the time since elections for the position have been held is committed to teaching English in the public schools. The Statehood Republican Party, in the governorship for one term since 1952, attempted to accelerate the English language program in the schools as a means of preparing Puerto Rico for statehood.

Before the late 1940's there was little industry in Puerto Rico. A glance at table 1 shows that as late as 1940 a near majority of the labor force were in agriculture which in most cases meant a few weeks of hard, poorly paid work and long stretches of unemployment during the off season (cf. Steward, et al., 1956). Many of the industrial operatives were women who did needlework in factories or on a put-out basis at home (Ross, 1969:16-17). The sharp stratification of the society into a small aristocracy and a large mass of rural poor that had existed in the Spanish
colonial period continued on past 1898.

The integration of Puerto Rico into the American economy after 1898 brought changes. Sugar companies bought up large tracts of land and converted them to sugar production. Processing plants for sugar, roads, bridges, port facilities, water purification plants, sewers, and a railroad were built. The U.S. federal government built a number of military bases. This kind of economic activity probably raised the incomes of many Puerto Ricans somewhat, but it did not transform the social structure (Ross, 1969: chap. 1; Wells, 1969:92). Most of the income gains by individuals were diluted by population increase. Puerto Ricans in 1930 were very little better off than they had been in 1898 and their situation worsened in the Great Depression. In this essentially stagnant agricultural economy, only the small elite had any need of proficiency in speaking English (Cebollero, 1945:114-115). They needed it for a political career or dealing with Americans as businessmen (Scheele, 1956:28-29,418; Cochran, 1959:83).

Following World War II, the colonial government and later the Commonwealth of Puerto Rico provided subsidies in the form of tax holidays for manufacturing industries to locate in Puerto Rico. In the beginning many of these plants were labor intensive, little capital making ample use of cheap hand labor. However, the base of industrialization broadened during the 1950's to include such capital intensive industries as oil refineries and petrochemical plants. Most of this direct investment is from U.S. corporations
The Puerto Rican subsidiaries of these corporations are often managed in their top positions by mainland Americans at least until a suitable ethnic Puerto Rican is found (cf. Puerto Rico, 1966; Hirsbrunner, 1971:65). It is quite conceivable that these mainlanders might attempt to recruit bilinguals for their whole operation. In fact, there is evidence of a shortage of bilingual workers, at least in the late 1950's (Hancock, 1960:154). Knowles (1965:113) as well as Hancock note that many companies attempt to hire ethnic Puerto Ricans who have migrated to mainland cities since these tend to be bilingual. Angle (1976b) has shown that there is a reward for bilingualism per se in terms of advantageous occupational placement and earnings in the labor market in Puerto Rico, but that for the labor force taken as a whole, as opposed to specific occupations such as managers and executives, it appears that mainland-owned companies are no more likely to recruit and reward bilinguals than are Puerto Rican owned companies. Bilingualism may have become so institutionalized in business (cf. Lieberson, 1970: chap. 5) and as a marker of higher social class (Epstein, 1966:222) that like excess educational qualifications, it is preferred by employers for few specific reasons. In any event, there is ample evidence that the notion that ability to speak English pays off in the labor market has taken root (Angle, 1976a: chap. 6). Epstein (1966:63) thinks that a major factor in the rapid growth of a private school system has been the concern of
parents that their children receive a more rigorous instruction in English than is available in the public schools.

Nearly everyone in Puerto Rico speaks Spanish, so Puerto Rico is not like the various linguistics sections of the Austro-Hungarian Empire, each controlled in government and business by a single mother-tongue group ruthlessly engaged in forcing all other people to learn and shift to its language (Inglehart and Woodward, 1967:34,35). There have been elements of the language policies of eastern European governments in Puerto Rico, however. During the colonial period, the public schools were used for Americanization just the way the public schools of Hungary were used for Magyarization. White collar government employment in the colonial period, a very attractive career in Puerto Rico at the time, required bilingualism. Mainland colonials were overbearing in the matter of speaking English, but their numbers were always small. Just at the moment when mainland political control was relaxing, rapid industrial growth created systems of reward in private industry, controlled as it is largely by mainlanders, which could be used to reward bilingualism. It is apparent that they were used that way but also that mainland companies did not give preference to bilinguals for their bilingualism per se more than did Puerto Rican owned companies. However, although there appears to be little direct influence of mainland Americans on the rewards for bilingualism, there is a clear, close relationship between bilingualism and social stratification: the higher the social class, the more bilingualism. Thus,
although Puerto Rico is not a society divided between two mother-tongue groups, one elite, the other subordinate, the model that underlies so much thinking on language diffusion, there are institutionalized economic rewards for speaking English and reason to expect these to drive the process of the diffusion of bilingualism.

