Distinguishing Between Proficiency, Choice, and Attitudes in Questions about Language for Bilinguals

Working Papers Series, No.32

by
Kenji Hakuta
School of Education, Stanford University

Summer, 1991

The purpose of the SCCR Working Paper Series is to publish works that significantly advance our knowledge about Chicanos and other Latinos. We invite your statements and critique. Please address your remarks to the author.

Stanford Center for Chicano Research, Cypress Hall "E" Wing, Stanford, CA 94309

DISTINGUISHING BETWEEN

PROFICIENCY, CHOICE, AND ATTITUDES

IN QUESTIONS ABOUT LANGUAGE FOR BILINGUALS

Kenji Hakuta Stanford University

Sociologists who study bilingual communities have long noted that stable and balanced bilingualism over generations is rarely observed (Fishman, 1966). The bilingualism of language minority communities in the United States is no exception, characterized by a rapid shift into monolingual usage of English within two to three generations, (Fishman, 1966; Veltman, 1983), even in the case of Spanish (Veltman, 1988). Indeed, Veltman (1988) concludes the following with respect to the status of Spanish in the United States: "Given the age structure of the immigrant population, more than three-fourths of any given group of immigrants will come to speak English on a regular basis after approximately 15 years of residence in the United States. more important, approximately 70 percent of the youngest immigrants and 40 percent of those aged 10-14 at time of arrival will make English their usual, personal language. As a result, they will give birth to children of English, not Spanish, mother tonque" (P. 45).

Variability in this shift from ethnic language to mainstream language is assumed to be correlated with acculturation, the extent of the correlation depending upon the theory of the relationship between language and culture to which one subscribes. Specific to the purposes of this volume, the phenomenon of language shift bears close scrutiny not only because of the intrinsic value of getting a better understanding of language, communication and acculturation of Puerto Rican families, but also because of the measurement issues involved. In analyses of the Hispanic Health and Nutrition Examination Survey (HHANES), for example, the language and acculturation variables have proven to be significant predictors of health preventive behavior (Solis, Marks, Garcia & Shelton, 1990), alcohol use (Marks, Garcia & Solis, 1990), marijuana and cocaine use (Amaro, Whitaker, Coffman & Heeren, 1990), and even use of curanderos (Higginbotham, Trevino & Ray, 1990). A closer look at the questions asked in the acculturation scale in light of our understanding of language shift may help sharpen the questions for future inquiry.

This paper presents data on language shift obtained in the course of research conducted in a Puerto Rican bilingual education program in New Haven, Connecticut. I will also make reference to data collected in Watsonville, California from students of Mexican background. The value of these data lie in the fact that they contain actual measures of English and Spanish

proficiency, whereas most of the quantitative research on language shift, such as the pioneering works of Joshua Fishman and Calvin Veltman, use self-reported language data collected by the Census Bureau and the High School and Beyond survey gathered by the National Opinion Research Center. Naturally, the latter data sets are infinitely superior with respect to sampling and generating population estimates, but they are possibly subject to distortion in self-report.

Methodological Issues in the Measurement of Language Shift

Language shift has at least three components that should be measured separately: (1) an individual's actual proficiency in the two languages, (2) an individual's choice to use differential amounts of the languages (in different discourse settings) given threshold proficiency in the languages, and (3) an individual's personal identification with the cultures associated with the These components are in principle separable (i.e., there may be an individual with high proficiency in both English and Spanish who chooses to use mostly English, but maintains an identity that is primarily Puerto Rican), but in reality, they are probably related. For example, Veltman focuses on language choice over language proficiency, not just because the census bureau questions have tended to ask the question on usual language practice (e.g., "What language does (this person) usually speak?" in the Survey of Income and Education, 1976), but also because he considers it a logical outcome that if a language is not usually spoken in the home, the children will not develop proficiency in it.

In addition to distinguishing between these components of language shift, it is important to ask whether the data are based on self-report or on direct observation. For example, the High School and Beyond survey asked "How well do you speak that language?" with response choices "very well," "pretty well," "not very well," and "not at all." How accurate would this self-report be when compared with direct observation of proficiency in the language? Obviously, self-reported data are the easiest to obtain, but they sacrifice objectivity; however, in some cases, direct observation may be extremely difficult or impractical, such as in the case of language identification.

