DOCUMENT RESUME

ED 245 575 FL 014 436

AUTHOR Cummins, Jim
TITLE Language and Literacy Learning in Bilingual Instruction: Policy Report.
INSTITUTION Southwest Educational Development Lab., Austin, TX. Div. of Language and Literacy.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Sep 83
CONTRACT 400-80-0043
NOTE 70p.; For related documents, see FL 014 432-437.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Information Analyses (070)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Bilingual Education; Educational Policy; English (Second Language); Immigrants; *Language Research; Linguistic Theory; *Literacy Education; Literature Reviews; *Policy Formation; *Public Policy; Research Utilization; Second Language Instruction; *Second Language Learning

ABSTRACT

The findings are presented of a study undertaken by the Southwest Educational Development Laboratory (SEDL) to assess the bilingual education services offered to limited English speaking students in three diverse communities and to study the extent to which the development of cognitive and academic skills in the native language and English are interrelated. The findings are integrated with current research and theory related to academic development in bilingual contexts, focusing on the policy implications of the SEDL findings. First, the general question of how research findings relate to educational policy is briefly considered. It is suggested that much of the confusion about the research basis for bilingual education derives from educators' and policy-makers' failure to appreciate the role of theory in the formulation of policy. The alternative policy-related theoretical arguments for and against bilingual education are outlined and the major SEDL study findings are examined in relation to these theoretical positions. The results of other related studies are reviewed and the extent to which policy-relevant theoretical principles emerge from the data is considered. (Author/MSE)

Policy Report:
Language and Literacy Learning in Bilingual Instruction

This report is one of a series of documents produced for the LANGUAGE AND LITERACY LEARNING IN BILINGUAL INSTRUCTION STUDY by the Southwest Educational Development Laboratory:

Descriptive Studies - Asian, Spanish, Navajo: Final Report

Cantonese Analytic Study: Final Report

Language and Literacy Learning in Bilingual Instruction: Executive Summary

Language and Literacy Learning in Bilingual Instruction: A Case Study of Practices and Outcomes

Policy Report: Language and Literacy Learning in Bilingual Instruction

Summary Report: Policy Report - Language and Literacy Learning in Bilingual Instruction

---

This document was produced under Contract No. 400-80-0043 from the National Institute of Education, U.S. Department of Education to the Southwest Educational Development Laboratory. However, it does not necessarily reflect the official views of either the Institute or the Department. Information about the study can be obtained from the National Institute of Education and from the following persons from the Southwest Educational Development Laboratory who have participated in the study: Wesley A. Hoover, Betty J. Mace-Matluck, Domingo Domínguez

POLICY REPORT: LANGUAGE & LITERACY LEARNING IN BILINGUAL INSTRUCTION

Jim Cummins

ONTARIO INSTITUTE FOR STUDIES IN EDUCATION

A Policy Review of the Southwest Educational Development Laboratory
Study submitted to the National Institute of Education, September 1983

TABLE OF CONTENTS

TABLE OF CONTENTS

POLICY REPORT: LANGUAGE AND LITERACY LEARNING IN BILINGUAL INSTRUCTION

Jim Cummins

ONTARIO INSTITUTE FOR STUDIES IN EDUCATION

I. INTRODUCTION

The purpose of this review is to integrate the findings of the SEDL study with current research and theory related to academic development in bilingual contexts in order to highlight the policy implications of the findings. The SEDL study of Cantonese-speaking students in Seattle directly addressed the question of the effects of bilingual education on academic development as well as issues central to the theoretical assumptions underlying bilingual education. Specifically, the study tests the validity of the "interdependence hypothesis" (Cummins, 1979; 1981a) by investigating the extent to which the development of cognitive/academic skills in English (L2) is a function of prior development of cognitive-academic skills in Cantonese (L1).

The policy report is organized as follows: first, in order to facilitate interpretation of the policy implications of the SEDL study, the general question of how research findings relate to educational policy is briefly considered. It is suggested that much of the confusion about the research basis for bilingual education derives from failure on the part of educators and policy-makers to appreciate the role of theory in the formulation of policy. Then the alternative policy-related theoretical arguments both for and against bilingual education are outlined. Next, the major findings of the SEDL study are examined in relation to these theoretical positions. Finally, the results of other related studies are reviewed and the extent to which policy-relevant theoretical principles emerge from the data is considered.

II. WHAT CONSTITUTES EVIDENCE? RESEARCH, THEORY AND POLICY

All educational programs are conceived within a particular societal context which significantly affects their implementation. For example, a bilingual program which is perceived as threatening by a community or by a school staff is likely to have different results from one which is enthusiastically supported by all concerned. Similarly, an L2 immersion program may have very different results when implemented for middle-class English-background students in Canada, minority Spanish-speakers in the United States and tribal language speakers in Africa, India or South America. Differences in language status, minority-majority group relations, student background characteristics, curriculum materials and teacher training are all likely to influence program outcomes.

The point is that research findings, in themselves, cannot be directly generalized across contexts. This does not of course mean that research findings have no relevance outside of the specific context in which they were obtained. They can become relevant to other contexts when they are integrated into a coherent theory from which predictions about program outcomes under different conditions can be generated.

The importance of theory for the policy-making process is often not appreciated. Policy-makers are interested in "getting the facts" and theoretical considerations are sometimes dismissed pejoratively as "just theory". Obviously, theories may be inadequate or invalid, but the point that must be emphasized is that the policy concern with "getting the facts" is futile unless the "facts" can be understood or explained within the context of a theory. It is the theory rather than the individual "facts" that has policy implications because it is the theory that allows the relevance of the "facts" in different contexts to be understood and consequently applied in the form of specific programs.

In the present context, it is clear that the results of, for example, a transitional bilingual program for Turkish and Moroccan students in Holland (Altena and Appel, 1982) or an enrichment bilingual program for Aboriginal students in Australia (Gale, McClay, Christie & Harris, 1981) cannot be directly generalized to the United States context. However, the theoretical or explanatory principles

underlying the apparent success of these programs are, almost by definition, generalizable across contexts insofar as the validity of any theoretical principle is assessed precisely by how well it is capable of accounting for data from a wide variety of contexts. If a theory cannot account for a particular set of research findings, then it is an inadequate or incomplete theory.

Thus, in reviewing the policy implications of the SEDL study, our primary concern is not whether the findings support or refute bilingual education as such; the focus is rather on the extent to which the findings are consistent with alternative theoretical predictions regarding the academic progress of minority students under different instructional conditions.

Viewed from within this perspective, the study becomes much more powerful in its policy implications. An adequate or valid psychoeducational theory of bilingual education must be capable of accounting for the SEDL data (as well as for the findings of other methodologically-sound studies). Thus, serious doubt is cast on any theory which cannot account for the findings. If the SEDL findings were not integrated into a theoretical framework, their policy impact would likely be minimal because they would become just one more set of findings to be placed on the weighing scale "for" or "against" the "effectiveness" of bilingual education. The problem with this "weighing scale" approach to drawing policy implications from research (employed in the Baker and de Kanter (1981) review) is that it deals only at the level of research findings which are not generalizable from one context (e.g. program, ethnic group, city, etc.) to another and it ignores the theoretical principles which, by definition, are generalizable across contexts and therefore relevant to policy.

In short, in reviewing research on the L2 acquisition of minority students in a wide variety of contexts, our purpose is to uncover the theoretical principles which can account for the pattern of findings. It is only on the basis of these explanatory principles that predictions can be made about the probable outcomes of a variety of program models for minority students in different U.S. contexts.

III. THEORETICAL ASSUMPTIONS UNDERLYING U.S. BILINGUAL EDUCATION POLICY

The controversial debate in the United States about bilingual education has been characterized by claims and counter-claims about the effectiveness of bilingual education in comparison to actual and potential alternatives in reducing academic failure among minority students. There exist a considerable number of evaluation reports which suggest that bilingual education resulted in improved academic performance among minority students (e.g. Troike, 1978); however, other reviews of the literature have interpreted the research findings much less favorably for bilingual education in that in many programs no improvement in student academic performance in comparison to monolingual English programs was noted (e.g. Baker and de Kanter, 1981). The result of these differences in interpretation of the research evidence has been considerable confusion and scepticism among policy-makers, many of whom suspect that bilingual education might be motivated more by political than educational considerations.

Compounding the confusion in the research findings is the intuitive appeal of the conventional wisdoms or theoretical assumptions underlying both sides of the debate. The argument for bilingual education is that children cannot learn through a language they do not understand and therefore require some mother-tongue (L1) instruction in the initial grades to make the academic content comprehensible. According to this line of reasoning the major cause of minority children's academic failure is mismatch between the language of the home and the language of schooling.

The altnernative argument is that if children are deficient in English then what they need is surely as much instruction in English as they can get; it makes very little sense, according to this view, to attempt to develop children's English skills by means of instruction through their L1. Proponents of monolingual English programs have increasingly cited the success of L2 "immersion" programs in Canada as evidence against bilingual education.

In short, the policy debate is currently characterized by an extreme polarization of views regarding the efficacy of bilingual education. One of the few issues upon which there appears to be some agreement is that there exists relatively

little well-controlled empirical research to inform policy decisions and that more and better research is required to resolve the issues.

In fact, as this review will attempt to make clear, there is an enormous amount of well-controlled policy-relevant research on bilingual education whose findings are consistent and unambiguous. This research includes the hundreds of evaluations of French "immersion" programs across Canada in which majority group children are educated through both French (L2) and English (L1) at no long-term cost to English and subject matter achievement. A clear pattern also emerges from well-controlled evaluations of bilingual programs for minority students in the United States, Canada, Europe, Australia and other parts of the world (see Cummins, 1983).

In the light of this research, it quickly becomes evident that the two conventional wisdoms which constitute the most frequently invoked arguments for and against bilingual education in the United States are examples of patently inadequate theoretical principles. The "linguistic mismatch" hypothesis that a home-school language switch results in academic retardation is refuted by the French immersion data as well as by the apparent success of a considerable number of language minority groups under home-school language switch conditions. The opposing "maximum exposure" argument that if minority students are deficient in English then they need as much exposure to English as possible is similarly refuted by the results of virtually every bilingual program that has ever been evaluated (including French immersion programs) all of which show that students instructed through a minority language (e.g. Spanish, French, etc.) for all or part of the school day perform as well or better in the majority language (e.g. English) as students instructed exclusively through that language. This pattern of findings suggests an alternative theortical principle, namely, the "interdependence" hypothesis.

The Interdependence Hypothesis

The fact that there is little relationship between amount of instructional time through the majority language and academic achievement in that language strongly suggests that L1 and L2 academic skills are interdependent, i.e. manifestations of a common underlying proficiency. The interdependence hypothesis has been stated formally as follows (Cummins, 1981a, p. 28):

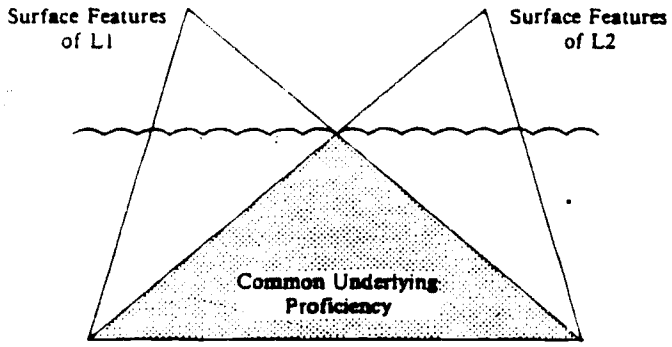
To the extent that instruction in Lx is effective in promoting proficiency in Lx, transfer of this proficiency to Ly will occur provided there is adequate exposure to Ly (either in school or environment) and adequate motivation to learn Ly.

In concrete terms what this hypothesis means is that in a bilingual program, Spanish instruction which develops L1 reading skills for Spanish-speaking students is not just developing Spanish skills, it is also developing a deeper conceptual and linguistic proficiency which is strongly related to the development of English literacy and general academic skills. In other words, although the surface aspects (e.g. pronunciation, fluency, etc.) of, for example, Spanish and English or Chinese and English are clearly separate, there is an underlying cognitive/academic proficiency which is common across languages. This "common underlying proficiency" (illustrated in Figure 1) makes possible the transfer of cognitive/academic or literacy-related skills across languages.

The three theoretical hypotheses which have been outlined, namely, the "linguistic mismatch" hypothesis, the "maximum exposure" hypothesis, and the "interdependence" hypothesis, form the context for reviewing the policy implications of the SEDL findings. The central policy question to be asked in relation to the findings is the extent to which they are consistent with each of these three hypotheses. Briefly, the predictions derived from each of the hypotheses are as follows: if the "linguistic mismatch" hypothesis is valid (i.e. generalizable across contexts), then Chinese-background students who have received bilingual instruction should perform better in English academic skills than those who have been subjected to a linguistic mismatch (i.e. English-only instruction); if the "maximum exposure" hypothesis is valid, then the opposite should be true: students who have been instructed exclusively through English should perform better than those who have been instructed through Chinese for part of the time; if the interdependence hypothesis is valid, then despite considerably less exposure to English, students who have received bilingual instruction should perform at least as well as those who have been instructed exclusively through English. The interdependence hypothesis would also predict a significant positive relationship between immigrant (i.e. schooled-abroad) students' Chinese academic skills and the development of English academic proficiency. For students who began schooling in the United States, a significant relationship between L1 and L2 academic skills is likely to be contingent on the extent to which Chinese literacy has actually been promoted in the bilingual program.