The Data

A 1:1000 sample was drawn of the lines of all the schedules of the 1935 Census of Puerto Rico stored in National Archives, Washington, D.C. Each line is either a person or a blank. No schedules were noticed missing. Sampling lines rather than people removed the problem of counting people. Each census schedule had a fixed number of lines making counting between sampled lines simple. The expected number of people sampled in a 1:1000 random sample of lines is 1/1000 of the total population of Puerto Rico. One one-thousandth of the population of Puerto Rico in 1935 according to the publication based on the 1935 Special Census is 1,724 (Puerto Rico Reconstruction Administration, 1938: 35). There are 1,757 people in the sample that was drawn. This figure is very close to expectation. Proportions calculated from the sample correspond very closely to proportions calculated from published census statistics. For example, the percent bilingual of the population 10 years old and older in the sample data is 22.57%. The percent of the population 10 years old and older who are bilingual in the published census figures is 22.93% (Puerto Rico Reconstruction Administration, 1938:32). The few people not born in
Puerto Rico and those 19 and younger are screened out. Those not born in Puerto Rico are removed because it is the part of the population who are labor force age who are of interest. Cases with missing data on any variable were dropped. 795 cases are left in the 1935 sample. Most of the difference between this number and 1,757 is due to the screening out of young people.

A 1:1000 sample of the 1970 Census of Population and Housing in Puerto Rico was taken by sub-sampling a public use sample of the "state characteristics" file for Puerto Rico, produced by the U.S. Bureau of the Census (1972a). This sample has 2,707 cases. Only people born in Puerto Rico with at least one Puerto Rican born parent are kept to insure that everyone in the sample speaks Spanish natively. Those 19 years of age and under are also screened out. 1278 cases are left in the 1970 sample. There are 2,073 people in the pooled 1935-1970 sample.

Variables which can be considered for a pooled sample of 1935 and 1970 data are limited to those the two censuses have in common. The questions asked in 1935 were few in number. Access to a flush toilet at home is used as a measure of standard of living. Literacy is an index to education. It measures a minimal level of education but since 42% of the sample of the 1935 adult population were illiterate, it is an appropriate measure. Occupation is measured by a trichotomy with categories 'white collar,' 'blue collar,' and 'no occupation.' 1970 Census occupation codes (U.S. Bureau of the Census, 1972b:100-110) were used to code
occupations in the 1935 sample, so there is no compatibility problem combining the two data sets. Similarly, the 1970 definition of what constitutes an urban area was applied to the 1935 census. Since towns back then tended to be well defined geographically, the definition is easy to apply. There are also dichotomous variables for age (36 years old or younger/older), sex, ability to speak English (yes/no), (the same question was asked in both censuses), and time (1935/1970). See table 2.

When the question is whether change in the occupational structure affects bilingualism, literacy and urban residence are the main controls for exposure to speakers of English and incidental re-inforcement for speaking English. Literacy and urban residence are closely associated in 1970 with education, being a veteran, or a return migrant from the mainland. Hernandez (1967; Appendix) documents the fact that in 1960 return migrants, particularly the bilinguals tended to stay in cities. Controlling for urban residence also controls for exposure to mainland tourists and businessmen who are concentrated in urban areas, particularly San Juan. Sex and age are important factors affecting how people enter the labor force. They are also related to bilingualism. See table 2. Men and the young tend to be more bilingual than women and the old. Literacy, urban residence, sex, and age as well as occupation are controlled for when the question is whether in 1935 or 1970 net bilingualism fetches a price in a higher standard of living, measured by whether a person has access to a flush toilet at home.
The Reward to Bilingualism per se

The first question to be addressed to the data is whether bilingualism per se has a reward in a higher standard of living, that is, whether net of factors which might be expected to affect standard of living, such as literacy, occupation, age, urban residence, bilingualism itself has an effect. Standard of living is operationalized as having a flush toilet. This question is answered by comparing two log-linear models.

