Studia Anglica Posnaniensia XXV-XXVII, 1991-1993

UNTANGLING THE BILINGUAL EDUCATION CONTROVERSY

SHARON J. LYNCH

The George Washington University

1. Introduction

In 1974, a United States of America Supreme Court decision (Lau versus Nichols) mandated that when minority children's inability to speak or understand English excluded them from successful participation in a school program, the local school district must take affirmative steps to rectify this situation in order to open its instructional program to these students. With this decision, bilingual education was initiated in America on a wide-scale, amidst considerable controversy.

The bilingual education debate has been the most heated and politically charged educational debate in America in recent memory. Perhaps only bussing and desegregation have received more national attention (Gorney 1985). To Europeans who value multilingualism, and who appreciate or at least have learned to live with multi-culturalism, the bilingual education controversy in pluralistic America may be difficult to comprehend. Why is bilingualism such a charged topic in meltingpot America at this time?

This controversy over bilingual education strikes at the heart of what it means to become "Americanized". For centuries, the country that ironically has been the new home to speakers of the greatest multiplicity of foreign languages has been monolingual, with few if any serious threats to its monolingualism (Gorney 1985). Indeed, learning to speak English has symbolized a break with the Old World and entry to the real or imagined prosperity of the New. Since America has always been English-speaking, to allow, encourage, or even mandate bilingual education – basic instruction in our public schools in a different tongue – seems to some people to be "anti-American", and to others counterproductive if the goal is assimilation.

In addition, recently two best-selling books, *The Closing of the American Mind* by Allan Bloom (1987) and *Cultural Literacy: What Every American*

Needs to Know by E. D. Hirsch, Jr. (1987) lamenting the demise of a common cultural frame of reference have added fuel to the fire. The supposition is that bilingualism (or in some areas of the United States, the gradual encroachment of Spanish as the dominant language used in virtually every aspect of daily life, from the written word to the airwaves) will accelerate the process of the disintegration of a common set of ethical and cultural values, and weaken the American sense of fundamental purpose. For example Hirsch (1987) writes:

> Tolerance of diversity is at the root of our society, but encouragement of multilingualism is contrary to our traditions and extremely unrealistic... I know that well-meaning bilingualism could unwittingly erect serious barriers to universal literacy at a mature level. I am opposed neither to biliteracy nor to the learning of foreign languages... But surely the first step in that direction must be for all of us to become literate in our own national language and culture.

But Hirsch seems to implicitly make two assumptions. The first is that teaching students in their native language will delay and hinder their acquisition of English, and the second is that this in turn will interfere with their full and rapid assimilation into the American mainstream. Diane Ravitch (1985), a prominent and vocal American educator from Columbia University, gives voice to these fears when she asks:

> Then there is the larger question of whether bilingual education actually promotes equality of educational opportunity. Unless it enables non-English speaking children to learn English and enter the mainstream of American society, it may hinder equality of educational opportunity... Fluency in these languages (other than English) will be little help to those who want to apply to American colleges, universities, graduate schools, or employers, unless they are also fluent in English.

It perhaps seems logical then, that concentration on English alone in school as soon as possible would be the most efficient way of eliminating inequities in opportunity, issues of pride in minority culture, notwithstanding.

For ambitious minority students, the problem may be further complicated by their desire not to fall behind in other subjects, while they are trying to learn English. Should then instruction in these other subjects be provided in the first language or in English?

The critical question, as Ravitch points out time and time again, has to do with the effectiveness of bilingual education in teaching English to Limited English Proficient (LEP) students. This is *first* a research issue related to developmental psychology and psycholinguistics that needs to be resolved. Political stances, educational policy, and even philosophy need to be informed by accurate information. This paper will show how difficult, unfortunately, it has been to obtain such information.