Figure 1

THE "DUAL-ICEBERG" REPRESENTATION OF BILINGUAL PROFICIENCY



IV. THE FINDINGS OF THE SEDL STUDY

(i) The Nature of Language Proficiency

The SEDL study is one of the few investigations of bilingual education that has included an adequate assessment of English language proficiency. The academic measures avoid the pitfalls of many standardized tests by assessing a wide range of reading and writing skills. Furthermore, the Interactive Reading Assessment Scale (IRAS) is based on an impirically-supported theory of the reading process and is individually administered, a fact which increases confidence in its validity. Students' formal and informal oral language is also assessed in considerable detail. Thus, the study permits the relationship between academic and communicative English L2 skills to be assessed.

The relevance of this issue for policy in bilingual education derives from the fact that the criteria for exit from bilingual programs have been problematic. It has been suggested, for example, that many students are exited before they have developed sufficient English academic proficiency to survive in an all-English classroom, despite the fact that their oral fluency in English may be well-developed (Cummins, 1981a). If the SEDL data showed extremely high correlations between academic and conversational English skills, the case for a distinction between these two aspects of proficiency would be weakened. Apart from the immediate policy implications of the SEDL findings in this regard, their more general theoretical implications are highly significant in that the nature of "language proficiency" has been a hotly debated issue among psycholinguists during the past decade.

The first point to note in relation to the language proficiency data is that the factor analysis of oral measures showed a clear distinction between "discourse", reflecting more school-based formal language, and both "grammar" and "interactional style" which are derived from more informal conversational measures. As the authors of the SEDL report point out, the distinction between these latter two dimensions replicates the findings of the Cummins et al (in press) study carried out with Japanese immigrant students.

Also consistent with the Cummins et al study is the finding that academic skills (e.g. as measured by the IRAS) are clearly separate from interactional style. Only about 25% of the variance is shared between these two dimensions, although about 50% is shared between academic proficiency and oral grammar.

A policy implication of these results is that students' ease in conversational situations, i.e., their interactional style, may easily give a misleading impression of their ability to participate academically in monolingual English instructional settings. From a more general theoretical perspective the data support the contention that the linguistic demands of academic and conversational tasks are distinct (see California State Department of Education, 1982).

(ii) The Effects of Bilingual Education and L1 Development

Previous studies carried out among Chinese background students suggest that they do not, as a group, experience disproportionate school failure, whether they are in bilingual or monolingual school programs (e.g. see Rosenthal et al, 1981; Vernon, 1982 for reviews). A variety of hypotheses have been suggested to account for the relative success of these students compared to some other minority language groups (see, for example, Ogbu, 1978; Wong Fillmore, 1983). Among the factors that appear to be important are the fact that education is highly valued in the Chinese culture and children tend to be considerably more adult-oriented than peer-oriented in school situations (Wong Fillmore, 1983).

To what extent is this pattern of relatively good school achievement replicated in the SEDL data? Predictions of grade 6 California Achievement Test (CAT) percentiles for the entire Chinese sample in grades 4 through 6 showed students performing at the 46th percentile in reading and at the 80th percentile in math. These figures are somewhat misleading, however, since they include both students who started schooling in the U.S. as well as more recent immigrant students who received some of their schooling abroad. When the percentile scores of these two groups are considered separately we find students who started schooling in the U.S. (in a bilingual program) scoring at the 58th percentile in reading whereas those who started school abroad were scoring at the 33rd percentile. Differences were not as large in math, with both groups scoring around the 80th percentile. The lower percentile rank in reading for

TABLE 1

MEANS OF CHINESE BACKGROUND STUDENTS
IN ENGLISH READING AND WRITING

Group	Amount of Bilingual or L1 Instruction (Semesters)	IRAS-E	IWI
U.S.A.-Schooled	0-2	6.2	-.01
	3-4	7.1	.18
	5-6	7.4	.56
	8	7.5	.53
Foreign-Schooled	2-3	6.1	-.13
	4-5	4.7	-.69
	6-8	5.1	-.51
	10-12	5.7	-.36

TABLE 2

CORRELATIONS BETWEEN ENGLISH ACADEMIC PROFICIENCY
AND AMOUNTS OF L1 AND L2 INSTRUCTION (N=112)

U.S.A.-Schooled	IRAS-E	IWI
Amount of L2 Instruction	.28*	.34*
Amount of L1 Instruction	.24*	.23*
Foreign-Schooled		
Amount of L2 Instruction	.36*	.33*
Amount of L1 Instruction	.00	-.02

* P<.05

those schooled abroad can be attributed to the fact that they have been in the U.S. for an insufficient amount of time to permit them to reach grade norms. Studies have found that between five and seven years, on the average, is required for immigrant students to approximate grade norms in English academic skills (Cummins, 1981b).

Clearly, the SEDL data are consistent with other studies in showing Chinese students who start school in the U.S. to be performing well academically. To what extent can this level of academic performance be attributed to the effects of bilingual instruction and/or L1 development?

In order to answer this question the data on students' reading and writing performance as measured by the IRAS and Informal Writing Inventory (IWI) will be examined. Means of students in different subgroups will first be presented, followed by simple correlations between amount of L1 instruction and English academic proficiency; finally, regression analyses with English reading and writing as dependent variables will be presented. The variables included in the regression analyses differ somewhat from those in the principle regression analyses in the SEDL report because here we are interested in testing very specific hypotheses about the relationships between L1 instruction and the development of L2 academic skills and also we are interested in making direct comparisons between the results of the SEDL study and those of previous studies. Thus, because length of residence (LOR) is a variable frequently included in previous studies, it is also included in these analyses rather than the variable "Amount of L2 Training" which is used in the main SEDL regression analyses.

(a) **IRAS-E and IWI Means.** The means of reading and writing scores for students who received different amounts of Cantonese instruction in the U.S.A.-schooled and foreign-schooled groups are presented in Table 1 (from Table 30 in the SEDL Preliminary Report, 1982). The IRAS-E scores are essentially grade levels while the IWI scores are standardized such that negative scores represent performance below the total group average whereas positive scores are above the group average.

For the U.S.A.-schooled subgroups it can be seen that those students who have had the greatest amount of bilingual instruction perform considerably better in

both English reading and writing than those who have had little bilingual instruction. Since all of these students started school in the U.S., their out-of-school exposure to English is likely to have been similar. Thus an examination of means suggests that Chinese students benefit academically from bilingual instruction.

For the foreign-schooled students, the more L1 instruction they have had in their homeland, the less their length of residence in the U.S. is likely to have been, and consequently, the less opportunity they would have had to acquire English. Thus, it is interesting to note the relatively good performance of the two subgroups who have had the most L1 instruction. These two subgroups perform better in both reading and writing than the subgroup that has had only 4-5 semesters of L1 instruction, and presumably considerably more exposure to English instruction as a consequence.

In summary, for the U.S.A.-schooled students, the means suggest a positive relationship between the amount of bilingual instruction and English achievement; for the foreign-schooled students, the relationship is non-linear in that the two subgroups that perform the best are those that have had the least and the most L1 instruction. It is worth noting also the generally high grade-level equivalents obtained by students on the IRAS-E which suggest that these students, as a group, are not experiencing major academic difficulties.

(b) Correlational Analyses. The correlations between IRAS-E and IWI scores on the one hand, and the amounts of L1 and L2 instruction, on the other, are presented in Table 2. It can be seen that within the U.S.A.-schooled group, English academic proficiency is significantly related to the amount of both L1 and L2 instruction received. The positive relationship between L2 instruction and proficiency is likely to be partially due to the effects of age or grade level, since students in grade 6 will have had considerably more L2 instruction than those in grade 4. This confounding is not present for the amount of L1 instruction since no bilingual instruction was offered after grade 3. Thus, the correlations are consistent with the trend that was apparent in the mean scores for IRAS-E and IWI, namely, students' academic development in English appears to have been positively affected by the Cantonese instruction they received.

A different pattern emerges for the foreign-schooled group. Here we find a significant relationship between amount of L2 instruction and achievement but no relationship between amount of L1 instruction and achievement. The former relationship presumably reflects the effects of both grade level and length of residence in the U.S.; by the same token, any effect of L1 instruction is confounded by the effects of length of residence since those who have had more L1 instruction abroad have had less exposure to English in the U.S. For the foreign-schooled group in grades 4-6 the mean length of residence was approximately 2.5 years and thus these students are likely to continue to approach grade norms in English academic proficiency for several more years. The influence of L1 development among this group can be ascertained more adequately in the regression analyses where the effects of length of residence (LOR) can be controlled.

(c) **Regression Analyses.** In order to assess the effects of L1 instruction and proficiency on English academic development, regression analyses were carried out first for the total group combined and then separately for each group. Two sets of variables were chosen for the analyses: first, LOR was chosen to assess the effects of exposure to English and opportunity to acquire English; second, three variables were chosen to represent an L1 "cognitive/academic" block. These variables were: 1. Amount of L1 instruction; 2. IRAS-Cantonese (C), which assesses Cantonese reading skills; and 3. Age of arrival (AOA), which for the foreign-schooled group is likely to reflect level of both cognitive maturity and L1 literacy skills. These variables were entered into the regression analyses after LOR in order to control for the effects of LOR in assessing their influence on students' English academic proficiency.

The summary table for the IRAS-E total score regression is presented in Table 3. The first variable entered in this regression was the group membership dummy variable (FORUSA), indicating whether a student was first schooled abroad or in the U.S.A. Then comes LOR for the foreign-schooled and U.S.-schooled students, followed by the amount of L1 instruction received by each group. IRAS-C and AOA were the final variables entered.

It can be seen that LOR is highly significant for both foreign-schooled and U.S.A.-schooled groups and the amount of L1 instruction received also attains

significance ($p<.05$) for both groups. The other four variables in the analysis add about 4% to the explanation of variance with a total of 57% of the variance explained. The fact that IRAS-C for the foreign-schooled group only approaches significance is due to the overlap between this variable and L1FOR. After LORFOR is entered into the regression equation on step 2, the partial correlation between IRASCFOR and IRAS-E becomes .23 which is significant at $p<.02$.

The analysis confirms the trends already noted and suggests that for minority students who come to school first in the U.S., bilingual instruction facilitates their acquisition of English academic skills, whereas for those who immigrate after having started schooling in their home country, the extent to which literacy in their L1 has been promoted is a significant factor in the development of English academic proficiency.

The summary table for the IWI regression is presented in Table 4. Overall, 54% of the variance is explained, the bulk of this being acounted for by the group membership variable (FORUSA) and by the LOR variables. L1FOR approaches significance ($p<.09$) but of the remaining variables only AOAUSA attains significance. Although negatively correlated with IWI, its partial correlation with IWI becomes positive (.32 $p<.001$) after LORUSA is entered into the equation on step 3.

In general, it appears that the influence of L1 instruction and of transfer from L1 to L2 is less in the case of English writing than reading. However, it is somewhat unclear what exactly is being included in the effects of group membership (FORUSA) and therefore it is useful to look at the separate regression analyses for each of the groups. These analyses are presented in Tables 5 and 6 for the U.S.A. first-schooled group and in Tables 7 and 8 for the foreign-schooled group.

For IRAS-E, LOR accounts for 19% of the variance in the U.S.A. first-schooled group, while the L1 cognitive/academic block together accounts for almost 14%. Within this block, the amount of bilingual instruction (AMTL1L2) attains significance ($p<.05$) while IRAS-C and AOA approach significance.

Only three variables enter the equation for IWI in the U.S.A. first-schooled group. LOR accounts for almost 16% of the variance while AMTL1L2 and AOA together account for about 11.5%; however, AMTL1L2 does not attain significance.

For the foreign first-schooled group, LOR explains 29% of the variance on IRAS-E, while the L1 cognitive/academic block explains almost 11%. LOR again explains the bulk of the variance (42.6%) on IWI for this group but the amount of exclusive L1 instruction (AMTEXCL1) also enters the equation, accounting for an additional 5.9% ($p<.05$). Thus, a significant effect is again noted for L1 development on English writing skills in addition to English reading skills.

The results of the regression analyses carried out on the L2 discourse, grammar and interactional style variables can be summarized briefly. In general, relatively little variance was explained in these analyses and the L1 cognitive/academic block, as expected, did not play a significant role. For both groups of students, the greatest amount of variance (<40%) was explained on the grammar variable with LOR accounting for most of that variance.

For Cantonese proficiency, only 15% and 21% of the variance was explained in the U.S.A. first-schooled and foreign-schooled analyses respectively. In the former analysis the major variable was AOA, accounting for 10% of the variance; this suggests that the older students were when they immigrated to the U.S. before schooling, the better their prospects for developing Chinese literacy. An implication of the findings is that bilingual education has not been particularly successful in developing students' Chinese literacy skills.