The rough measurement model in the study of language shift, then, can be thought of as a 2x3 table as follows:

Type of Data

Observation

Self-Report

Aspect of Bilinqualism		
Language Proficiency	1	2
Language Choice	3	4
Language Identification	5	. 6

In an ideal situation, one would look at the correlations between 2, 4, and 6 based on actual observations. However, we are often dependent on self-report and other indirect means of inference.

Puerto Rican Children, New Haven, Connecticut

R

The data on Puerto Rican children reported here were collected in the fall of 1980 in a bilingual education program in New Haven. The program uses a transitional rather than a maintenance model of bilingual education. Students are placed in the program on the basis of the home language reported by the parents, and exited to English-only classes on the basis of their English proficiency as determined by a standardized language test and teacher assessment (see Hakuta, Ferdman & Diaz, 1987 for more information on bilingualism in the New Haven Puerto Rican community). The primary purpose of the study was to test the hypothesis that degree of bilingualism would be related positively to performance on various measures of cognitive flexibility (Hakuta & Diaz, 1985; Hakuta, 1987). Because the study was not originally designed to investigate language shift, the data have severe limitations, but they are nevertheless informative because language proficiency and language choice measures were obtained.

The <u>S</u>s reported in this analysis were 226 children in Kindergarten (N=54, Mean Age=5.23, SD=.43), Grade 1 (N=60, Mean Age=6.55, SD=.55), Grade 4 (N=63, Mean Age=10.02, SD=.78), and Grade 5 (N=49, Mean Age=10.76 SD=.65). In order to obtain

background information on the \underline{S} s, a brief questionnaire was distributed to the parents, on which the following questions were asked:

- What language does your child use with you and the other adults in the household? (Only Spanish, Mostly Spanish, Both English and Spanish, Mostly English, Only English) [¿Qué idioma usa su hijo/a con usted y con otras personas adultas en la casa? (español solamente, mayormente español, igual español que inglés, mayormente inglés, inglés solamente)]
- What language does your child use with his or her brothers and sisters? [Qué idioma usa su hijo/a con sus hermanos y hermanas?]
- What language do the adults in your household use with each other? [¿Qué idioma usan las personas adultas en la casa?] In addition, parents were asked to indicate the number of years that the family had lived on the mainland (we did not ask for the entire history of migration, although clearly such data would be highly relevant).

In order to assess the relative bilingual proficiency of the Ss, the Peabody Picture Vocabulary Test was administered in both languages by separate testers (Dunn, 1965 for English, and Wiener, Simmond & Weiss, 1978 for Spanish). The PPVT measures receptive vocabulary; uncertainty about the validity of standardized norms in either language led us to use the raw scores as estimates of language proficiency.

With reference to the measurement model described above, we have parental report of language used (Cell 3, in this case as reported by the parents), and actual observations of language proficiency in English and Spanish (Cell 2). We also have some background information, including years of residence on the mainland.

1

Figure 1 displays Spanish and English proficiency as a function of years of residence on the mainland, reported separately by grade level (K, 1, 4, 5). The smoothing function fitted on the data in this and all other figures used in this

INSERT FIGURE 1 ABOUT HERE

paper applies a locally weighted regression method (Cleveland, 1981) available on SYGRAPH (Wilkinson, 1987). The first notable feature of these figures is the fact that there is a clear subtractive relationship between English and Spanish proficiency. Years of residence on the mainland is associated with higher English proficiency and lower Spanish proficiency. The second notable feature is that there is a sharp break at about 10 years of residence. The relationship between years of residence and proficiency is evident in those who have been on the mainland for about 10 years or less, but if one considers only those whose families have been here for over 10 years (i.e., in the case of this sample, this would mean that on average most of them would

have been born on the mainland), the trends in both English and Spanish proficiency are flat.

Figure 2 displays the data on language choice as a function of years of residence on the mainland. The data were combined for the four cohort groups since they were obtained from parents, and cohort differences were not evident. Several features are

INSERT FIGURE 2 ABOUT HERE

notable. First, the language used by the adults (Panel A) who have been on the mainland for under about 15 years is mostly Spanish, but there is encroachment of English among those who have been on the mainland for over 15 years. Nevertheless, there are few values that use "mostly" or "only" English. Second, the language reportedly used among siblings (Panel C) shifts much earlier (within 10 years of residence) to a value of approximately somewhere between "mostly Spanish" and "both equally", and then remains at about that level. Among the siblings, there are considerably more individuals who reportedly use "mostly" or "only" English. The adult language practice with children (Panel B) remains somewhere between these two.