But first, it is useful to separate the bilingual education research issues from the political rhetoric, and the sometimes legitimate complaints about hopeless bilingual educational programs that have arisen in attempt to comply with various state laws. For example, in the State of California, bilingual education

has been mandated whenever there are 10 or more students (non-native -speakers) in a school who are below a prescribed level of language proficiency (Gorney 1985). Instruction must be provided by a teacher (certified by State standards) in their native language whether its Lao, Cantonese, Hmong, Tangira (Ethiopian), Tagalog, or Vietnamese. But, for example, there have been few (or no) certified Hmong teachers. So a designated teacher must sign a pledge agreeing to learn the language, although classes in the language may not exist, and some teachers have neither the time nor the motivation to learn the language. Furthermore, a third of the students in these bilingual classrooms must be native speakers of English. These conditions are a formula for the creation of unhappy teachers, irate parents, and poorly functioning classrooms.

Adding to the potential for confusion, the program offerings for LEP students in America vary considerably, and this in turn helps to create misunderstandings of the research results. What are the current educational possibilities for the education of LEP students?

- "Submersion" which essentially places the LEP student in regular classes to sink or swim, with no special program. This up until recently had been the solution for rapid Americanization.
- "English as a Second Language (ESL)" which consists of special classes for English language instruction which may or may not include the students' first language. Other school subjects are taken in the regular classes subsequently or concurrently, and are in English.
- "Structured Immersion", the Canadian model, where the young students only hear the second language spoken. (This will be discussed in more detail later.) But note that in Canada, instruction is done entirely in the target language for the first three years; while in American immersion programs, explanations often are given in the first language in order to clarify the lessons.
- "Transitional Bilingual Education" which includes instruction in both the first language and in English, and encompasses all academic subjects.

It is understandable how bilingual education controversy may be exacerbated by confusion among the public and professionals alike, over what bilingual education is, or of what bilingual education of high quality might consist.

2. No research consensus

Assuming that at least the experts were able to adequately define bilingual education, still in the 15 years of state and federal funding of bilingual education programs, no consensus on efficacy of such programs has been reached (Baez et al. 1985; Ravitch 1985). Various authors of review articles have variously concluded that bilingual programs were helpful to LEP students if the programs were of a high quality; that the programs had no effect at all; or that they were harmful (Willing 1985).

Why all the disagreement about the efficacy of bilingual programs? The main reason is that the research has been of such questionable quality that it has been impossible to reach valid and defensible conclusions. For example, of three review studies, one found only 7 of 150 primary studies methodologically adequate, another found 12 of 180, and a third by Baker and deKantor, which will be discussed in detail later, judged 300 studies and found only 28 acceptable (Willig 1985). Some of the problems of the primary research studies included: bad research design, failure to adequately describe the study, incorrect and inadequate use of statistics, and the failure to equate treatment group and comparison groups on important factors such as socio-economic status or language proficiency. Of course, the reviewers also chose different studies for analysis and the goals of reviewers have often been quite different.

2.1. US Department of Education study

In 1979, two researchers, Baker and deKantor, were commissioned by US Department of Education to evaluate bilingual education programs and to resolve the controversy over these programs. Having focused upon two criteria, rate of English acquisition and mathematics achievement (Willig 1987), they concluded that the case for transitional bilingual education programs was very weak indeed.

Among the studies that they reviewed, they found that transitional bilingual education may have no effect at all, may have a positive result in one of three instances, and may be harmful in one of four instances. While recognizing the value of some of the programs, they concluded that there was insufficient evidence that transitional bilingual education was superior to all other methods, and thus it did not deserve a federal mandate (Baker 1987). They further concluded that other methodologies might also be effective, and that it should be up to the individual school district to make the decision regarding which method to use (Ravitch 1985).

Baker and deKantor used a traditional method of research review, which is sound, and statistically conservative since it only accepts a result as statistically significant if one is 95% or more sure that it ought not to have occurred by chance. This process excludes completely from consideration the grey area of results that are less powerful than this 95% level of certainty, a conservatism that is probably well-justified given the purpose of the Baker and deKantor study – to guide federal policy formation.