The only variable to enter the equation in the foreign first-schooled analysis was AMTEXCL1 which accounted for 21%. It is perhaps surprising that more of the variance was not accounted for by this variable.

TABLE 3

CANTONESE SITE DEPENDENT AND INDEPENDENT MEASURES 28 JUN 83 10.02.05. PAGE 20
IRAS-E REGRESSED
FILE NONAME (CREATION DATE = 28 JUN 83)

* M U L T I P L E R E G R E S S I O N *

DEPENDENT VARIABLE.. IRASETOT AVE IRAS-E SCORE OVER 6 SCALES

S U M M A R Y T A B L E

STEP	VARIABLE ENTERED	VARIABLE REMOVED	F TO ENTER OR REMOVE	SIGNIFICANCE	MULTIPLE R	R SQUARE	R SQUARE CHANGE	SIMPLE R	OVERALL F	SIGNIFICANCE
1	FORUSA		52.20316	.000	.55731	.32184	.32184	.56731	52.20316	.000
2	LORFOR		22.63247	.000	.66215	.43844	.11660	.34143	42.55092	.000
3	LORUSA		11.51011	.001	.70180	.49253	.05409	.23255	34.93960	.000
4	L1FOR		5.24070	.024	.71849	.51622	.02369	-.00897	28.54382	.000
5	L1USA		4.03303	.047	.73072	.53395	.01773	.23773	24.28894	.000
6	IRASCFOR		3.03542	.084	.73963	.54705	.01309	.13845	21.13536	.000
7	IRASCUSA		2.49759	.117	.74677	.55767	.01062	.07367	18.73120	0
8	AOAUSA		2.59810	.110	.75402	.56855	.01088	-.16189	16.96642	.000
9	AOAFOR		1.49614	.224	.75815	.57479	.00624	-.19112	15.32014	.000

TABLE 4

CANTONESE SITE DEPENDENT AND INDEPENDENT MEASURES 28 JUN 83 10.02.05. PAGE 33
IMITOT REGRESSED
FILE NONAME (CREATION DATE = 28 JUN 83)

* M U L T I P L E R E G R E S S I O N *

DEPENDENT VARIABLE.. IMITOT AVE IMI TOTAL SCORE

S U M M A R Y T A B L E

STEP	VARIABLE ENTERED	VARIABLE REMOVED	F TO ENTER OR REMOVE	SIGNIFICANCE	MULTIPLE R	R SQUARE	R SQUARE CHANGE	SIMPLE R	OVERALL F	SIGNIFICANCE
1	FORUSA		44.55980	.000	.53694	.28830	.28830	.53694	44.55980	.000
2	LORFOR		19.31286	.000	.62883	.39542	.10712	.32725	35.64550	.000
3	LORUSA		14.60185	.000	.68369	.46743	.07201	.26833	31.59636	.000
4	L1FOR		3.04577	.084	.69438	.48217	.01474	-.02473	24.90760	.000
5	L1USA		2.10118	.150	.70159	.49223	.01007	.22941	20.55138	.000
6	IRASCUSA		1.24807	.266	.70583	.49820	.00596	.03764	17.37424	.000
7	AOAUSA		8.46534	.004	.73210	.53597	.03777	-.17977	17.16036	.000

TABLE 5

CANTONESE SITE DEPENDENT AND INDEPENDENT MEASURES 01 SEP 83 13.58.22. PAGE 50
IRAS-E REGRESSED (USA FIRST-SCHOOLED)
FILE NONAME (CREATION DATE = 01 SEP 83)

* MULTIPLE REGRESSION *

DEPENDENT VARIABLE.. IRASETOT AVE IRAS-E SCORE OVER 6 SCALES

SUMMARY TABLE

STEP	VARIABLE ENTERED	VARIABLE REMOVED	F TO ENTER OR REMOVE	SIGNIFICANCE	MULTIPLE R	R SQUARE	R SQUARE CHANGE	SIMPLE R	OVERALL F	SIGNIFICANCE
1	LOR		15.36589	.000	.44001	.19361	.19361	.44001	15.36589	.000
2	AMTL1L2		5.38253	.024	.50703	.25708	.06347	.44966	10.90031	.000
3	IRASCTOT		3.34453	.072	.54324	.29511	.03802	.13938	8.65215	.000
4	ADA		3.56840	.064	.57798	.33406	.03896	-.30624	7.65003	.000

TABLE 6

CANTONESE SITE DEPENDENT AND INDEPENDENT MEASURES 01 SEP 83 13.58.22. PAGE 58
IWITOT REGRESSED (USA FIRST-SCHOOLED)
FILE NONAME (CREATION DATE = 01 SEP 83)

* MULTIPLE REGRESSION *

DEPENDENT VARIABLE.. IWITOT AVE IWI TOTAL SCORE

SUMMARY TABLE

STEP	VARIABLE ENTERED	VARIABLE REMOVED	F TO ENTER OR REMOVE	SIGNIFICANCE	MULTIPLE R	R SQUARE	R SQUARE CHANGE	SIMPLE R	OVERALL E	SIGNIFICANCE
1	LOR		11.86404	.001	.39546	.15639	.15639	.39546	11.86404	.001
2	AMTL1L2		1.67592	.200	.42219	.17825	.02186	.33800	6.83263	.002
3	ADA		8.08329	.006	.52252	.27303	.09478	-.26489	7.76166	.000

TABLE 7

CANTONESE SITE DEPENDENT AND INDEPENDENT MEASURES 01 SEP 83 13.58.22. PAGE 12
IRAS-E REGRESSED (FOREIGN FIRST-SCHOOLED)
FILE NONAME (CREATION DATE = 01 SEP 83)

* MULTIPLE REGRESSION *

DEPENDENT VARIABLE.. IRASETOT AVE IRAS-E SCORE OVER 6 SCALES

SUMMARY TABLE

STEP	VARIABLE ENTERED	VARIABLE REMOVED	F TO ENTER OR REMOVE	SIGNIFICANCE	MULTIPLE R	R SQUARE	R SQUARE CHANGE	SIMPLE R	OVERALL F	SIGNIFICANCE
1	LOR		18.18029	.000	.54072	.29238	.29238	.54072	18.18029	.000
2	AMTEXCUR		3.94144	.054	.59312	.35180	.05942	-.00164	11.66854	.000
3	IRASCTOT		2.24100	.142	.62019	.38463	.03283	.21905	8.75053	.000
4	ABA		1.06918	.307	.63267	.40027	.01564	-.30242	6.84100	.000

- 18 -

TABLE 8

CANTONESE SITE DEPENDENT AND INDEPENDENT MEASURES 01 SEP 83 13.58.22. PAGE 18
IWITOT REGRESSED (FOREIGN FIRST-SCHOOLED)
FILE NONAME (CREATION DATE = 01 SEP 83)

* MULTIPLE REGRESSION *

DEPENDENT VARIABLE.. IWITOT AVE IWI TOTAL SCORE

SUMMARY TABLE

STEP	VARIABLE ENTERED	VARIABLE REMOVED	F TO ENTER OR REMOVE	SIGNIFICANCE	MULTIPLE R	R SQUARE	R SQUARE CHANGE	SIMPLE R	OVERALL F	SIGNIFICANCE
1	LOR		32.69838	.000	.65294	.42632	.42632	.65294	32.69838	.000
2	AMTEXCU1		4.89804	.032	.69641	.48499	.05866	-.04945	20.24661	.000

(iii) Summary of the SEDL Findings

The findings of the SEDL analytic study of Chinese students provide support for the distinction made between academic and conversational L2 skills (Cummins, 1981a) and they also confirm previous indications that Chinese background students tend to perform relatively well in English academic skills. Students' performance tends to be exceptionally high in mathematical tasks. The analyses also tend to show a consistent moderate effect of bilingual instruction (for the U.S.A. first-schooled group) and of L1 instruction (for the foreign-schooled group) on the development of English reading and, to a lesser extent, writing skills. The distinction between academic and conversational skills is further supported by the fact that the L1 cognitive/academic variables exerted virtually no effect on the English conversational variables.

(iv) Policy Implications of the SEDL Findings

The policy implications of the study can be assessed by the extent to which the findings are compatible with the three theoretical hypotheses outlined earlier, namely, the linguistic mismatch hypothesis, the maximum exposure hypothesis, and the interdependence hypothesis.

Although other studies refute the generalizability of the linguistic mismatch hypothesis, the present findings appear consistent with the hypothesis in that students first schooled in the U.S., who had received little or no bilingual instruction, performed considerably worse on English literacy tasks than those who had received bilingual instruction. As outlined in Table 1, the group who received minimal bilingual instruction obtained a score of 6.2 on the IRAS-E compared to score of more than 7 for the other three groups. The consistent positive effect for amount of bilingual instruction in the regression analyses is also clearly compatible with predictions derived from the linguistic mismatch hypothesis.

The maximum exposure hypothesis is clearly refuted by the SEDL findings. If the hypothesis were valid then there should be a negative effect for amount of bilingual or L1 instruction on student English achievement. The opposite, in fact, is the case. Thus, the SEDL findings refute the common opinion of many

press commentators and policy-makers that bilingual programs should be replaced by monolingual programs which provide maximum exposure to English for minority students.

The SEDL findings are also consistent with the interdependence hypothesis. No negative consequences were found as a result of bilingual instruction and a positive relationship was observed between the promotion of L1 academic proficiency and the development of L2 literacy skills. The results suggest that transfer of underlying cognitive/academic skills occurred between languages despite the surface structure dissimilarity between Chinese and English.

In summary, the findings suggest that even in situations where minority students are performing relatively well academically, bilingual instruction can exert a moderate positive effect on English academic achievement. The next section analyzes the extent to which this conclusion and the three policy-related hypotheses discussed above are compatible with the results of other studies which have investigated similar issues.

V. PREVIOUS RESEARCH REGARDING THE EFFECTS OF BILINGUAL EDUCATION

The most prominent reviews of research findings on bilingual education in the United States are those of Troike (1978), Dulay and Burt (1979), Zappert and Cruz (1980) and Baker and de Kanter (1981). The earlier review by Engle (1975) focused largely on research conducted outside the United States. Of the four prominent U.S. reviews, the first three reach conclusions supporting the effectiveness of bilingual education in promoting academic skills among minority students; however, the Baker/de Kanter review is by far the most comprehensive and its conclusions are not supportive of transitional bilingual education. All reviewers comment on the poor methodological quality of many Title VII (bilingual education) evaluations and base their conclusions only on those they consider acceptable. However, the criteria of methodological acceptability adopted by Baker and de Kanter differ in many respects from those in the other reviews.

Dulay and Burt (1979, p. 2) summarize the results of their review and Troike's as follows:

> "Despite the recentness of this complex innovation, more than half of the findings show that bilingual education worked significantly better than monolingual programs for LES/NES (limited English-speaking/non-English speaking) students. Another recent survey has revealed 12 more research studies conducted since 1976 which demonstrate the effectiveness of bilingual instruction for LES/NES students (Troike, 1978)."

On the basis of the fact that several evaluations showed that students in successful bilingual programs attained grade norms in English academic skills only in the later grades of elementary school, Troike (1978) suggested that bilingual instruction may exert a cumulative effect. Thus, an early exit into English-only programs may prevent minority students from benefitting fully from the bilingual program. Cummins (1981a) has also suggested that a quick-exit transitional policy is often educationally inappropriate in view of the fact that it takes minority students considerably longer to attain age-appropriate levels in English academic skills than in face-to-face communicative skills.

In summary, the positive reviews claim that there are sufficient individual bilingual program evaluations showing significantly improved academic progress by minority students to support the general educational principles underlying bilingual education. The relative paucity of positive evaluations in comparison to the total number of bilingual programs in operation is attributed to (i) the recency of bilingual programs, (ii) initial implementation difficulties and (iii) lack of usable data from the majority of program evaluations because of methodological inadequacies.

The Baker/de Kanter report was released in September 1981 and its major conclusions are very different from those of previous research reviews. These conclusions are as follows:

- "Schools can improve the achievement level of language-minority children through special programs.
- The case for the effectiveness of transitional bilingual education is so weak that exclusive reliance on this instructional method is clearly not justified . . . Therefore . . . each school district should decide what type of special program is most appropriate for its own unique setting.
- There is no justification for assuming that it is necessary to teach nonlanguage subjects in the child's native tongue in order for the language-minority child to make satisfactory progress in school . . .
- Immersion programs, which involve structured curriculums in English for both language and nonlanguage subject areas, show promising results and should be given more attention in program development" (de Kanter and Baker, Education Times, October 5, 1981).

The many critiques and rebuttals of the Baker/de Kanter report are too extensive to discuss in detail here (see e.g. American Psychological Association, 1982; Willig, 1981/82). Common to all the critiques, however, is an objection to Baker and de Kanter's definition of what constitutes methodologically-acceptable data.