In order to understand better the relationship between language proficiency and language choice (reported by parents), a measure combining the proficiency scores of English and Spanish was constructed, by subtracting Spanish from English. Recall

that the choice scales run from 1=only Spanish to 3=both equally to 5=only English. Thus, to the extent that proficiency and choice are related, the correlation between the English minus Spanish scale and the choice measures should be a positive one. Figure 3(a) displays the data for the Ss from Grades K-1, and Figure 3(b) for the Ss from Grades 4-5. The magnitude of the linear correlations are r=.32 and r=.23 respectively. We must consider this to be both good news and bad news. The good news is that they are positively related, but the bad news is that there is a large amount of variance (over 90 percent) still left to be explained.

INSERT FIGURE 3 ABOUT HERE

There are probably at least two factors that account for this low correlation. The first is that the language choice scale is very crude -- one question that uses a five-point scale. More questions would undoubtedly increase the reliability. The second fact is that this sample of bilingual children is quite selective, i.e., they are students who are in a bilingual education program, and by definition, they are dominant in English. If the entire Puerto Rican community in New Haven had been included, there would have been considerably higher variances associated with both the language proficiency and

language choice scores. These factors are controllable by sharpening the instruments and the sampling procedures.

Mexican Background High School Students, Watsonville, California

In our Watsonville data (Hakuta & D'Andrea, in press), we examined the language proficiency, choice, and identification of over 300 high school students of Mexican descent from different generational backgrounds. With reference to the measurement model above, we directly assessed English and Spanish proficiency by one of three measures (vocabulary, grammatical knowledge, and a measure of whole language knowledge called "cloze"), i.e., Cell 2. In addition, to study the validity of self-reported ability in the languages, we asked them to rate their own ability in reading, writing, and speaking the two languages (Cell 1) through the following questions:

- How well do you speak and understand English/Spanish? (not at all, hardly at all, not so good, so-so, good, well, but not perfect, perfect)
- How well do you read in English/Spanish?
- How well do you write in English/Spanish?

To assess language choice, we asked for a self-report (Cell 3) of what language they used with various parties including adults at home, their siblings, and their peers at school, using the following questions:

- The adults in my home usually speak with each other in...(only Spanish, mostly Spanish, both languages equally, mostly English, only English, not applicable)
- I speak with my father in...
- I speak with my mother in ...
- I speak with my older brothers and sisters in...
- · I speak with my younger brothers and sisters in...
- At school I speak with my friends in...
- Outside of school I speak with my friends in...
- My friends usually speak with me in...

Finally, we obtained their attitudes towards maintenance of bilingualism and Mexican culture by Likert scale responses to the following questions:

- Knowing how to speak Spanish is important to understand a person's family history. (strongly disagree/ strongly agree)
- How important is it for you to know Spanish well? (not at all / very much)
- A person who knows Spanish, in addition to English, has more chances to express his or her feelings. (strongly disagree/ strongly agree)
- How important is it for you to know both English and Spanish
 well? (not at all / very much)
- Using Spanish allows a person to feel good about him or herself. (strongly disagree/ strongly agree)

 People who know Spanish should use it daily, especially at home. (strongly disagree/ strongly agree)

8

- A person often needs to use Spanish for daily communication.
 (strongly disagree/ strongly agree)
- It's O.K. if a person grows up speaking Spanish, and later forgets it. (strongly agree/ strongly disagree)

These attitude questions are clearly an aspect of identification, and are probably best classified as self-report on this aspect of bilingualism (Cell 5).

The first major finding with respect to generational background was that Spanish proficiency appeared to be maintained up through the second generation (the group who were born in the United States, but whose parents were born in Mexico), but drops sharply for the third generation. This dip in proficiency was associated with the reported home language choice of the adults, very much along the lines predicted by Veltman (1988). Language shift, as defined by a drop in Spanish proficiency, appears discontinuous. On the other hand, when we looked at language choice across the different generational backgrounds, the shift appeared incremental and synchronous to the acquisition of English.