Because of the impact of the Baker and deKantor review on federal policy, the criteria that they used in the review are important to note. Baker and deKantor were *not* interested in literacy in the first language, or in the effect of transitional bilingual education on other school subjects, self-esteem measures, or drop-out rates (which have been a classic reason for bilingual programs for Hispanics who have a drop-out rate of 40% (Gorney 1985). For Baker and

deKantor, a program was successful only if it accelerated the acquisition of English better than *no* program (Baker 1987). Their review also considered mathematics achievement as another indicator of success. (This choice presents additional problems, to be discussed later.)

2.2. Willig's meta-analysis

Far from settling the matter, the Baker and deKantor study only stirred up the controversy, and drew considerable criticism on methodological grounds. Ann Willig, a researcher from the University of Illinois, decided to tackle the problem using a relatively new review technique, meta-analysis, which relies on the consolidation of data from many studies, and their subsequent statistical re-analysis.

Willig (1985) decided to base her meta-analysis on the 28 studies used by Baker and deKantor. The criteria they had used for inclusion could be easily adapted to meta-analysis and their study had had a great policy impact. In Willig's study, the main issues guiding the synthesis were more expansive than Baker and deKantor's criteria (see Baker (1987) for the rationale behind their more circumscribed approach). Willig's review would encompass achievement in English; achievement in mathematics and other subjects; school attendance; mediating effects of substantive variables, i.e. program design; and the mediating effects of research design.

Willig decided to limit her analysis to US studies only, and had to eliminate five of the original 28 because they were done in countries other than America, including the Canadian immersion studies, probably the best-documented research in this area available. The differences in goals and social conditions in the two countries makes Willig's decision to concentrate on US programs for her meta-analysis a wise one (Willig 1985).

In Willig's chosen method, meta-analysis, a computer is used to synthesize the data for all studies chosen for scrutiny. The use of the computer allows important variables to be held constant or considered simultaneously. In meta-analysis, group data of similar studies are combined to produce larger *n*'s. This has the effect of increasing sample size.

While objective, explicit, and systematic, meta-analysis still relies heavily on human judgement, since the researchers themselves must establish the criteria for the inclusion of primary studies for review, code the variables of interest (Willig coded 183 different variables from 23 different studies), and, after re-analysis by computer, draw conclusions. Meta-analysis has been shown to possess more power than traditional methods, and is likely to produce summary statements that are more precise, dispassionate, and detached than other techniques (Kavale & Forness 1990). If one compares Willig's review with Baker and deKantor's, one finds that Willig's yields a great deal more information than Baker's and deKantor's. While there are various advantages to the two review procedures used by Baker and deKantor, and Willig (Baker

1987; Willig 1987); on the whole, it appears that Willig's analysis covered a broader range of criteria for success, yielded more specific information, and allowed manipulation of the pooled data that resulted in a much better picture of the effects of bilingual education (Secada 1987).

2.3. Results of the meta-analysis

In overview, the results of the meta-analysis showed that the type of program and the type of tests used to evaluate the programs made the biggest differences when comparing transitional bilingual education to the submersion method. There were significant positive effects for bilingual programs over submersion. If tests were in English, positive results were shown for reading, language arts, mathematics, and total achievement. If the tests were in the students' first language, positive results were shown for listening comprehension, reading, writing, total language, mathematics, social studies, and attitudes towards school and self. Thus, Willig's meta-analysis was able to detect positive results by, in effect, using larger (combined) samples and holding various factors constant.

But Willig's results were not sufficiently robust, nor was her technique sufficiently free from criticism (Baker 1987) to lay the matter to rest. Rather, her results pointed in a positive direction for bilingual education, but their impact was mediated by her main conclusion: *The primary research included in the reviews (both her own and Baker's and deKantor's, of course) was of such poor quality that all of the results were somewhat suspect.*

3. Immersion versus transitional bilingual education

It seems ironic that in North America, there currently exist the most highly respected and best documented programs for teaching foreign language to small children, the Canadian immersion studies; as well as the most problem-ridden and controversial programs in the US bilingual education debate. Why not simply use the Canadian immersion model and apply it to LEP children in the United States (Gorney 1985)? The profound differences between the situations in Canada and the United States must be fully taken into account.