From the present perspective most of the literature on the effectiveness of bilingual education suffers from a serious flaw insofar as reviewers have used a "weighing scale" approach to judging whether, on balance, bilingual education has been effective. Thus, studies have been categorized as supporting or refuting

"bilingual education" itself rather than particular theoretical hypotheses regarding the effects of different types of instructional programs. This is one of the major reasons, we would argue, that most policy-makers consider research evidence on bilingual education to be lacking and what is available to be of poor quality. When the research findings are interpreted in the light of specific theoretical hypotheses, the apparent inconsistencies in the data disappear.

Both the Baker and de Kanter (1981) and Cummins (1983) reviews show that in the vast majority of programs minority students instructed bilingually performed at least as well as those who received all their instruction through the majority language. In a substantial number of cases (but by no means all) students in the bilingual program evidenced better academic achievement than those in monolingual programs.

These results are capable of different interpretations depending upon the questions that are being asked. Baker and de Kanter's question, for example, was whether transitional bilingual education (TBE), as a federally-sponsored program, was achieving its goal of improving minority children's achievement. Thus, evaluations in which no differences were found between bilingually- and monolingually-instructed students were interpreted as negative results for bilingual education. Hence, their conclusion that TBE is not working well enough to make it the exclusive program eligible for federal funding.

One of the major problems with the Baker and de Kanter report is their classification of vastly different types of programs in the same category (e.g. "transitional bilingual education", "structured immersion"). The potential confusion from a policy perspective can be seen from the fact that the "immersion" program in McAllen, Texas, that Baker and de Kanter describe as a "promising" alternative to transitional bilingual education actually involved more L1-medium instruction (about 60 minutes per day) than probably most transitional bilingual programs in the U.S.

The data reviewed by Baker and de Kanter and other investigators are, in fact, easily interpreted when the right questions are asked, specifically questions relating to the theoretical assumptions that inform policy. Thus, the data clearly show that the linguistic mismatch hypothesis is inadequate as a basis for

policy since in many cases students who experienced a linguistic mismatch performed as well academically as those who received some form of bilingual instruction.

The maximum exposure hypothesis is also clearly refuted by the data since in almost no case did students instructed bilingually perform worse than those who received all their instruction through English. The one or two cases where this did happen (e.g. Cohen, 1975) can be attributed to the use of a very inferior form of bilingual education, namely, the concurrent or translation method. It is sobering that arguments reflecting the maximum exposure hypothesis continue to play a major role in the policy debate in view of the fact that it is refuted by virtually every evaluation of bilingual education (including immersion programs) ever conducted.

The interdependence hypothesis is the converse of the maximum exposure hypothesis and is consistent with the evaluation results from bilingual programs implemented in vastly different contexts both in the U.S. and abroad (see Cummins, 1983).

An important point in relation to the interpretation of U.S. evaluation results is that the adequacy of the evaluation design depends on the question that is being asked. Thus, designs that may not be adequate to show that bilingual education worked better than a monolingual program (e.g. because of inadequate control over possible initial group differences) may be quite adequate to test the interdependence hypothesis. The prediction that derives from this hypothesis is that students will perform as well in English academic skills despite considerably less instructional time through the medium of English. Thus, all the findings of no difference between bilingual and monolingual groups are consistent with the interdependence hypothesis. Also, studies where bilingual program students attained grade norms and/or outperformed comparison groups (e.g. the Rock Point study rejected by Baker and de Kanter because of inadequate matching of experiemental and comparison groups) also support the interdependence hypothesis although the designs may not be adequate to demonstrate that the bilingual program is <u>more</u> effective than the monolingual comparison program.

The policy implication of this pattern of findings is that students in bilingual programs will perform at least as well as those in monolingual programs. The finding that in many cases bilingual programs appear more effective than monolingual programs is probably due as much to sociological as to linguistic factors (see Cummins, 1981a; Ogbu, 1978). Although these sociological factors are as yet inadequately understood, they must nevertheless be taken into account if the policy mistakes of the past (e.g. basing bilingual education exclusively on the linguistic mismatch hypothesis) are to be avoided.

In summary, the data do not demonstrate the effectiveness of bilingual instruction in any absolute sense, nor do they show that L2-only programs are necessarily ineffective or inappropriate. However, the data do clearly dispel the myth that L1-medium instruction will impede the acquisition of English academic skills. As with the Canadian immersion results, some form of transfer of academic skills across languages must be invoked to explain the fact that less time through the medium of English exerts no negative effects on the development of English academic skills.

The pattern of findings from previous studies is clearly compatible with the SEDL results and suggests that bilingual education programs can improve students' academic performance, and will certainly not result in any disadvantage to students compared to monolingual programs.

In the next section the research basis for the interdependence hypothesis is further explored by examining studies which investigated the learning of L2 academic skills by immigrant students.

VI. PREVIOUS STUDIES OF LINGUISTIC INTERDEPENDENCE AMONG IMMIGRANT STUDENTS

One of the major influences on the U.S. bilingual education debate was the series of studies of Finnish immigrant students in Sweden reported by Skutnabb-Kangas and Toukomaa (1976). These investigators reported a substantial relationship between immigrant students' L1 cognitive /academic skills and their subsequent acquisition of Swedish cognitive/academic skills. Students who immigrated to Sweden after several years schooling in Finland (i.e. 10-12 year olds) seemed to have better academic prospects than either Finnish students born in Sweden or students who immigrated at younger ages. The findings of this study were highlighted by Troike (1978) and Cummins (1979) in arguing for the importance of promoting minority students' L1 academic skills within the context of U.S. bilingual programs.

The SEDL study addresses this underlying theoretical issue among Cantonese-speaking immigrant students. Since the Skutnabb-Kangas and Toukomaa report, several other studies have investigated determinants of the achievement of immigrant students, but few within the context of U.S. bilingual education policy. The findings of all of these studies will be reviewed here in order to assess the extent to which they are consistent with the SEDL results. To the extent that findings from several very different contexts converge, confidence in the generalizability and applicability of the theoretical principles increases.

The central issue in reviewing these studies is the extent to which their findings are consistent with the principle of linguistic interdependence. The issue of how long it takes immigrant students to approximate grade norms in English academic skills will also be examined. Where findings appear inconsistent across contexts, ways of reconciling the findings will be discussed. This will involve an examination of the interaction between socio-cultural and cognitive/academic influences on the development of bilingual proficiency among language minority students.

Finnish Immigrant Students in Sweden (Skutnabb-Kangas and Toukomaa, 1976)

The data reported in the Skutnabb-Kangas and Toukomaa (1976) report and in a subsequent paper by Skutnabb-Kangas (1979) derived from empirical studies conducted by Toukomaa and other researchers in Sweden in the early 1970's. In general, these studies were designed to determine the level of Finnish students' academic achievement in both Finnish and Swedish and to explore some of the determinants of achievement. A variety of tests in both languages were used; most assessed cognitive/academic proficiencies such as vocabulary knowledge, synonyms, antonyms, etc. as well as academic achievement in reading, math and other school subjects.

L1/L2 Achievement. Finnish students in grades 4-6 were found to perform, on average, approximately 3-4 years behind Finnish students in Finland in Finnish verbal-academic skills (Table 4, p. 51). The retardation was somewhat less in Gothenburg than in Olofstrom, probably as a result of the fact that students in Gothenburg had received two hours teaching of Finnish per week for several years. However, the authors comment that the considerable retardation still evident in the Gothenburg sample suggests that two hours per week of L1 instruction is "totally insufficient" (p. 52) to maintain or adequately develop L1 skills among immigrant pupils.

Performance on Swedish verbal-academic tests was similarly poor. In the Olofstrom sample the average level of Finnish students was equivalent to the lowest 10 percent of native Swedish speakers. On a listening comprehension test administered in Gothenburgh (Jauho and Loikkanen, 1974) only 15 percent of 191 Finnish students tested performed as well as their Swedish peers. In the same sample, 25 percent of grades 2-6 Finnish pupils performed as well as the Swedish average on what Skutnabb-Kangas and Toukomaa describe as a "very simple" reading comprehension test. On non-verbal cognitive measures, minority students performed close to Swedish and Finnish norms.

The picture that emerges from several studies using a variety of verbal-academic measures is of low academic performance by Finnish minority students in both Finnish and Swedish. Other data reported by Skutnabb-Kangas and Toukomaa, however, suggest that Finnish students consider their everyday conversational skills in Swedish to be quite fluent.

Factors That Influence Achievement. The influence of four major interrelated factors was examined in the Skutnabb-Kangas and Toukomaa reports: age on arrival (AOA) in Sweden, length of residence (LOR), Finnish cognitive/academic proficiency, and Finnish instruction in Sweden.

The Swedish performance of grades 3-6 Finnish minority students in Gothenburg (Jauho and Loikkanen, 1974) on reading and listening comprehension tests is presented in Tables 9 and 10. The following quotation from Skutnabb-Kangas and Toukomaa explains the meaning of the scores in Tables 9 and 10.

> "A test average of 1 means that the pupils cannot manage even the simplest Swedish-language tasks and the highest point average 5 that the pupil has the normal Swedish-language skills of his Swedish peers as far as reading simple daily language goes. A pupil with the average 3 already needs remedial Swedish instruction. The table combines the two lowest and two highest values.
>
> According to the Table (9), those who attended school in Finland approached the level of achievement of normal Swedish pupils (4 or 5) in the written comprehension test considerably more often than those who began school in Sweden. Those who had attended school in Finland for at least three years did best. The explanation for this can perhaps be found in their better skill in their mother tongue, which laid the basis for understanding a test written in Swedish. Two years in a Finnish class in Sweden did not, on the other hand, make for as good a basis for learning Swedish as the corresponding time in Finland." (1976, p. 65-66)

The results for listening comprehension presented in Table 2 appear to show somewhat more optimistic trends in terms of the efficacy of Finnish classes and less optimistic results for Swedish classes. It is not clear, however, what the relationship is between Group 1 in Table 10 and the Swedish-only schooled students in Table 9; for example, whether the Table 10 data also derive from the Jauho and Loikkanen study or whether more subjects were added at a later date.

Baker and de Kanter have applied statistical tests to the data presented in Tables 9 and 10 and conclude that although the group differences in listening comprehension are statistically significant, those in reading comprehension are generally not. In terms of the interdependence hypothesis, Baker and de Kanter's analyses are irrelevant because the issue is <u>not</u> whether those who attended school in Finland performed better than those who attended school only in Sweden. The issue is why students who have considerably less exposure to

TABLE 9

Results of Reading Comprehension Tests According to Location of School (Jauho & Koikkanen 1974)

Points	Attended school only in Sweden		Attended school in Finland	
	Swedish-lang. class	Finnish-lang. class	1-2 yrs	3 yrs or more
1-2 (-)	12%	5%	14%	6%
3	26%	40%	11%	12%
4-5 (+)	62%	55%	75%	82%
Total	100%	100%	100%	100%
N	65	40	28	17

(From Skutnabb-Kangas and Toukomaa, 1976, Table 13, p. 65)

TABLE 10

Listening Comprehension (Swedish) and the Language of School Entry

	Group 1 Started school in Sweden		Group 2 Started school in Finland School years in Finland before emigration	
Test mark (%)	Swedish classes	Finnish classes	1-2	3 or more
1-2 (-)	12	4	14	12
3	50	33	17	12
4-5 (+)	38	63	69	76
	100%	100%	100%	100%
N	82	49	29	17

(From Skutnabb-Kangas, 1979, Table 7, p. 16)

Swedish perform as well or better than students who have had much greater exposure to Swedish.

Other analyses presented by Skutnabb-Kangas and Toukomaa (1976) suggest answers to this question. First, length of residence (LOR) was found to be important for the development of both Finnish and Swedish language skills in that skills in Finnish decrease and skills in Swedish increase with time (1976, Table 7, p. 57). The negative impact of LOR on Finnish increased across the grades. In grade 6, for example, the correlations between LOR and four measures of Finnish verbal-academic skills in both Olofstrom and Gothenburg ranged from -.42 to -.72 with six out of eight correlations in the -.6 to -.7 range. In grade three, by contrast, correlations ranged from -.16 to -.33. Correlations of LOR with Swedish verbal-academic skills between grades 2 and 6 in Olofstrom ranged from .34 to .57 while with teacher evaluations of Swedish language skills in Gothenburg the correlations of LOR ranged from .56 to .83. In the latter case all correlations between grades 4 and 6 were greater than .8. Skutnabb-Kangas and Toukomaa attribute the greater impact of LOR on teachers' evaluation of Swedish skills to "the fact that teachers have been found to pay more attention to verbal fluency and the right accent than to understanding concepts" (1976, P. 58).

The fact that the negative correlations of LOR with Finnish verbal-academic skills were greater than the positive correlations of LOR with Swedish verbal-academic skills led Skutnabb-Kangas and Toukomaa to conclude that L1 verbal-academic skills regress faster than L2 verbal-academic skills are acquired. This does not appear to be a general pattern among immigrant students, however, since the SEDL study and other studies (e.g. Cummins et al, in press) found a greater impact of LOR on the acquisition of L2 academic skills than on the regression of L1 academic skills. Community and educational support for the maintenance of L1 may be a factor in explaining these differences across contexts. However, Skutnabb-Kangas and Toukomaa's finding that L2 conversational fluency develops largely as a function of exposure to L2 is supported by the findings of other studies (see below).