With regard to the interrelationship between the different aspects of language, language choice was predicted by language proficiency in both languages and by language attitude; however, neither proficiency in Spanish nor English was predicted by

language attitude. Thus, we concluded that language choice is a socially mediated variable having to do with ethnic identification, but it is constrained by the limits of proficiency in the two languages (i.e., if one does not have proficiency in either language, the choices are not even available).

8

Some of the interesting properties of the time course of shift in language proficiency and choice can be observed in Figure 4 (not reported in Hakuta & D'Andrea, in press), which plots these variables as a function of the age at which they started learning English. (We obtained an estimate of this age from the Ss.) A quick look at the figures shows that age of exposure to English is a powerful predictor for all of the variables, although with differently shaped functions.

INSERT FIGURE 4 ABOUT HERE

Early exposure to English before the age of 5 shows negative relationship with Spanish proficiency (Panel A). It is likely that those who were exposed before age 5 never fully acquired proficiency, especially in light of the reported language choice in favor of English of the adults at home (Panel C). The English slope (Panel B) requires a more careful interpretation, because

those with exposure after age 10 are still in the process of second language acquisition.

₹

Figure 4c shows the pattern for adult language choice, and suggests that the low Spanish proficiency in the early exposure group is most likely due to the adult language practice. Indeed, as can be seen in Figure 5, when English and Spanish proficiency are plotted as a function of the language choice of adults in the home (selecting only those Ss who were exposed to English before age 6), it is evident that adult language practice maps directly onto Spanish proficiency in their children. Equally interesting from the viewpoint of the debate about the role of home language practice in accelerating the children's English proficiency is the fact that adult language practice evidently has no effect on English proficiency.

INSERT FIGURE 5 ABOUT HERE

Returning to Figure 4, the data on language choice for siblings and peers (Panels D, E) suggest a linear trend towards English with earlier exposure. Notable (though predictable) is the discrepancy in language choice with that for the adults (Panel C), with a stronger gravitation towards English when the choosing language with siblings and peers — this underscores the

importance of specifying the interlocutors in asking questions about language choice.

To parallel the analysis of the relationship between reported language choice and observed language proficiency that was reported for the Puerto Rican sample, Spanish proficiency scores were subtracted from English proficiency scores, and used to predict self-reported language choice with siblings. The results can be found in Figure 6, which also incorporates some data collected from 7th and 8th grade students (N=95) from two feeder schools in the course of a pilot work (the language proficiency measure was a test of productive vocabulary, and questions were also asked about language use with siblings and friends). The linear correlations are N=39 for the Grade 7-8 sample, and N=62 for the Grades 9-12 sample. In the discussion of the results from the Puerto Rican sample, I speculated that the low correlation might be due to the coarseness of measurement as well as restriction of range. The Watsonville sample is

INSERT FIGURE 6 ABOUT HERE

reasonably representative of the Latino population in the community (see Hakuta & D'Andrea for discussion), and in the case of the high school study, two questions were asked about language choice with siblings. Thus, both factors may have been at work,

although there are other significant differences between the samples.

Finally, the study revealed an important methodological finding about the relationship between self-reported and observed language proficiency (the relationship between Cells 1 and 2 in the measurement model above). Regression analyses were conducting using the following model with simultaneous entry of the predictors:

Self-reported Spanish (or English) proficiency =

Constant +

Spanish (or English) Proficiency +

Maintenance Attitude +

Subtractive Attitude +

Pragmatic Attitude

(the latter two attitudinal orientations were also extracted from the attitude questionnaire using factor analytic techniques).

The results indicated a significant contribution of the attitudinal variables independent of actual measured proficiency, especially for Spanish. Thus, for Spanish self-reported proficiency, the b's were as follows: Constant (-1.889), Spanish Proficiency (.53), Maintenance Attitude (.407), Subtractive Attitude (-.135), Pragmatic Attitude (.007). All but the last predictor were significant, and it is indeed notable that the attitudinal variables were about as predictive of self-reported proficiency as actual measured proficiency itself. Thus, we

concluded that Cell 1 is contaminated by Cell 5, and as such, it is an imperfect approximation of Cell 2.