Log-linear models are sets of contingency table marginals from which expected frequencies for table cells can be generated (cf. Goodman, 1972, 1973). The usual way to designate a marginal, referred to as a 'term,' is to assign a letter to each variable, such as: A = Access to Flush Toilet, B = Sex, C = Age, E = Literacy, G = Bilingualism, H = Occupation, J = Urban/Rural Residence, and T = Time. A clump of letters together such as BCE indicates that the three variables Sex, Age, and Literacy and all their interactions are being used to generate expected frequencies. The particular marginals used to generate expected frequencies correspond to hypotheses about what explains variation in the frequencies. The way to test for the significance of a particular term is to generate expecteds without it, then with it, and compare the chi-square measures of the fit to the actual data. The maximum likelihood chi-square is the measure of fit. The difference between the chi-square of the first model and the chi-square of the second
model is itself a chi-square statistic distributed with the difference of the degrees of freedom of the two models. This chi-square of the difference is used to test the statistical significance of the term under examination.

To test whether net bilingualism has an effect on standard of living, operationalized as whether a person has access to a flush toilet at home, a number of marginal effects need to be "fixed" or held invariant, because they are not themselves of interest. The joint distribution of the predictor variables of Sex, Age, Literacy, Bilingualism, Occupation, Urban/Rural Residence and Time is of no interest. Thus the BCEGHJT term is inserted into the model. The fact that access to a flush toilet depends on a number of factors other than bilingualism needs to be controlled for. Thus an ABCEHJ term is inserted. Time is not included in this last term since it is not, properly speaking, an explanatory variable.

The stage is now set to test for the Bilingualism-Flush Toilet term. Bilingualism, if it has an effect, will have it net of Sex, Age, Literacy, Occupation, and Urban/Rural Residence. Model 1 in table 3 does not have the Bilingualism-Flush Toilet term. It is the "baseline" model. Model 2 does. The chi-square of model 2 is compared with that of model 1. If there is a significant reduction for the difference in the degrees of freedom, then the Bilingualism-Flush Toilet term is statistically significant. In fact it is. See table 3. The Bilingualism-Flush Toilet term has a chi-square of 108 with 1 degree of freedom, which
is significant far beyond the .001 level, and an adjusted partial r-square, a measure of how much the term improves the fit of the baseline model, of .22, indicating that it has a fairly large effect. The lambda coefficient between the two variables shows that the categories are associated in the expected way, that is, that being bilingual is associated with having access to a flush toilet, and, by implication, other material advantages. This finding confirms for the pooled 1935-1970 data what Angle (1976b) found to be the case with more stringent controls for 1970 data alone.

First Differencing with Log-linear Models: Does Change in Employment and Occupation Explain Change in Bilingualism?

Bilingualism affects standard of living, but does change in the occupational structure in the direction of more employment, particularly white collar employment, affect bilingualism? Bilingualism presumably leads to a higher standard of living through placement into higher paying occupations and, among those in similar occupations, in higher earnings. Expansion of employment, particularly white collar employment, generates more opportunities to reward bilinguals. If the economic reward theory of the diffusion of bilingualism is true an expansion of employment and white collar employment ought to be related to an increase in the likelihood of people being bilingual. It is not enough to ask if changes in the occupational structure are related to an increase in bilingualism. What is of interest is whether holding cross-sectional effects fixed, a change in occupation is related to a change in
bilingualism—that is, whether, as far as can be ascertained, a net change in occupation caused a net change in bilingualism.

How do we know it is occupation causing bilingualism and not vice versa? Nearly all writers on the subject of language diffusion assume economic determinism. In Puerto Rico, it is known that the inception of the development program could not have been in any sense caused by a change in bilingualism. The development program was undertaken because Puerto Ricans wanted a higher standard of living and the U.S. government was not enthusiastic about giving up its military bases there or the Puerto Rican market. None of which is related to how many people learned English during the period of industrialization.

A number of different effects have to be held invariant in order to see whether there is variation in the terms of interest, change in the occupational distribution and change in bilingualism, expressions similar to the difference in a variable between two consecutive time points, a first difference. Both Occupation and Bilingualism are taken as dependent on other variables. Occupation is seen as dependent on Sex, Age, and Literacy, and their mutual interactions. All three variables affect occupational placement. Occupation is not taken as dependent on Urban/Rural Residence since both are a result of industrialization. Bilingualism is taken as dependent on Sex, Age, Literacy, and Urban/Rural Residence. A Bilingualism–Occupation interaction term is added to the model to hold this relationship
invariant. White collar people are more bilingual than blue collar people, who, in turn, are more bilingual than people with no occupation. Since this relationship is not at issue, it is held invariant.