In the original Canadian program, majority language (English-speaking) children who lived in the bilingual (French and English) provinces were totally immersed in the French language. They received all instruction in French from specially trained teachers for their first three or four years of school. In the next several years, they received instruction in both French and English, with an approximately 50:50 mix. The result of this immersion method is that students are truly proficient in both languages, without any of their other school subjects being adversely effected. These results are extremely well-documented, robust, and consistent, and as a result, the program has expanded to other regions and the method has been attempted with other languages, for instance Native American languages (Genesee 1985).

Canadian French Immersion	*US Transitional Bilingual Ed*
– enrichment	– compensatory or remedial
– middle/upper class English speaking children learn French through complete immersion for all school subjects until 3rd or 4th year of school. Then they study both languages. They know no French when they begin the program.	– lower SES children are instructed in two different languages in school. Great variability in initial knowledge of both languages.
– goal is to learn second language with no decrease in proficiency in the first or in other subjects.	– goal is to learn English and not fall behind in other school subjects, sothat eventually the student will speak only English.
– children are volunteered	– children who are least proficient are designated
– both languages have high status	– first language generally has low-status and is discarded as soon as possible
– additive bilingualism	– subtractive bilingualism

The above comparison shows that the conditions in the Canadian French immersion studies – involving two stable, legalized, and esteemed languages – are profoundly different from those of America with one official language and a plethora of other tongues, new ones being introduced with each new wave of immigration. Lambert, one of the main authors and proponents of the Canadian immersion model emphatically states:

> The story about the positive effects of immersion is completely different for language minority young people. Immersion programs were not designed or meant for ethnolinguistic groups in North America that have some language other than English as the main language used in the home. To place such children in an initially all–English instructional program would be to misapply the immersion process in a harmful subtractive way.

3.1. Immersion projects in the United States

Nonetheless, there have been some successful and interesting exeriments in the immersion method in the United States, as well. Genesee (1985) reports on some of these immersion programs, such as the Spanish immersion program in Culver City, California, or Maryland's Montgomery County French immersion program. In such programs, middle-class English-speaking children learned a second language through immersion, following rather closely the Canadian model. The results of these programs were positive, and looked very much like Canadian studies.

An interesting variation of this approach which included children from lower socio-economic classes was the " magnet school" program in Cincinnati, Ohio. Here, the purpose was to increase racial integration by offering attractive programs that made mandated racial integration (bussing) palatable. The Cincinnati program offered a 50: 50 French immersion program to middle and working class black and white students. In almost all comparison (immersion versus non-emersion, white versus black, and by social class) groups of students did equally well on English assessments. Moreover, an analysis of the immersion students by group showed all groups to be successful in both English and in French; except for black middle class immersion students who were more successful in French than the black working class immersion students (Genesee 1985).

In San Diego, California, a two-way Spanish and English bilingual program was initiated in 1975. This was a Spanish immersion program following essentially the Canadian model, except that 40% of the children were English speakers and 60% were LEP Spanish speakers. Games and activities were included to encourage all students to interact with one another. All children, regardless of first language, were instructed mostly in Spanish until grade three, then the division was 50% English and 50% Spanish. Achievement results were reported for only the total groups, however, which make them impossible to meaningfully interpret. This is unfortunate, since the model is extremely interesting.

The Canadian immersion method seems most feasible for upper and middle class students from majority language homes, who are volunteered, and where there is much value placed on both languages. The verdict is still out on the immmersion method for LEP children. Once again, poor research methodology tends to obfuscate the salient issues.

3.2. Collier's studies on age and rate of language acquisition

A final important component in understanding the complexities of teaching LEP children in the United States comes, curiously enough, from a study by Collier (1987) done on English as a Second Language (ESL) students, not transitional bilingual education students, and her follow-up review article (1989). In the Collier study, 1, 548 children who recently arrived in America (predominantly Spanish, Korean, and Vietnamese speakers), were followed longitudinally in order to determine the length of time it took them to learn *English*, and to score at the achievement means for their grade level in all *other academic subjects*. In other words, how many years did it take them to perform as well as *average* native-born children in school?