Discussion

The New Haven Puerto Rican and the Watsonville Mexican-American data both support the claims of demographers that there is considerable on-going language shift from Spanish to English in the United States. This shift is evident both in language choice and language proficiency measures. The time course of shift in these two measures, however, appear to be different. the Watsonville high school data, for example, the shift in both language use with siblings and with peers were quite linear and gradual (Figure 4d, 4e), whereas the shift in Spanish proficiency was characterized by a relatively sharp break at age of exposure to English of five years (Figure 4a). Likewise, the Puerto Rican sample showed discontinuities in English and Spanish proficiencies at about 10 years on the mainland (Figure 1), whereas the shift in choice was more gradual (Figure 2). remains for future research to sort out the finer differences between these two aspects of language, but I would speculate that a main source of difference might lie in the fact that language proficiency is more cognitive in nature, and thus it is bounded by maturational factors associated with first and second language acquisition. On the other language, language choice is primarily social in nature, and may exhibit more of a dosage relationship with experience.

Finally, we are in a position to examine the language and acculturation questions used by the HHANES survey. According to Delgado, Johnson, Roy and Treviño (1990), the following questions comprise the acculturation scale:

- (1) What language do you speak?
- (2) What language do you prefer?
- (3) Can you read Spanish? Can you read English? Which do you read better?
- (4) Can you write Spanish? Can you write English? Which do you write better?
- (5) What ethnic identification do you use?
- (6) What ethnic identification does/did your mother use?
- (7) What ethnic identification does/did your father use?
- (8) Where was the birthplace of yourself, your mother, your father?

Perhaps the most informative and instructive finding to emerge from analyses of the HHANES dataset to this point is the work by Solis, Marks, Garcia and Shelton (1990). These investigators factor analyzed these questions, and found three groups: language preferred and spoken (Items 1 and 2); language written and read (Items 3 and 4), and ethnic identification (Items 5-8). They compared the relative contributions of these factors in predicting preventive health behavior in the Mexican-American sample, and found spoken language to be better than reading and

writing ability, and they also found that acculturation was not a significant factor.

In the terminology that we have been using in this paper, the first two questions on spoken and preferred language are the closest to questions about language choice, whereas questions 3 and 4 are most closely associated with proficiency. The last four questions on identification would most likely be related to language attitude, but probably only indirectly so. Indeed, based on the Watsonville study, we might argue that the questions on language choice are related to both language attitude and language proficiency, and that the question about language proficiency, because it is based on self-report, probably also picks up on language attitude. It is then not altogether surprising that language choice emerges as the most powerful of their three factors.

The power of the language variables is attested further by the findings of Marks, Garcia and Solis (1990) who were able to use the full scale for the Mexican-American sample in predicting alcohol use, but were only able to use the language items for the Cuban-Americans and the Puerto Ricans due to highly restricted variance on the identification items. Nevertheless, the magnitude of prediction was similar across the three groups, suggesting that the identification items (5-8) added little predictive power for the Mexican-American sample.

In conclude, there is considerable evidence for language shift among various Hispanic communities in the United States, and this shift process is usefully characterized as a combination of processes related to proficiency, choice, and attitudes.

These components of shift may have different developmental time courses, and therefore should be measured separately. Further, self-report of the proficiency variable is likely to contain a substantial attitudinal component, and therefore should be measured through observation to obtain valid proficiency estimates. A better understanding of the community and psychological processes involved in language shift will hopefully contribute to a better understanding of the acculturation variable in studies of health in language minority populations.

FOOTNOTES

Preparation of this paper was supported in part by a grant from the Spencer Foundation.

REFERENCES

- Acculturation and marijuana and cocaine use: Findings from
 HHANES 1982-84. American Journal of Public Health, 80.
 Supplement, 54-60.
- Cleveland, W. S. (1981). LOWESS: A program for smoothing scatterplots by robust locally weighted regression. The American Statistician, 35, 54.
- Delgado, J., Johnson, C., Roy, I. & Treviño, F. (1990). Hispanic

 Health and Nutrition Examination Survey: Methodological

 considerations. American Journal of Public Health, 80.