The terms which hold these given interactions and the joint distribution of predictor variables constant are: BCEJT BCEF BCEGJ GH. This model, model #3 in table 3, is the baseline model. Model #4 in table 3 tests whether there is a significant change in the occupational distribution between 1935 and 1970. There is. The lambdas of Occupation with Time(1970) are: -.10 for 'no occupation,' -.10 for 'blue collar' and +.20 for 'white collar,' indicating that unemployment and blue collar employment decreased while white collar employment increased. Model #5 in table 3 tests whether there is a significant change in net bilingualism between 1935 and 1970. There is a significant change in net bilingualism. Its lambda with Time(1970), .23, indicates that net bilingualism increased between 1935 and 1970.

Model #6 shows that adding an Occupation-Time term to model #5 does not significantly change its chi-square, which means that the Occupation-Time term is not statistically independent of the Bilingualism-Time term. Thus, bringing in causal imagery, a change in the occupational distribution can be said to explain part of the change in bilingualism, 23% of the change. This finding was predicted by the economic reward theory of the diffusion of bilingualism and confirms it. However, only part of the change in
net bilingualism, that is, bilingualism not explained by the other predictor variables, is accounted for. Three-quarters of it is still unexplained.

Larger Effects for Young than for Old

A third deduction from the economic reward theory of the diffusion of bilingualism is that the measured reward for bilingualism will be greater for the younger half of the adult population than for the older half, who, it is theorized, are more settled in occupations and more removed from the labor market. It is also predicted that change in the occupational distribution will have a greater impact over time in change in net bilingualism for the young than for the old, again for the same reason. The population under study was split into two groups, one 20 to 36 years of age, the other 37 and older. The appropriate log-linear models were fitted to these data but the hypothesis that the predicted effects would be larger for the younger half of the population was not born out. Perhaps the reason that this third hypothesis is not confirmed is that the rapidity of economic development between 1935 and 1970 was such that the premise on which the hypothesis is based, that the older half of the adult population do not change occupations, is a very poor approximation of reality.

Conclusions

Economic development, however measured, is related to a rising level of bilingualism in English among native speakers of Spanish in Puerto Rico. Every indicator of economic development is associated with bilingualism in the
expected direction: the more development, the more bilingualism. Lieberson et al.'s (1974,1975) finding that economic development is not related over-time to a change in mother-tongue diversity in nations, and by implication, to language shift, has not been replicated in Puerto Rico for the diffusion of bilingualism. On the other hand, Puerto Rico does illustrate the point that in the short run the relationship between economic development and language shift does not fluctuate in lockstep. Puerto Rico has industrialized almost without language shift.

This paper has presented evidence that confirms the economic reward theory of the diffusion of bilingualism. The economic reward theory states that it is the reality and the perception of a labor market reward for bilingualism that is the main determinant of the effort made to become bilingual in the population as a whole, and consequently, of the rate at which bilingualism increases in the population. Controls have to be applied for situations where other kinds of reinforcements may be at work. Comparisons of log-linear models fitted to contingency tables containing a pooled sample of the 1935-1970 censuses of Puerto Rico are used to establish whether several effects, predicted by the economic reward theory, do exist.

First, is the question of whether net of other factors which affect both standard of living and bilingualism, bilinguals do enjoy a higher standard of living than Spanish monolinguals. They do. Second, is the question of whether, over-time, improvement in the occupational structure, that
is, higher levels of employment, particularly white collar employment, are related to an increase in bilingualism. More employment, especially in white collar occupations, offers more opportunities to reward bilinguals and ought to be a cause of an increase in bilingualism. Changes in other variables which affect bilingualism have to be controlled for. In fact, change in the occupational structure, net of several controls, does explain part of the change in bilingualism, net of quite a few controls.