Skutnabb-Kangas and Toukomaa also report data derived from questionnaires (N= 166) administered to grades 3-6 students in Olofstrom regarding which language

they felt they spoke better. Seventy-three percent of students with LOR 3-4 reported that they spoke Finnish better compared to only 18 percent of students with LOR 5-6. Thus, a dramatic language shift appears to occur between LOR 4 and 5.

The significance of LOR for Swedish verbal-academic skills development would lead us to expect that students initially schooled in Finland would perform worse in Swedish than students who attended school only in Sweden (Tables 9 and 10) since the former group has had less exposure to Swedish. Clearly, other factors are at work. Several analyses reported by Skutnabb-Kangas and Toukomaa on both the Olofstrom and Gothenburg data suggest that one of these factors is the level of cognitive/academic proficiency students had attained in Finnish. For example, among the Olofstrom sample, the partial correlations between Finnish and Swedish verbal-academic proficiency for grades 3-6 students (N=165) with LOR held constant (Table 8, p. 60) ranged from .20 ($p<.01$) to .41 ($p<.001$). The partial correlation of age on arrival (AOA) for this group was .23 ($p<.01$). For the grades 7-9 sample (N=78), five out of six partial correlations of Finnish with Swedish verbal-academic skills were significant. This pattern of significant L1-L2 partial correlations was also found for the Gothenburg sample (Tables 11 and 12, p. 64 and 64).

The interdependence of L1 and L2 verbal-academic skills is further indicated by the fact that Finnish skills correlated about as highly with subject matter achievement as did Swedish skills, despite the fact that all subjects were taught in Swedish (Table 14, p. 68).

Analysis of sibling data also supports the authors' conclusions regarding the effects of AOA. They also cite a study by Gustavsen (1974) involving 183 migrant students in Oslo which showed that students who immigrated at older ages made more rapid progress in acquiring Norwegian than students who immigrated at younger ages (i.e. early grades of elementary school).

Criticism of the Skutnabb-Kangas and Toukomaa Report. The main criticisms of the Skutnabb-Kangas and Toukomaa report concern the validity of the test instruments used (Brent-Palmer, 1979; Carlson, 1981; Ouvinen-Birgerstam and Wigforss, 1978) and the interpretation of the statistical analyses (Baker and de

Kanter, 1981). These criticisms need not be discussed in detail here. It is sufficient to note that Skutnab-Kangas and Toukomaa made no attempt to assess the totality of "language proficienty" in either L1 or L2; what they did attempt to assess were academically-related aspects of language proficiency both with tests standardized for Finnish and Swedish students amd measures designed specifically for immigrant students (Jauho and Loikkanen, 1974). Most of the objections to the measures employed are not based explicitly on any theory about how language proficiency relates to academic achievement, whereas the Skutnabb-Kangas and Toukomaa data can be readily interpreted from within such a theoretical framework (Cummins, 1981a).

Baker and de Kanter do not dispute the validity of the correlational analysis presented by Skutnabb-Kangas and Toukomaa, although they do question its interpretation. Similarly, their objections to the data presented in Tables 1 and 2 concern interpretation rather than validity and do not alter the fact that these data support the interdependence hypothesis.

While the objections to the Skutnabb-Kangas and Toukomaa report do not seriously affect the confidence that can be placed in the pattern of findings that emerged consistently across several sites, some other points of qualification can be briefly noted. First, the use of the emotively-laden term "semilingualism" to characterize the relatively low L1 and L2 verbal-academic skills of Finnish immigrant students is unfortunate, and, I believe, accounts for some of the vehemence with which the Skutnabb-Kangas and Toukomaa findings have been rejected (e.g. Brent-Palmer, 1979). Second, interpretation of their data would have been facilitated had more powerful statistical tools such as multiple regression analysis been used. Third, although the interdependence of L1/L2 cognitive/academic proficiencies which is supported by the data is consistent with the policy recommendation that minority students' L1 be strongly promoted, it does not imply that this approach is the only valid one nor that it is necessarily the most appropriate one for all language minority groups under all conditions. In other words, although clarification of the theoretical constructs helps explain the development of bilingual proficiency, it does not substitute for actual evaluation of bilingual programs themselves.

In summary, the Skutnabb-Kangas and Toukomaa data are consistent with the SEDL findings in showing: 1. that immigrant students whose L1 cognitive/academic skills were better developed made more rapid progress in acquiring L2 cognitive/academic skills; 2. that LOR is significantly related both to L2 cognitive/academic skills and to L2 conversational skills, the relationship in the latter case being greater.

Achievement of Immigrant Students in Canada (Cummins, 1981b)

The effects of age on arrival (AOA) in Canada and length of residence (LOR) on English achievement were investigated by reanalyzing data from the 1969 Every Student Survey carried out by the Toronto Board of Education (Wright and Ramsey, 1970). This survey involved 25% of the Toronto Board's grade 5, 7 and 9 classrooms which included over 1,200 immigrant students from a large variety of language backgrounds. Wright and Ramsey reported that students who arrived in Canada after the age of 6 performed progressively worse in English academic skills in relation to grade norms than those who arrived at younger ages. However, the possible effects of LOR were not investigated.[1]

Based on the data presented by Wright and Ramsey, it was possible to estimate LOR and separate the effects of AOA and LOR on immigrant students' performance both in relation to grade norms and in terms of absolute scores. The measures of English proficiency consisted of a group adaptation of the Ammons Picture Vocabulary Test (PVT) and a six-part test of English language skills developed by the board. Only the PVT data are reported here since the same patterns emerged on all the measures tapping English cognitive/academic proficiency.

Figure 2 shows that it took immigrant students who arrived in Canada after the age of 6-7 between 5 and 7 years, on the average, to approach grade norms in English cognitive/academic skills. However, with the exception of the LOR:5 group there is no clear tendency in relation to AOA. Given the difference in cognitive/academic skills between older and younger native English speakers, we would thus expect the older immigrant students to have learnt more English in absolute terms than younger students. This is confirmed in Figure 3 where older students show clear superiority in PVT scores when LOR is controlled. The

extent of the advantage can be seen in the fact that the AOA:14-15 group scores higher on the vocabulary test after 1 year LOR than the AOA:4-5 group does after 7 years LOR (27.1 vs. 26.3). Older AOA students performed better than younger AOA students on 89 out of 90 comparisons involving English cognitive/academic performance.

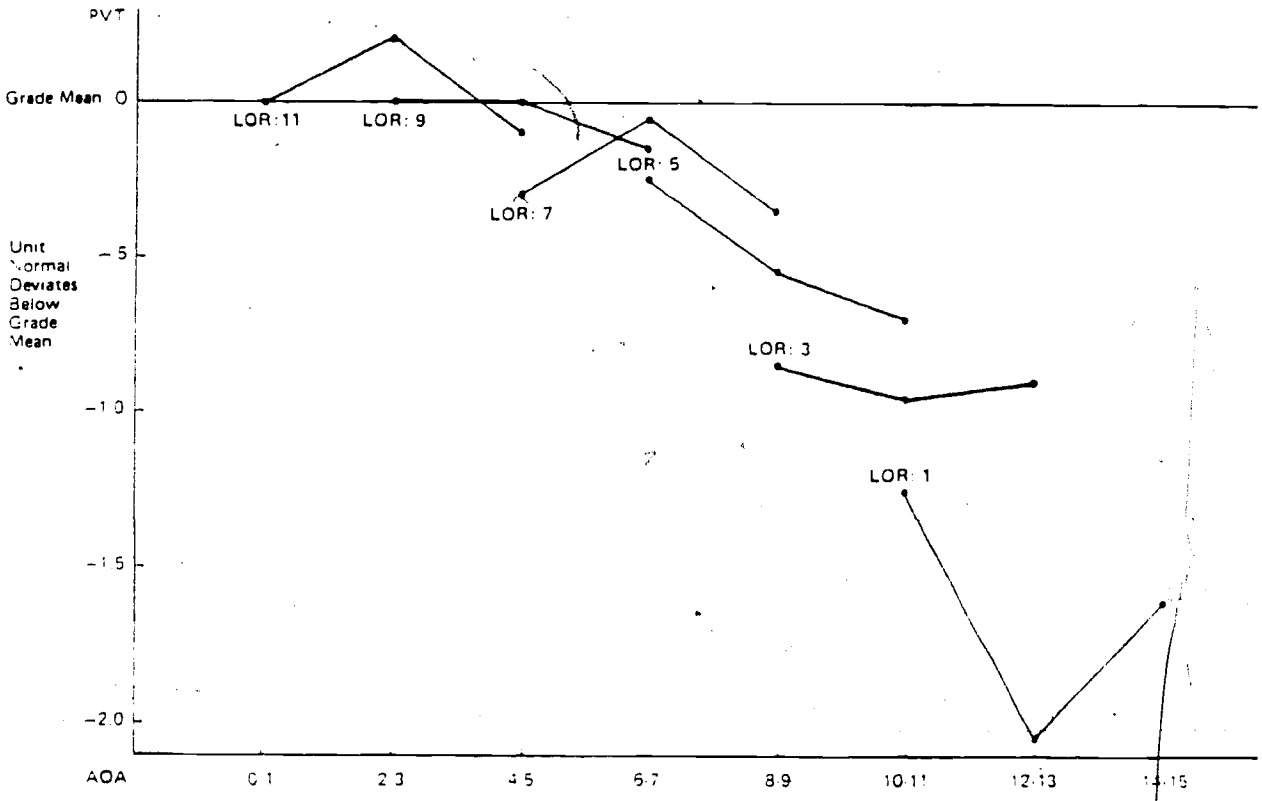


Figure 1: Age on Arrival, Length of Residence, and PVT Standard Scores

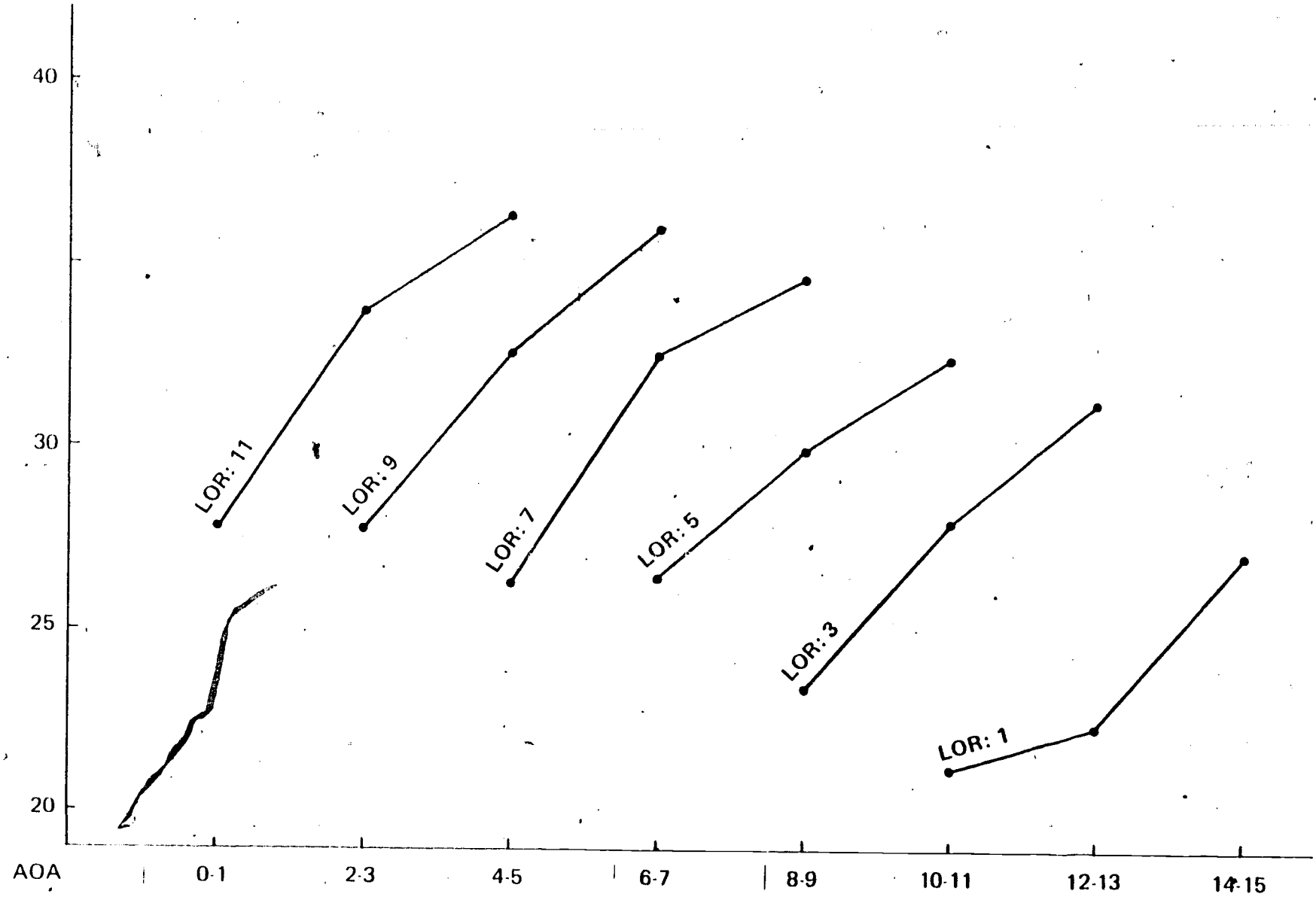


Figure 3. Age on Arrival, Length of Residence, and PVT Raw Scores

These data are consistent with many other recent studies showing that older learners make more rapid progress than younger learners in acquiring cognitive/academic aspects of L2 skills (see reviews by Cummins, 1981b; Ekstrand, 1979; Krashen et al, 1979) and are clearly consistent with the interdependence hypothesis.