Collier only used for her study subjects who were doing well in school prior to arrival in America, primarily upper and middle class children of educated parents in their country of birth, although their incomes tended to be much lower in the United States, at least initially. (It would be interesting to see this study replicated for children who come from uneducated, lower income families.) The more advantaged students studied by Collier were not exposed to a transitional bilingual education program. Rather, all students had been placed in ESL classes. This meant that most of their school work being done in typical academic classes, where only English was spoken.

Collier found that the students who reached grade level (the mean achievement level by academic subject according to American norms) fastest were those who arrived in US at ages 8-11, followed by 5-7 year-old arrivals. Poorest were those who arrived at ages 12-15 years, who had the greatest demands placed on them in that the level of English needed for them to be able to succeed in junior high and high school level academic classes was far greater than the younger children's. Consequently, it took longer for them to reach the grade level norms in most subjects, mathematics being the exception.

Although this was not a transitional bilingual education program, Collier's conclusions for acquisition of *English* nonetheless seem to support the principles of transitional bilingual education. She concluded (1989) that:

> Before puberty, it does not matter for overall long-term academic achievement when one is initially exposed to (or first receives instruction in) the second language, as long as L1 (first language*) cognitive development is continued through age 12 (the age by which language acquisition is largely completed). When children's L1 development is discontinued before it is completed, they may experience negative cognitive effects in L2 (second language*) development; conversely, children who have reached full cognitive development in two languages enjoy cognitive advantages over monolinguals.

For other *academic areas*, students generally needed 4-7 years to reach national achievement norms in most subjects if schooled in two languages (except for mathematics and language arts). Collier recommends that children older than age 12 who speak no English when they arrive in America probably need academic instruction in their first language in order to maintain their academic progress in other subjects. Collier also says that children younger than age eight might acquire English for academic purposes faster if they had two more years of instruction in their first language, since they seem to need more extensive cognitive development in their first language in order for effective transfer to other subjects and the second language.

Mathematics is an exception to this general rule – LEP students may achieve at the average for their grade level in as little as two years after arrival in the US. Collier's findings for mathematics achievement in light of previous discussions in this paper deserve note. The students were able to reach grade level in mathematics in two years. Mathematics is a subject that seems least dependent upon mastery of English, since it depends to a great extent on its own symbol system which is largely international. This fact particularly calls into question Baker and deKantor's choice to use mathematics as one of the two parameters for measuring the efficacy of transitional bilingual education

* parentheses mine

programs, since mathematics is a subject that is most independent of the effects of second language learning. Almost any other academic subject would have been a better choice, if the goal was to measure LEP students' overall successful functioning in school as a results of acquisition of the second language.

Thus, Collier's conclusions and recommendations on the basis of her work and her review of other studies including those from countries other than the United States (1989), are consistent with the attributes of bilingual education programs. They suggest that instruction in both languages seems to be most effective in teaching children English, as well as aiding their continuous progress in other school subjects. Collier does not directly address the best way of doing this. But the important point is that the common wisdom – that all immigrant children should learn English as quickly as possible, disregarding their first language – may be counter–productive for both English language acquisition and their academic mastery of other subjects.

It should also be pointed out that many research methodological issues that Willig (1985) raises were not addressed in Collier's work, and a direct test of transitional bilingual education program success versus ESL program success was never done by Collier. Thus, despite the quality of Collier's synthesis, straightforward research on the effectiveness of transitional bilingual education using true experimental design remains scarce.

4. Conclusion

Research on bilingual education, taken overall, has been of too poor quality to convincingly settle the debate; Collier's, Willig's, and a few other researchers' work notwithstanding. The lack of quality in the primary research has resulted in the amplification and extension of the bilingual education controversy, rather than providing important information, critical to making policy decisions. Glass and Willig (Willig 1985) make an excellent argument for the need for leadership in providing research of the highest quality at the outset of such programs. Funding large-scale evaluations of many improperly designed programs seems a waste of money and time, and only fans the flames of controversy. The first important conclusion is really a recommendation: The United States Department of Education should take a more active role in funding and maintaining quality control of primary research studies on controversial programs such as bilingual education.