The hypothesis that the predicted effects of the economic reward theory are larger for younger adults because they are, or have been recently, more oriented to the labor market than older adults was not confirmed. Perhaps the premise that younger adults are more oriented to the labor market than older adults was not valid in Puerto Rico between 1935 and 1970.
Table 1. Occupational Distribution of the Employed Labor Force, Puerto Rico (in percentages)\(^a\)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>professional, technical, and kindred managers</td>
<td>2.5</td>
<td>3.1</td>
<td>4.9</td>
<td>7.8</td>
<td>12.0</td>
</tr>
<tr>
<td>sales workers</td>
<td>2.9</td>
<td>4.8</td>
<td>6.0</td>
<td>7.7</td>
<td>7.2</td>
</tr>
<tr>
<td>clericals</td>
<td>3.9</td>
<td>5.0</td>
<td>5.3</td>
<td>6.3</td>
<td>7.5</td>
</tr>
<tr>
<td>craftsmen, foremen, and minor operatives</td>
<td>3.1</td>
<td>3.1</td>
<td>5.0</td>
<td>7.7</td>
<td>12.3</td>
</tr>
<tr>
<td>laborers</td>
<td>6.4</td>
<td>5.5</td>
<td>7.6</td>
<td>11.5</td>
<td>15.0</td>
</tr>
<tr>
<td>farmers</td>
<td>18.6</td>
<td>18.0</td>
<td>16.7</td>
<td>17.5</td>
<td>19.4</td>
</tr>
<tr>
<td>farm laborers</td>
<td>7.8</td>
<td>5.2</td>
<td>5.6</td>
<td>6.5</td>
<td>6.6</td>
</tr>
<tr>
<td>service workers</td>
<td>8.4</td>
<td>9.4</td>
<td>6.5</td>
<td>3.3</td>
<td>1.4</td>
</tr>
<tr>
<td>service workers</td>
<td>35.5</td>
<td>35.2</td>
<td>31.2</td>
<td>20.2</td>
<td>6.1</td>
</tr>
<tr>
<td>service workers</td>
<td>4.1</td>
<td>3.0</td>
<td>5.4</td>
<td>8.0</td>
<td>11.0</td>
</tr>
<tr>
<td>private household workers</td>
<td>6.6</td>
<td>7.8</td>
<td>5.9</td>
<td>3.4</td>
<td>1.6</td>
</tr>
</tbody>
</table>

Base \((512)^b(506,942)^c(554,691)^c(544,264)(591,619)\)

\(\text{aSources:}\)

1935: 1:1000 sample of 1935 census of population
1950: 
1970: 

\(\text{bSample data, anyone with an occupation is included.}\)

\(\text{ctotal of those reporting an occupation.}\)
<table>
<thead>
<tr>
<th></th>
<th>1935</th>
<th>1970</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Distrib-</td>
<td>% Bilingual</td>
</tr>
<tr>
<td></td>
<td>in English</td>
<td>in English</td>
</tr>
<tr>
<td>Bilingualism in English</td>
<td></td>
<td></td>
</tr>
<tr>
<td>yes</td>
<td>14.5</td>
<td>*</td>
</tr>
<tr>
<td>no</td>
<td>85.5</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Literacy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>yes</td>
<td>58.1</td>
<td>24.9</td>
</tr>
<tr>
<td>no</td>
<td>41.9</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 to 36</td>
<td>56.1</td>
<td>21.1</td>
</tr>
<tr>
<td>37 and older</td>
<td>43.9</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>50.2</td>
<td>18.5</td>
</tr>
<tr>
<td>female</td>
<td>49.8</td>
<td>10.4</td>
</tr>
<tr>
<td></td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Occupation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>white collar</td>
<td>7.0</td>
<td>51.8</td>
</tr>
<tr>
<td>blue collar</td>
<td>48.3</td>
<td>9.9</td>
</tr>
<tr>
<td>no occupation</td>
<td>44.7</td>
<td>6.8</td>
</tr>
<tr>
<td></td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Urban Residence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>urban</td>
<td>31.8</td>
<td>24.9</td>
</tr>
<tr>
<td>rural</td>
<td>68.2</td>
<td>9.6</td>
</tr>
<tr>
<td></td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Access to Flush Toilet</td>
<td></td>
<td></td>
</tr>
<tr>
<td>yes</td>
<td>13.8</td>
<td>42.7</td>
</tr>
<tr>
<td>no</td>
<td>86.2</td>
<td>10.0</td>
</tr>
<tr>
<td></td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1935</td>
<td>38.4</td>
<td>14.5</td>
</tr>
<tr>
<td>1970</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Source is a pooled 1:1000 sample of 1935 and 1970 censuses of population in Puerto Rico. N of sample is 2,073. Allocated data are used in 1970 sample.

b indicates table entry does not exist.
Table 3. Comparisons of Log-linear Models to Test the Significance of Particular Terms, Pooled 1935, 1970 Census Data, Puerto Rico N=2,073<sup>a</sup>