 Supplement, 6-10.
- Dunn, L. (1965). <u>Peabody Picture Vocabulary Test</u>. Circle Pines, MN: American Guidance Service.
- Fishman, J. (1966). Language loyalty in the United States. The Hague: Mouton.
- Hakuta, K. (1987). Degree of bilingualism and cognitive ability in mainland Puerto Rican children. Child Development, 58, 1372-1388.
- Hakuta, K. & D'Andrea, D. (in press). Some properties of bilingual maintenance and loss in Mexican background high school students. Applied Linguistics.
- Hakuta, K. & Diaz, R. (1985). The relationship between degree of bilingualism and cognitive ability: a critical discussion and some new longitudinal data. In K. E. Nelson (Ed.),

- Children's Language, Volume 5 (Pp. 319-344). Hillsdale, N. J.: Lawrence Erlbaum Associates.
- Hakuta, K., Ferdman, B. M. & Diaz, R. M. (1987). Bilingualism and cognitive development: Three perspectives. In S.

 Rosenberg (Ed.), Advances in Applied Psycholinguistics

 Volume II: Reading, Writing and Language Learning. (pp. 284-319). Cambridge: Cambridge University Press.
- Higginbotham, J., Treviño, F. & Ray, L. (1990). Utilization of curanderos by Mexican Americans: Prevalence and predictors: Findings from HHANES 1982-84. American Journal of Public Health, 80, Supplement, 32-35.
- Marks, G., Garcia, M. & Solis, J. (1990). Health risk behaviors of Hispanics in the United States: Findings from HHANES,

 1982-84. American Journal of Public Health, 80, Supplement,
 20-26.
- Solis, J., Marks, G., Garcia, M. & Shelton, D. (1990).

 Acculturation, access to care, and use of preventive services by Hispanics: Findings from HHANES 1982-84.

 American Journal of Public Health, 80, Supplement, 11-19.
- Veltman, C. (1983). <u>Language shift in the United States</u>.

 Berlin, New York, and Amsterdam: Mouton.
- Veltman, C. (1988). The Future of the Spanish Language in the United States. Hispanic Policy Development Project: New York City and Washington, DC.

- 'iener, F. D., Simmond, A. J. & Weiss, F. L. (1978). Spanish

 <u>Picture Vocabulary Test</u>. New York: Marymount Manhattan

 College.
- Wilkinson, L. (1987). SYGRAPH. Evanston, IL: SYSTAT, Inc.

FIGURE CAPTIONS

- Figure 1. Spanish and English proficiency (as measured by raw scores on receptive vocabulary tests, Dunn, 1965 for English, Wiener, Simmonds & Weiss, 1978 for Spanish) as a function of years of residence on the mainland, separately by grade level. A locally weighted regression was applied to summarize the data trend (Wilkinson, 1987). So were Puerto Ricán children enrolled in a bilingual education program in New Haven, Connecticut.
- Figure 2. Language choice (as reported by parents, 1=only Spanish, 2=mostly Spanish, 3=both English and Spanish, 4=mostly English, 5=only English) as a function of years of residence on mainland. A locally weighted regression was applied to summarize the data trend (Wilkinson, 1987). The three panels are for (a) language among adults, (b) language between children and adults, and (c) language among siblings. Respondents were parents of Puerto Rican children enrolled in a bilingual education program in New Haven, Connecticut.
- Figure 3. Correlation between reported language choice among siblings and proficiency difference between English and Spanish (English raw vocabulary score minus English raw vocabulary score). A locally weighted regression was applied to summarize the data trend (Wilkinson, 1987). Linear correlation coefficient for Grades K-1 was r=.32, and for Grades 4-5, r=.23.
- Figure 4. Language proficiency (a,b) and language choice (c,d,e) measures as a function of reported age when subject was exposed to English. A locally weighted regression was applied to summarize the data trend (Wilkinson, 1987). So were Mexican-American high school students in Watsonville, California.

- 'igure 5. Spanish and English proficiency as a function of reported language choice with adults at home (1=only Spanish, 5=only English, selecting only Ss who were exposed to English before age 6). A locally weighted regression was applied to summarize the data trend (Wilkinson, 1987).
- siblings and proficiency difference between English and Spanish (English raw vocabulary score minus English raw vocabulary score). A locally weighted regression was applied to summarize the data trend (Wilkinson, 1987). Data for Grades 7-8 were obtained from feeder schools to the high school in Watsonville, California. Linear correlation coefficient for Grades 7-8 was <u>r</u>=.39, and for Grades 9-12, <u>r</u>=.62.

