The second conclusion is that the bulk of the current evidence points in the direction that transitional bilingual education of high quality may be the most efficient way for LEP students to learn English and keep up with other subjects in school. An important bonus is the development of proficiency in the first language, as well. Immersion, Canadian-style, may be best for teaching young middle-class native speakers a second language.

The discussion of the bilingual education controversy must be informed by systematic, sound research. At first blush, it seems logical that bilingual education (education in the second language as well as the continuation of the development of the first) might somehow interfere with the rapid mastery of English, and delay or inhibit progress in other academic subjects. This is implicitly the same worrisome conclusion that Hirsch (1987) and other Americans intuitively reach when considering transitional bilingual education. But the best evidence points in the counter-intuitive direction: Instruction in both languages may be the most effective method for teaching English and other subjects to American LEP children. Children's language and cognitive development seems to require a different solution for language learning than for adult language learners where it is clear the immersion in the new language results in the most rapid mastery.

If this conclusion was supported by the strongest research evidence (which admittedly is not currently the case), would there be any real argument against bilingual education, politically and philosophically? That is, if we knew that a given educational program could be provided which resulted in maximum mastery of English and all other academic subjects for LEP students, would any but the most jingoistic and threatened persons begrude the side-effect of proficiency in another language (the students' first language)? In the author's opinion, this would not trouble Hirsch, Ravitch, or many others, as long as they felt certain that the primary goal of Americanization– moving the LEP children into the mainstream of American society with mature understanding of ideas fundamental to the American heritage –was being attained. The result would be a better educated populace, well-suited to functioning in an increasingly international world climate.

REFERENCES

Baez, T., Fernandez, R. R., Navarro, R. A., & Rice. R. L. 1989. "Bilingual education." In Noll (ed.) 1989. 308-20. (Reprinted from *Issues in education* 3. (1985).

Baker, K. 1987. "Comment on Willig's 'A meta-analysis of selected studies in the effectiveness of bilingual education'". *Review of educational research*, 57/3. 351-62.

Bloom, A. 1987. *The closing of the American mind*. New York: Simon and Schuster.

Collier, V.P. 1987. "Age and rate of acquisition of second language for academic purposes". *TESOL quarterly*, *21/4*. 617-41.

Collier, V. P. 1989. "How long? A synthesis of research on academic achievement in a second language". *TESOL quarterly*, 23/3. 509-31.

Genesee, F. 1985. "Second language learning through immersion: A review of U.S. Programs.". *Review of educational research*, *55/4*. 541-61.

Gorney, C. 1985. "The bilingual education battle." *The Washington Post national weekly edition*, July 29, 1985. 6-10.

Hirsch, Jr. E.D. 1987. *Cultural literacy: What every American needs to know.* New York: Vintage Books.

Kavale, K. A. & Forness, S. R. 1990. "Substance over style: A rejoinder to Dunn's animadversions." *Exceptional children, 56/4.* 357 - 61.

Noll, J. P. (ed.) 1989. *Taking sides.* Guliford, Ca.: The Duskin Publishing Group.

Ravitch, D. 1989. "Politicization and bilingual education." In Noll (ed.) 1989. 308 - 20. (Reprinted from *Proceedings of the American Philosophical Society,* 1985, *129/2*).

Secada, W. C. 1987. "This is 1987, not 1980 a comment on a comment." *Review of educational research,* 57/3. 377 - 84.

Willig, A. C. 1985. "Meta-analysis of selected studies on the effectiveness of bilingual education". *Review of educational research.,* 55/3. 269 - 317.

Willig, A.C. 1987. "Examining bilingual education research through meta-analysis and narrative review: A response to Baker." *Review of educational research,* 55/3. 363 - 76.