<table>
<thead>
<tr>
<th>variables</th>
<th>code letter</th>
<th>number of categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access to Flush Toilet at Home</td>
<td>A</td>
<td>2</td>
</tr>
<tr>
<td>Sex</td>
<td>B</td>
<td>2</td>
</tr>
<tr>
<td>Age</td>
<td>C</td>
<td>2</td>
</tr>
<tr>
<td>Literacy</td>
<td>E</td>
<td>2</td>
</tr>
<tr>
<td>Bilingualism</td>
<td>G</td>
<td>2</td>
</tr>
<tr>
<td>Occupation</td>
<td>H</td>
<td>3</td>
</tr>
<tr>
<td>Urban/Rural Residence</td>
<td>J</td>
<td>2</td>
</tr>
<tr>
<td>Time</td>
<td>T</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>models (fitted marginals)</th>
<th>x&lt;sup&gt;2&lt;/sup&gt;</th>
<th>d.f.</th>
<th>p</th>
<th>adjusted partial r-square</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. BCEGHJT ABCEHJ</td>
<td>478</td>
<td>144</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>2. BCEGHJT ABCEHJ AG</td>
<td>370</td>
<td>143</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>Bilingualism-Flush Toilet term, AG, (model #2 vs. model #1)</td>
<td>108</td>
<td>1</td>
<td>&lt;.001</td>
<td>.221</td>
</tr>
<tr>
<td>3. BCEJT BCEH BCEGJ GH</td>
<td>243</td>
<td>126</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>4. BCEJT BCEH BCEGJ GH HT</td>
<td>225</td>
<td>124</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>5. BCEJT BCEH BCEGJ GH GT</td>
<td>183</td>
<td>125</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>6. BCEJT BCEH BCEGJ GH GT HT</td>
<td>179</td>
<td>123</td>
<td>&lt;.001</td>
<td></td>
</tr>
</tbody>
</table>

Evidence that Occupation-Time term, HT, is not independent of Bilingualism-Time term, GT, (model #6 vs. model #5) 4 2 >.5 .006

<sup>a</sup>Models are fitted to contingency tables defined by the categories of only the variables specified in the model.

<sup>b</sup>maximum likelihood chi-square

<sup>c</sup>degrees of freedom

<sup>d</sup>significance level of chi-square

<sup>e</sup>adjusted partial r-square = (X<sub>0</sub>/df<sub>0</sub>-X<sub>1</sub>/df<sub>1</sub>)/(X<sub>0</sub>/df<sub>0</sub>), where X<sub>0</sub> is chi-square of model with the term and X<sub>1</sub> is the chi-square of the model with the term, and so on for the degrees of freedom, df.
REFERENCES

Angle, John


Brun, Auguste


Brunot, Ferdinand


Carlsson, Gosta


Carlsson, Gosta and Katarina Karlsson


Cochran, Thomas

Deutsch, Karl

Epstein, Erwin

Etzioni, Amitai

Fishman, Joshua

Goodman, Leo

Greenberg, Joseph

Hancock, Ralph

Haugen, Einar
Hernandez Alvarez, Jose

Hertzler, Joyce

Hirsbrunner, Jack

Inglehart, Ronald and Margaret Woodward

Jespersen, Otto

Knowles, William

Leach, Edmund

Lewis, Gordon

Lieberson, Stanley
Lieberson, Stanley and Timothy Curry

Lieberson, Stanley and Lynn Hansen

Lieberson, Stanley, Guy Dalto, and Mary Ellen Johnston

Maldonado-Denis, Manuel

Meillet, Antoine

Muñiz Souffrant, Luis

Novicow, Jacques

Osuna, John Joseph
1923 Education in Puerto Rico. New York: Teacher's College.

Puerto Rico, Department of Labor
1966 Establecimientos Manufactureros en Puerto Rico. San Juan: Department of Labor.

Puerto Rico Reconstruction Administration

Peyre, Henri
Pool, Jonathan

Rodriguez Bou, Ismael

Ross, Daniel

Scheele, Raymond

Stanford Klapper Associates
1970    Who Reads the San Juan Star?  San Juan: The San Juan Star.
1973    Focus Puerto Rico.  San Juan: The San Juan Star.

Steward, Julian, et al.

U.S. Bureau of the Census
U.S. Congress, Senate


Wagenheim, Kal


Wells, Henry