The Toronto data show many consistencies with both the SEDL and Scandinavian data, but also some differences. First, all three studies support the interdependence hypotheses in showing that older immigrant students whose L1 cognitive/academic proficiency is better developed make more rapid progress than younger immigrant students in acquiring L2 cognitive/academic proficiency. Second, all three studies suggest differences between L2 cognitive/academic and L2 conversational skills in terms of patterns of acquisition. It is clear both from several studies (see Cummins, 1981b) and anecdotal accounts that it normally takes considerably less than 5-7 years for immigrant students to acquire fluent native-like L2 communicative skills in everyday situations. Third, LOR is clearly a major determinant of L2 acquisition in all three contexts.

An important difference between the Scandinavian data and both the Toronto and SEDL studies, however, is the fact that immigrant students in these latter contexts tended to approximate grade norms in L2 cognitive/academic skills given sufficient LOR. Students who arrived before school age (with consequent long LOR) were performing at grade level. The academic success of students under home-school mismatch conditions in the Toronto situation is clearly inconsistent with the linguistic mismatch hypothesis. By contrast, however, Finnish students who were born in Sweden or arrived at an early age tended to reach a low level plateau in the development of L2 cognitive/academic skills and generally to have worse academic prospects in both L1 and L2 than students who immigrated at older age levels. In short, whereas in the Toronto and SEDL data, LOR appears to exert a strong linear effect until students' L2 skills approximate native L2 norms, in the Scandinavian data the impact of LOR is less overall and appears to be considerably modified by sociocultural variables. A large-scale study carried out in Sweden (Ekstrand, 1979) also suggests that the influence of LOR can be modified by other factors.

Ekstrand's (1979) Study of Immigrant Students in Sweden

Ekstrand's data were collected in the late sixties from a large number of school boards in Sweden. Overall, more than 2,000 students were assessed on 22 test variables designed to measure Swedish and more general cognitive skills. The general pattern of results was that students' age was a more powerful predictor of L2 performance than was LOR. Age was independent of LOR in its effect. On only two variables did LOR exert a strong impact, namely, pronunciation and free oral production. On cognitive/academic measures the effect of LOR was weak and tended not to exert any influence beyond one or two years LOR.

Ekstrand's study supports the distinction made between cognitive/academic and conversational L2 skills and it indirectly supports the interdependence hypothesis by showing older students whose L1 literacy skills have been promoted in schools in their home countries to make more rapid progress in acquiring L2 cognitive/academic skills. The study is somewhat inconsistent with other studies in reporting only a weak effect of LOR on L2 cognitive/academic development.

Russian Immigrant Students in the United States (Carlson, 1981)

Carlson investigated the effects of LOR and AOA among 31 Russian immigrant students in the U.S. who ranged in age from 7 to 18 (mean 11.8). The range in AOA was from 6 to 17 (mean 9.7) while LOR ranged from 11 to 60 months (mean 27.2). The dependent variables were subtests from the Language Assessment Scales (LAS), school grades and the Peabody Vocabulary Test.

LOR was found to exert a stronger impact on English skills acquisition than AOA. The correlations of AOA with English skills (partialling out LOR) and of LOR with English skills (partialling out AOA) are shown in Table 11. The Pearson correlation of grade in English classes with AOA was also significant (r= .32*) but partial correlations involving this variable are not reported.

TABLE 11

PARTIAL CORRELATIONS OF ENGLISH VARIABLES WITH AOA AND LOR AMONG RUSSIAN IMMIGRANT STUDENTS (N=31)
(Adapted from Carlson, 1981, Table 12, p. 52)

LAS	AOA (with LOR controlled)	LOR (with AOA controlled)
Minimal pairs recognition	.51*	.45**
Lexical production	.02	.51**
Phoneme production	-.33*	.52**
Sentence comprehension	.22	.48**
Oral production	-.19	.70**
Peabody Vocabulary	.35*	.85*

* $p<.05$
** $P<.01$

Carlson suggests that the significant correlations between AOA and both minimal pairs recognition and Peabody Vocabulary are related to cognitive developmental and metalinguistic factors, an interpretation supported by the fact that correlations between Age and these measures among an American monolingual sample of similar age range (N=31) were considerably higher (r=.7 for minimal pairs, .8 for Peabody vocabulary) than between Age and the other English variables (range: .14 - .45). The lowest correlation was between phoneme production and Age (r=.14). There was no variability in the oral production of the American sample and thus this variable was excluded from the correlational analysis. Analyses of variance for the total sample (Russian and American) also showed that Age exerted a significant positive effect (older students better) for minimal pairs and Peabody Vocabulary and a negative effect for phoneme production.

Carlson interprets the negative correlation between AOA and phoneme production as evidence of a critical prepubertal period for acquiring a native-like accent. Her data, however, do not support this interpretation. Age (and AOA) exerts an impact only for students with LOR less than 24 months. Immigrant

students with LOR greater than 24 months performed equivalently to the American group and, among this group, no differences were evident between immigrant students older and younger than 11 years. These data suggest that older immigrant students take longer than younger students to acquire a native-like accent but attain similar ultimate levels. Carlson also reports that parental English proficiency was strongly (r range: .34 - .67) related to all aspects of English proficiency.

Several aspects of Carlson's study are interesting. First, her finding of negative correlations between AOA and both phoneme production and oral production is consistent with Ekstrand's (1979) finding that it was only on measures assessing these two skills that older immigrant students did not show a significant advantage in comparison to younger students. As in the studies already discussed, these differences suggest the need to distinguish between L2 cognitive/academic and everyday conversational L2 skills.

Second, the significant relationship between AOA and cognitive/academic aspects of L2 proficiency are consistent with the findings of studies already discussed and with the interdependence hypothesis.

Third, LOR exerts similarly strong effects to those observed in the Seattle and Toronto data and stronger effects than in the Scandinavian data.

Fourth, the fact that Russian students attained high levels of L2 skills within two years of arrival suggests that the average figure of 5-7 years LOR which Cummins (1981b) reported can be shortened considerably for students who are of relatively high SES and highly motivated to succeed, as presumably the Russian immigrant students are.

Carlson regards her findings as challenging the validity of those reported by Skutnabb-Kangas and Toukomaa insofar as no evidence of "double semi-lingualism" was found among the Russian students despite rapid language shift. She also regards the greater impact of LOR in comparison to AOA as inconsistent with the Scandinavian findings.

However, as pointed out in discussing the Toronto data, different patterns of findings across contexts do not imply that one set of findings is necessarily invalid. Differences are more likely to point towards the influence of socio-cultural factors in modifying the effects of individual factors such as AOA and LOR. The differences in the LOR required to attain L2 norms in cognitive/academic proficiency between the Scandinavian, Toronto and Pittsburg (Russian) groups clearly illustrate the role of sociocultural factors.

What is more remarkable is the robustness of predictions derived from the interdependence hypothesis in accounting for the data in these very different language acquisition contexts.

Linguistic Interdependence Among Japanese and Vietnamese Immigrant Students in Canada (Cummins et al, in press)

Cummins et al set out to explicitly test the interdependence hypothesis among 91 Japanese and 45 Vietnamese students in Toronto. The Japanese students were sons and daughters of temporary residents who were in Canada for business or professional reasons whereas the Vietnamese sample consisted of refugee students. The Japanese students attended a part-time (Saturday) Japanese school and were selected from grades 2-3 and 5-6 (Canadian school grades) in order to allow the effects of AOA and LOR to be separated. Thus, a grade 2 student with two years LOR has an AOA of about 5 years whereas the AOA for a grade 6 student with two years LOR is about 9 years. All the Vietnamese sample were recent arrivals (LOR 5-22 months) and ranged in age between 9 and 17 years. Thus, all the sample had received at least some education in Vietnamese prior to arrival in Canada.

The dependent variables for the Japanese group consisted of five English cognitive/academic measures (two reading measures from the grade 2 Gates McGinitie reading test and three oral tasks) and measures derived from ratings of student interviews (administered to a subsample, N=59). Students were also interviewed in Japanese and were administered a Japanese standardized diagnostic reading measure.

Because the interview ratings in English were similar to those devised for the SEDL study, the factor analysis (N=59) of English measures among the Japanese students is shown in Table 12. The first two factors are similar to factors obtained for the Chinese students in Seattle; factor 1 represents grammatical competence while factor 2 represents an interactional style dimension related to the amount of detail and elaboration volunteered by students in the interview (e.g. conversational "richness"). The third factor is English cognitive/academic proficiency.

Analyses involving partial correlations, multiple regression and sibling comparisons were all found to support the interdependence hypothesis insofar as both AOA and Japanese Academic Proficiency were significantly related to English cognitive/academic proficiency. Regression analyses of the three English factor scores on individual exposure (LOR) and attribute variables are shown in Table 13. JFAC1 represented the first Japanese proficiency factor which had high loadings on aspects of proficiency similar to those of both factors 1 and 2 for the English analysis. It can be seen that the L1 cognitive/academic variables (Japanese Academic Proficiency, AOA and Age) collectively account for 18 percent of the variance in English cognitive/academic proficiency, one percent more than LOR. However, LOR accounts for 26 percent of the variance in EFAC1. Non-cognitive individual attributes accounted for a substantial amount of the L2 interactional style (EFAC2) variance.

In terms of Japanese proficiency, LOR was found to exert a negative effect only on fluency, pronunciation and, to a lesser extent, Japanese Academic Proficiency. Other aspects of Japanese proficiency (e.g. conversational syntax and "richness") were strongly related to personality, age and AOA. Students' academic performance was close to the average level (mean=50.0) for Japanese students in Japan (grades 2-3 mean=48.2; grades 5-6 mean=55.3). There was considerable evidence that level of Japanese proficiency on arrival in Canada was an important factor in maintaining the language.

In short, the pattern of relationships between AOA and both L1 maintenance and L2 acquisition are similar to those reported by Skutnabb-Kangas and Toukomaa; the difference, however, is that the Japanese students appear to be performing well in both English and Japanese academic skills whereas low levels of both Swedish and Finnish characterize Finnish students in Sweden.

TABLE 12

FACTOR ANALYSIS OF ENGLISH ACADEMIC LANGUAGE PROFICIENCY (17-21) AND INTERVIEW MEASURES (1-16)

Variable		Varimax Rotation 1	2	3
1.	Pronunciation	.69	.26	.22
2.	Interviewer Speech	.48	.41	.32
3.	Semantically Appropriate Responses	.17	.58	.22
4.	Conversational Richness	.16	.77	.08
5.	Inflections of Verbs and Nouns	.67	.14	.35
6.	Article Use	.73	.24	.20
7.	Conversational Syntax	.80	.23	.44
8.	Ease	.23	.73	.16
9.	Tel. Question Formation	.57	.31	.26
10.	Tel. Appropriateness	.39	.20	.48
11.	Picture Sequence Syntax	.73	.22	.44
12.	Picture Sequence Cohesion	.14	.62	.52
13.	Picture Sequence Richness	.32	.66	.31
14.	Picture Description Syntax	.81	.36	.33
15.	Picture Description Strategies	.56	.49	.25
16.	Picture Description Richness	.40	.64	.24
17.	English Vocabulary	.46	.33	.72
18.	English Reading	.39	.29	.84
19.	English Prepositions	.34	.17	.50
20.	English Antonyms	.52	.41	.60
21.	Sentence Repetition	.49	.50	.56

TABLE 13

Exposure and Attribute Predictors of English Proficiency

	EFAC 1			EFAC 2			EFAC3		
	R square	Rsq change	Beta	R square	Rsq change	Beta	R square	Rsq change	Beta
1. LOR	.26	.26	.54	.21	.21	.49	.17	.17	.73
2. Japanese Academic Proficiency	.26	.00	.13	.25	.04	.11	.26	.09	.25
3. AOA: Older Group	.28	.02	-.06	.27	.01	.06	.35	.09	.40
4. Age in Months	.29	.01	-.05	.27	.00	-.14	.35	.00	.08
5. Personality*	.30	.01	-.03	.32	.05	.09	.37	.01	-.09
6. JFAC 1	.33	.03	-.21	.44	.12	.42	.37	.01	-.11
7. Sex**	.36	.02	-.16	.48	.04	.21	.38	.01	.07

*5 point scale, 1 = very shy, 5 = very outgoing

**2 = Female, 1 = Male

Strong support for the interdependence hypothesis was also found among the Vietnamese sample where performance on a Vietnamese Antonyms test together with Age accounted for 61 percent of the variance in English Antonyms, compared to only 6 percent for LOR (entered first into the equation). The relatively small effect of LOR is clearly due to the restricted range of LOR in the Vietnamese sample.

Among the conclusions drawn by the investigators in this study were: 1. the interdependence hypothesis was strongly supported among two very different samples whose L1 bears little relationship to English; 2. not only are cognitive/academic aspects of L1 and L2 interdependent, L1/L2 interactional style is also interdependent insofar as it is a manifestation of personality attributes of the individual; 3. exposure to L2 in the environment and actual behavior in the language learning situation are strongly related to the acquisition of grammatical conversational skills but less related to cognitive/academic and interactional "attribute-based" aspects of L2 proficiency (EFAC2 and EFAC3).

A follow-up study (Fukushima, 1983) of 17 of the original Japanese sample about 30 months after the initial testing provided further support for the Interdependence hypothesis. Sixty-seven percent of the variance in English reading scores at Time 2 was explained by LOR, AOA and Japanese reading. LOR entered first into the equation accounted for 45% while the cognitive/academic block accounted for an additional 22%.

Japanese Kaigaishijo in New York (Iwasaki, 1981)

Using a similar sample of Kaigaishijo (children of temporary residents abroad) Iwasaki investigated the effect of age of intensive exposure to L2 and transferability of cognitive/academic skills across languages among grades 7 and 8 students in both full-time (N=76) and part-time (N=72) Japanese schools. The Gates McGinitie level E reading comprehension test was used as the dependent measure for English cognitive/academic proficiency. Unlike the Cummins et al study where raw scores were used, Iwasaki used Normal Curve Equivalent (NCE) scores normed for grade, with the result that age (grade) variance was removed from these scores. Different Japanese measures for grades 7 and 8 were developed by the investigator and were converted into grade-normed scores so that the grade 7 and 8 data could be combined.

For the most part, data were analyzed separately for students in part-time and full-time schools, thus considerably reducing the power of the analyses (school type could have been entered as the first variable in the regression analysis, allowing its effect to be ascertained in relation to other predictor variables). However, the general pattern of findings is clear and consistent for both groups. First, length of local schooling (strongly related to LOR) was the best predictor of English reading performance. For full-time students this variable accounted for 46 percent of the variance in English scores (df 1, 51). Second, grade-norm acquisition of English reading occurred at 27 months of local schooling, a figure similar to that reported by Carlson for Russian immigrant students. Third, relatively little variance was accounted for on Japanese scores compared to English scores (as with Cantonese L1 in the SEDL study). For part-time students, number of siblings was positively related to Japanese proficiency, possibly Iwasaki suggests, because students get to know more Japanese friends through siblings and also experience more active promotion of Japanese at home by parents. For full-time students, AOA was positively ($r=.27^*$) and LOR negatively ($r=-.24^*$) related to Japanese scores. This pattern is similar to that reported by Cummins et al.

The interdependence hypothesis was tested by means of correlational and multiple regression analyses as well as by Chi-square analysis. When the sample (both full- and part-time) was broken down by length of local schooling categories, significant correlations between Japanese and English scores were obtained for most categories. Multiple regression analysis (N=106) showed that length of local schooling accounted for 52 percent of the variance in English reading with Japanese score adding 9 percent ($p<.001$). The Chi-square analysis for 39 students with more than 27 months local schooling is shown in Table 14 (from Iwasaki, Table 22, p. 99).

In summary, the findings of this study, in combination with those already discussed, suggest that although the rate at which L2 cognitive/academic skills are acquired, and the ultimate level attained, varies considerably across socio cultural contexts, the transferability of L1/L2 cognitive/academic skills is found in a wide variety of linguistic and sociocultural contexts.

TABLE 14

TABLE OF FREQUENCIES RELATING ENGLISH SCORES TO JAPANESE SCORES

		English Scores 51-100		0-50		Total Number of Students
Japanese Scores	High and Medium	18		6		24
	Low	4		11		15
Total Number of Students		22	+	17	=	39 = N

$X^2 = 8.92^{**}$

** $p<.01$

Other Studies of Immigrant Students' L2 Acquisition

In this section a variety of studies which have some relevance to the policy issues addressed by the SEDL study will be briefly reviewed.

It is first appropriate to note that, consistent with the Skutnabb-Kangas and Toukomaa findings, there is considerable anecdotal evidence that immigrant students from Mexico fare better educationally than native-born Mexican Americans. Troike (1978), for example, stated that "it is a common experience that . . . children who immigrate to the United States after grade six . . . rather quickly acquire English and soon outperform Chicano students who have been in the United States schools since grade one" (p. 21). Based on a survey of school personnel in four southwestern states, Carter (1970) similarly reported that many teachers and administrators believed that older immigrant students achieved better than native-born Chicano students.

Findings reported by Baral (1979) appear at first sight inconsistent with these beliefs. Baral found that immigrant students who had at least two years of schooling in Mexico performed significantly lower in English academic skills than native-born Mexican students. However, the socioeconomic status of the immigrant students was significantly lower than that of the native-born students and their LOR (2-5 years) is likely to have been insufficient to catch up with the native-born students.

In a study involving 91 limited English proficient students, Connor (1983) reported that present grade level was the best predictor of English reading skills. Grade level was interpreted to include the age of the students as well as L1 reading skills acquired prior to immigration or in non-ESL classes. Length of residence was not found to exert a significant effect in this study, a finding which was attributed to the rapid progress of recently-arrived Vietnamese students in comparison to the relatively slower progress of Spanish-background students whose length of residence was longer. Connor attributes this pattern, in part, to the fact that

> "the Vietnamese-speakers were able to read in their L1, while the Spanish-speakers lacked these skills in their L1. . . . the Vietnamese children were recent immigrants while most Spanish-speaking subjects were long-term residents in the U.S. whose education had been in English-only programmes" (1983, p. 285-286).

The findings of this study again points to the influence of sociocultural factors on LOR and to the robustness of the effects of AOA and transfer of skills from L1 to L2.

A study of 48 grade 6 Japanese students on the U.S. west coast reported by Okamura (1981) is also relevant to the interdependence hypothesis. The subjects were similar to those in the Cummins et at (in press) and Iwasaki (1981) studies in that their parents were temporary residents in North America. Students were administered a Japanese academic proficiency measure, a non-verbal I.Q. test and also a questionnaire about language use and attitudes. They also rated their proficiency in English language skills. It was intended to administer an English reading test to students but it was possible to do this for only ten students. Although significant correlations were obtained between the English self-ratings and reading test performance, the small number of subjects involved considerably reduces the reliability of the results. Thus, it is not clear what aspects of English proficiency are in fact being assessed by the self-ratings.

Several of Okamura's results are of interest. First, the non-verbal I.Q. score did not relate significantly to either Japanese or English proficiency. Second, correlations between the English reading self-rating and Japanese proficiency were non-significant ($r=.11$, $N=48$) but the partial correlation holding length of U.S. schooling constant approached significance ($r=.24$, $p<.10$).[2] These results

offer only minimal support for the interdependence hypothesis but the weakness in the English proficiency variable makes the findings difficult to interpret.

Nomoto (1966) conducted a questionnaire survey involving 125 Japanese children residing in London, England. Parents were asked to give subjective evaluations of their children's competence in both Japanese and English in six areas of language skills: speaking, hearing, reading, writing, counting and thinking. Nomoto tried to determine the relevance of the children's age on arrival, length of residence, and home-language use on English and Japanese. Some of his findings were: (a) reading and writing skills in Japanese rapidly drop as children's period of residence becomes longer; (b) among the six areas, counting shows most clearly the shift of language competence from L1 (Japanese) to L2 (English). For the children who have already reached the age of eight before extensive exposure to L2, counting remains as the strongest area in Japanese, whereas for the children who were below eight at the time of entry, counting is the weakest of the Japanese language areas; (c) children's L2 becomes stronger than L1 when their period of stay extends longer than 2 1/2 years; (d) Japanese parents on the whole easily lose control over their children's home language. When they remain in London more than 2 1/2 years, 50% of the children use L2 even with parents and 69% of them use L2 among brothers and sisters. As for the arrival age, a child who is older than ten is the least affected in terms of Japanese retention but for the school child who is less than six on arrival, the degree of retention of Japanese as home language is less than 50%. These findings are consistent with Skutnabb-Kangas and Toukomaa (1976) and Cummins et al in suggesting that age on arrival is related to L1 retention and development.

A recent study of adult Hmong refugees in South East Asian camps in which it was possible to separate the effects of L1 literacy and previous formal education on the learning of English in a formal classroom setting also supports the interdependence hypothesis (Robson, 1981). It was found that both previous formal education and literacy in Hmong independently predicted progress in learning English. Students who had neither L1 literacy nor formal education learned little or no English (oral or written) from the formal classroom experience.

A study carried out in Sweden by Lofgren (1981) among Finnish minority students in a bilingual education program, as reported by Ekstrand (1983), showed that Finnish academic proficiency was the most powerful predictor of Swedish academic proficiency as well as of overall academic achievement at the grade 3 level (path coefficients were in the region of .7). This pattern of findings was confirmed in a larger sample of students at the grade 6 level.

Appel (1979) investigated the oral Dutch language proficiency of two groups of Turkish and Moroccan immigrant workers' children (N=57 in total) after eight months experience within two different school models. The first school model was an experimental transitional bilingual program in which Turkish or Moroccan Arabic was the main initial language of instruction while the second model was the normal program in which Dutch was the main (or sometimes only) language of instruction.

Appel reports that older immigrant students made more rapid progress in acquiring oral Dutch skills compared to younger students whose LOR (approximately 8 months) was similar. An evaluation of the program after three years (by which time the proportions of Dutch and L1 instruction were reversed) showed that students in the bilingual program were performing significantly better in Dutch academic skills (Altena & Appel, 1982). The consistency with the SEDL findings is apparent.

Finally, Fradd (1983) investigated the relationship between Spanish and English reading skills among 41 Cuban-background junior high school students in Florida. A significant correlation was obtained ($r=.43$, $p<.01$), consistent with the interdependence hypothesis.

In summary, there is a high degree of consistency in the research findings which tend to strongly support predictions derived from the interdependence hypothesis.

VII. SUMMARY AND CONCLUSIONS

The SEDL analytic study of language and literacy learning in bilingual education involved a relatively small number (N=112) of grades 4-6 elementary school students, from one ethnic group (Chinese) in a particular urban center (Seattle). The central issue for this policy report is whether it is possible to draw any conclusions from this study which are relevant not just to Chinese students in Seattle and to the particular type of bilingual program model they experienced, but which are generalizable to other language minority contexts and have implications for U.S. bilingual policy as a whole. This issue requires consideration of the role of theory in the policy-making process.

Research, Theory and Policy

Two immediate issues must be addressed by educators and politicians in assessing the policy implications of any study: first, to what extent are the research findings and interpretation of those findings valid? Second, to what extent do the findings have any broader application or generalizability beyond the specific context in which the research was carried out?

Within the context of U.S. bilingual education policy, the first issue has been extremely controversial since the publication of the Baker and de Kanter (1981) report. Baker and de Kanter disputed the interpretation of many studies that previously had been viewed as supportive of bilingual education policy, and debate has since continued on the appropriate criteria of methodological adequacy in assessing bilingual education evaluations.

Although it is clearly crucial to assess the methodological adequacy of the research design and consequent validity of the results, meaningful interpretation of those results requires consideration of the theoretical principles which permit explanation and generalizability to other contexts. The omission of any coherent examination of <u>theory</u> in bilingual education is a major flaw which limits the usefulness of the Baker and de Kanter literature review. Methodological adequacy is a necessary condition for research to become relevant to educational policy; however, contrary to the implicit assumption of many commentators, it is not a sufficient condition. Policy-making requires that research findings, often

from very different contexts, be related to one another and integrated into a coherent theory. Individual findings or sets of findings are useless in isolation since they cannot be generalized across contexts. Thus, for example, the results of the Chinese bilingual program in Seattle, evaluated in the SEDL study, cannot be directly generalized to programs for Puerto Rican students in New York. However, the theoretical principles underlying the success of these two programs are, almost by definition, generalizable across contexts. If a particular theoretical hypothesis cannot account for a set of research findings, then the hypothesis is inadequate and requires revision.

In short, part of the controversy and confusion in regard to the evidence for and against bilingual education derives from the fact that educators and policy-makers have tried to go directly from research findings to policy implications, without the intermediate step of theoretical interpretation of the findings. Elucidation of the explanatory principles that underlie the research findings in very different contexts is necessary before any policy implications can be drawn from the findings. This becomes clear in examining the policy implications of the SEDL study.

Major Findings of the SEDL Study

The SEDL study used regression analysis to examine the influence on English proficiency of variables such as length of residence in the U.S., amount of bilingual or mother tongue (L1) instruction received, current and preschool exposure to English, L1 academic proficiency, etc. The study is exceptional in the range of both academic and oral communicative measures of English proficiency it included. It also differs from many previous evaluations in defining bilingual instruction as a continuous variable (i.e. number of semesters) rather than as a discrete program. This is largely a function of the particular conditions prevailing in the Seattle site, but it has the advantage of largely (but not entirely) avoiding thorny problems of control group comparability. Thus, in terms of research design, the SEDL study compares well with most previous studies in regard to conceptualization and operationalization of independent and dependent variables. In addition, its pseudo-longitudinal design and range of subjects (U.S. first-schooled and those first-schooled abroad) provides an exceptionally rich data base.

The major policy-related findings are as follows:

1. For both U.S. first-schooled and foreign first-schooled, length of residence (LOR) in the U.S. is strong predictor of English proficiency.

2. For the U.S. first-schooled students, the amount of bilingual education received was significantly related, in a positive direction, to the development of English academic skills, but not to English conversational skills. In other words, a moderate improvement in students' English academic achievement was noted as a result of bilingual education.

3. For the foreign-first schooled students, when LOR was held constant, the extent to which Cantonese literacy skills had been promoted prior to immigration to the U.S. was significantly related to the development of English literacy skills.

4. As a group, the Chinese students were performing exceptionally well in math (approximately 80th percentile) and adequately in English literacy skills. Those first-schooled in the U.S. appeared to be above grade norms in English reading skills and those first-schooled abroad slightly below grade norms.

Relationship of SEDL Findings to Previous Research

The SEDL findings appear quite consistent with trends that have emerged in previous studies. These trends can be summarized as follows:

1. Minority students instructed bilingually tend to perform at least as well, or better, than comparable students instructed in monolingual English programs (Baker and de Kanter, 1981; Cummins, 1983).

2. Among immigrant students, LOR or amount of exposure to L2, has generally been strongly related to academic and especially conversational L2 skills.

3. Among immigrant students, the extent to which L1 cognitive/academic skills have been promoted has generally been moderately related to the acquisition of L2 cognitive/academic skills.

4. Asian students have tended to show adequate academic achievement, whether in bilingual or English monolingual programs (e.g. Vernon, 1983).

Relationship of SEDL Findings to Theory

Three psychoeducational hypotheses are currently prominent in the research and policy debate regarding bilingual education. First, the linguistic mismatch hypothesis suggests that minority students' academic failure is caused by instruction through a language they do not understand; the hypothesis would predict that home-school language switching will lead to academic retardation. This hypothesis has usually been advanced as a justification for transitional bilingual education.

The second hypothesis is termed the maximum exposure hypothesis and argues that if minority students are deficient in English then they need as much exposure to English as they can get; consequently, diluting this exposure by means of bilingual education appears counter productive.

These two hypotheses represent intuitively-appealing conventional wisdoms. Despite their intuitive appeal and prominence in the policy debate each hypothesis is refuted by a vast amount of data. In other words, as theoretical principles or generalizations they have little validity. The linguistic mismatch hypothesis is refuted by the data from French immersion programs, showing that majority language children can be successfully instructed through a second language, and by the success of many groups of minority students under home-school language switch conditions (see Cummins, 1981a). The maximum exposure hypothesis is similarly refuted by the results of virtually every bilingual program ever evaluated (including French immersion programs) which show that students taught for all or part of the school day through a minority language suffer no detrimental consequences for the development of academic skills in the majority language.

The research data are fully consistent with the interdependence hypothesis which is more limited in its predictions than the linguistic mismatch hypothesis. The interdependence hypothesis argues that L1 and L2 academic skills are interdependent, or manifestations of a common underlying proficiency. This permits considerable transfer of academic skills from one language to another, given sufficient exposure and motivation to learn each. Thus, the hypothesis predicts that minority students taught through L1 will not lose out in the development of L2 academic skills (given exposure and motivation) despite considerably less time through L2.

The positive relationships between L1 promotion and English academic proficiency among both the U.S.A. first-schooled and foreign first-schooled groups in the SEDL study support the interdependence hypothesis. The other studies reviewed from both bilingual education and immigrant language learning situations also support the interdependence hypothesis.

Thus, the SEDL findings assume considerable importance for U.S. educational policy because they add to the evidence that the interdependence hypothesis is generalizable to an extremely large number of contexts. Considerable confidence can therefore be placed by policy-makers in predictions derived from this hypothesis regarding the outcomes of bilingual education programs.

The SEDL data are also consistent with the linguistic mismatch hypothesis in showing a positive relationship between bilingual instruction and achievement. However, unlike the interdependence hypothesis, the linguistic mismatch hypothesis is refuted by other data and consequently does not provide a basis for policy decisions.

This analysis is clearly not meant to imply that linguistic mismatch and L2 exposure play no role in accounting for student achievement. It is very likely more difficult to learn through L2 (other things being equal) and, as the SEDL study and many others demonstrate, exposure to L2 is important. What is being implied, rather, is that linguistic mismatch does not always result in academic retardation and is certainly not the only, or even the major cause of minority student underachievement. Similarly, although a second language clearly cannot be acquired in the absence of exposure to that language, much more than just

exposure is required to promote L2 academic proficiency. In this regard, the data strongly suggest that students' L1 academic proficiency (among other factors) plays an important role in making L2 academic proficiency comprehensible (see California State Department of Education, 1982). This "linguistic interdependence" factor helps account for the fact that minority students with less exposure to English often perform better in English academic skills.

In short, rational policy in regard to the education of minority students must abandon simplistic conventional wisdoms and acknowledge both what is known and what is not yet known. We know (i.e. we can predict with confidence), for example, that promotion of minority children's L1 in school will not in any way inhibit the development of English academic skills, despite the fact that less instructional time is spent through English. However, we do not, as yet, fully understand the causes of academic underachievement by some groups of minority students; bilingual programs appear to be effective, in many cases, in ameliorating this underachievement, but again we do not yet fully understand exactly how nor under what conditions bilingual programs achieve this positive effect. However, a considerable amount of research is being carried out internationally in this area and further empirical investigation and theoretical integration will doubtless continue to advance our understanding and ability to plan rationally for excellence in educating minority and majority children.

FOOTNOTES

1. The instructional pattern for all students was English-only with some ESL assistance when it was available and/or deemed necessary.

2. This partial correlation analysis was carried out by the writer on the basis of the Pearson correlations, reported by Okamura (1981).

REFERENCES

Altena, N. and Appel, R. Mother tongue teaching and the acquisition of Dutch by Turkish and Moroccan immigrant workers' children. Journal of Multilingual and Multicultural Development, 1982, 3, 315-322.

American Psychological Association. Review of Department of Education Report entitled "Effectiveness of Bilingual Education: A Review of the Literature". Letter to Congressional Hispanic Caucus, April 22, 1982.

Appel, R. The acquisition of Dutch by Turkish and Moroccan children in two different school models. Unpublished research report, State University of Utrecht, 1978.

Baker, K.A. and de Kanter, A.A. Effectiveness of Bilingual Education: A Review of the Literature. Final draft report, Office of Technical and Analytic Systems, Office of Planning and Budget, U.S. Department of Education, 1981.

Baral, D.P. Academic achievement of recent immigrants from Mexico. NABE Journal, 1979, 3, 1-13.

Brent-Palmer, C. A sociolinguistic assessment of the notion "immigrant semi-lingualism" from a social conflict perspective. Working Papers in Bilingualism, no. 17, 1979, 135-180.

California State Department of Education. Basic Principles for the Education of Language-Minority Students: An Overview. Sacraments: California State Department of Education, 1982.

Carlson, P.E. Field-Testing the theory of double semi-lingualism: An evaluation of Toukomaa's findings. M.A. Long Paper, Department of General Linguistics, University of Pittsburg, February 1981.

Carter, T.P. Mexican-Americans in School: A History of Educational Neglect. New York: College Entrance Examination Board, 1970.

Cohen, A. A Sociolinguistic Approach to Bilingual Education. Rowley, Mass.: Newbury House, 1975.

Connor, U. Predictors of second-language reading performance. Journal of Multilingual and Multicultural Development, 1983, 4, 271-288.

Cummins, J. Linguistic interdependence and the educational development of bilingual children. Review of Educational Research, 1979, 49, 222-251.

Cummins, J. The role of primary language development in promoting educational success for language minority students. In California State Department of Education, Schooling and Language Minority Students: A Theoretical Framework. Los Angeles: National Dissemination and Assessment Center, 1981. (a)

Cummins, J. Age on arrival and immigrant second language learning in Canada: A reassessment. Applied Linguistics, 1981, 2, 132-149. (b)

REFERENCES (cont'd)

Cummins, J. Heritage Language Education: A Literature Review. Toronto: Ontario Ministry of Education, 1983.

Cummins, J., Swain, M. Nakajima, K., Handscombe, J., Green, D. and Tran, C. linguistics interdependence among Japanese and Vietnamese immigrant students. In C. Rivera (Ed.) Communicative Competence Approaches to Language Proficiency Assessment: Research and Application. Clevedon, Avon: Multilingual Matters, in press.

Dulay, H. and Burt, M. Bilingual education: A close look at its effects. Focus, 1979, No. 1.

Ekstrand, L.H. Replacing the critical period and optimum age theories of second language acquisition with a theory of ontogenetic development beyond puberty. Bulletin No. 69, Malmo School of Education, Lund University, 1979.

Engle, P.L. The Use of Vernacular Languages in Education: Language Medium in Early School Years for Minority Language Groups. Arlington, Va.: Center for Applied Linguistics, 1975.

Fradd, S.H. Language acquisition of 1980 Cuban immigrant junior high school students. Doctoral dissertation, University of Florida, 1983.

Fukushima, M. The development of first and second language reading skills among Japanese children in Toronto. Qualifying Research Paper, O.I.S.E., 1983.

Gale, K., McClay, D., Christie, M., Harris, S. Academic achievement in the Milingimbi bilingual education program. TESOL Quarterly, 1981, 15, 297-314.

Gustavsen, R. Utvikling i norskkunskaper hos fremmedspraklige elever. Oslo Kommune, Skolesjefen, Skolesjefen, Undervisnigen av fremmedspraklige elever. Oslo, 1974.

Iwasaki, M. A study of sociological and developmental factors in first language maintenance and second language acquisition of kaigaishijo in New York. Unpublished M.A. Thesis, University of Hawaii, August 1981.

Jauho, S. and Loikkanen, M. Finska invandrarbarn i svensk skola: en sprakundersokning bland finska invandrarbarn i Goteborg. Pro gradu-av-handling i nordisk filologi. Tammerfors, 1974.

Krashen, S., Long, M.A. and Scarcella, R.C. Age, rate and eventual attainment in second language acquisition. TESOL Quarterly, 1979, 13, 573-582.

Nomoto, K. (Language maintenance among Japanese children in London). In C. Hayashi (Ed.) (A comparative study of Japanese as ethnic group: A survey conducted in Japan and Hawaii). Tokyo: Chuo Koran, 1973. (In Japanese)

Okamura, F. Mother tongue maintenance and development among Japanese children living in the United States. Doctoral dissertation, University of California at Los Angeles, 1981.

REFERENCES (cont'd)

Ogbu, J. Minority Education and Caste. New York: Academic Press, 1978.

Ouvinen-Birgerstam, P. and Wigforss, E. A critical study of Toukomaa's investigation of the bilingual development of Finnish immigrant children in Sweden. Pedagogical Bulletin No. 6, Department of Education, University of Lund, Sweden, 1978.

Robson, B. Alternatives in ESL and literacy: Ban Vinai. Final report to Asian Foundation, Centre for Applied Linguistics, 1981.

Rosenthal, A.S., Milne, A., Ginsburg, A. and Baker, K. A Comparison of the Effects of Language Background and Socioeconomic Status on Achievement Among Elementary School Students. Washington, D.C.: Applied Urbanetics, Inc., 1981.

Skutnabb-Kangas, T. Language in the Process of Cultural Assimilation and Structural Incorporation of Linguistic Minorities. Rosslyn, VA.: National Clearinghouse for Bilingual Education, 1979.

Skutnabb-Kangas, T. and Toukomaa, P. 1976 Teaching Migrant Children's Mother Tongue and Learning the Language of the Host Country in the Context of the Sociocultural Situation of the Migrant Family. Helsinki: The Finnish National Commission for UNESCO.

Snow, C.E. and Hoefnagel-Hohle, M. The critical period for language acquisition: Evidence from second language learning. Child Development, 1978, 49, 1114-1128.

Troike, R. Research evidence for the effectiveness of bilingual education. NABE Journal, 1978, 3, 13-24.

Vernon, P.E. The Abilities and Achievements of Orientals in North America. New York: Academic Press, 1982.

Willig, A. The effectiveness of bilingual education: Review of a report. NABE Journal, 1981-82, 6, 1-20.

Wong Fillmore, L. The language learner as an individual: Implications of research on individual differences for the ESL teacher. In M.A. Clarke and J. Handscombe (Eds.) On TESOL '82: Pacific Perspectives on Language Learning and Teaching. Washington, D.C.: TESOL, 1983.

Wright, E.N. and Ramsey, C.A. Students of non-Canadian origin: Age on arrival, academic achievement and ability. Research report No. 88, Toronto Board of Education, 1970.

Zappert, L.T. and Cruz, B.R. Bilingual Education: An Appraisal of Empirical Research. Berkeley: Bay Area Bilingual Education League, 1977.