

Written Testimony of Kenji Hakuta<sup>1</sup>

to

The New York City Mayor's Task Force on Bilingual Education

October 30, 2000

My name is Kenji Hakuta. I am the Vida Jacks Professor of Education at Stanford University. I have conducted research and published articles on the development of language minority students for over twenty-five years. Especially after the divisive, politically-charged debates we have recently experienced over bilingual education, I believe it is critical to bring research into discussions concerning the education of English Language Learners (ELLs). I am pleased to offer my written testimony to the New York City Mayor's Task Force on Bilingual Education from that perspective.

*Length of time to learn English*

I am often asked by the media and by policymakers: "How long should we expect it to take for English learners to learn English?" Unfortunately, there is no simple "formula", but we do know some things. For example, it depends on whether ELLs are proficient in their first language upon their arrival to the U.S. and at what age they enter U.S. schools. In a recent paper I co-authored with Yuko Goto Butler and Daria Witt, we estimated that it would take ELL students between 4 to 7 years to learn English, and that this time period varies considerably by the socioeconomic status of the students as well as the aforementioned variables. The study further showed that this estimate would hold both in districts that have bilingual education as well as those that use English-only and have intensive English as a Second Language instruction.

Basic research has corroborated my findings, showing that ELL students are learning English at a rapid and natural rate of development ("at the speed limit"), regardless of relative amounts of exposure to English vs. native language in school. It appears then that a simplistic, "time on task" theory fails to predict English language development, just as it fails to predict patterns in the acquisition of a first language. It also appears unlikely that any further intensifying of instruction in English would cause the rate of acquisition to become much faster.

Looking specifically at New York, data I have reviewed from the New York State Education Department (NYSED) and Board of Education of the City of New York on the

---

<sup>1</sup> I would like to acknowledge the help of my research assistant, Michele Bousquet Gutierrez, in preparing this written testimony.

progress of English learners in New York City clearly support my findings of the length of time it takes for students to attain English language proficiency.

### *Controversy over bilingual education*

The reported research findings on the effectiveness of bilingual education present mixed results. However, the best research (i.e., research that investigates programs that are representative of their “bilingual” label and that are properly implemented) suggests that bilingual education *is* a successful model and *does* produce measurably better outcomes in academic achievement. This finding has been supported by two independent committees of the National Academy of Sciences (I chaired one of these committees, and served on the other). In fact, it is often the studies with the weaker research design that have reported inconclusive findings on the success of bilingual education (the old research saying, that “garbage in, garbage out” seems to hold). Additionally, several meta-analyses (Green, 1998 and Willig, 1985) that methodically pool the findings across multiple studies have consistently supported bilingual education over English-only alternatives.

That said, the comparison of bilingual programs with English-only programs, as I have argued elsewhere, has come to serve as a distraction from attention to the much larger problem of educating children from poor socioeconomic backgrounds in under-resourced schools. I need only point to the recent studies reported commissioned by the ELL Subcommittee of the New York City Board of Education which suggests that consistent programming and school quality (probably teacher quality, though the analyses do not get into that) are key to successful outcomes. In that study, mixed programs yielded the worst outcomes. There were only minor differences in outcome between the bilingual and ESL students, most of this probably due to their differences in socioeconomic status.

Research also suggests that success in non-bilingual education settings most likely occurs when children come from homes in which there is a high degree of education, where second language only programs are taught by teachers who have the skills to communicate with the child in his or her first language, and where there is a school-wide climate that supports high degrees of learning. However, absent these conditions—as is usually the case, the research strongly suggests that children are best taught at least to some degree in their first language until they have acquired proficiency in English. Because of this, native language instruction is one method of teaching that *educators must be free to use* in constructing effective programs for ELLs. Depriving educators of this method summarily removes one of the many useful tools that can be used to improve learning outcomes for these students.

### *Identifying components of successful programs for ELL students*

Having defended the need for native language support, it is important to recognize that more will be gained by focusing not on program type but on program components. This is important for several reasons. First, not all program labels are representative of the

instructional methods they employ. Second, even those programs that are representative of their labels vary widely in their quality, and third, successful components are likely to be found in a variety of programs that are being run effectively. In effect, there is likely no ONE best model for educating ELL students. What is critical is finding a set of program components that works for the children in the community of interest, given the goals, demographics, and resources of that community. This set of components will (and should) vary depending on factors that differ not only across but within immigrant groups, such as students' first language, SES, previous academic experience, community and parental socio-linguistic climate, learner styles, and goals for proficiency (additive v. subtractive). Teacher availability/ qualifications also play an important role. The best approach to take is to look to basic research, which will inform us as to how children learn best and under what conditions.

What has emerged from this thinking is a set of generally agreed upon “best practices” that can and should be found across program types to encourage the success of language minority students. A typical list of key components would include the following: some use of native language and culture in the instruction of language minority students, a balanced curriculum that incorporates both basic and higher-order skills, explicit skills instruction, opportunities for student-directed activities, use of instructional strategies that enhance understanding, opportunities for practice, systematic student assessment, staff development, and home and parent involvement. Since the success of language minority students does not hinge solely on the classroom environment but also on that of the school, criteria for “good” schools for language minority students have been developed as well. The following are recommended school attributes: a supportive school-wide climate, school leadership, a customized learning environment, articulation and coordination between and among schools in the district, school-wide coherence, rigorous standards for teaching and learning, assessment and accountability, continuous evaluation, and research of program effectiveness. Many of these strategies for overall improvement of schools were derived and adapted from successful improvements for high-poverty schools, where approximately 75% of ELL students attend.

### *Assessing ELL students performance*

Up to now, program evaluations have pitted bilingual education against English-only programs and have evaluated their success based on students' English language development. The identification of effective components for ELL student programs signals a shift in thinking away from program labels. What is needed now is a shift in focus for the evaluation of ELL programs—a departure from gauging overall program effectiveness by looking simply at students' English language development and a commitment to determining success based on academic standards.

Currently, however, in typical programs for ELL students—regardless of program type—high levels of academic learning are not promoted. In order for standards-based evaluations of ELL programs to work, then, it stands to reason that ELL students need to be included in standards-based reform movements in the various content areas and in

academic measures of competitiveness. Clear academic standards for ELL students must be in place, confirming the need to set the same expectations for this population as for mainstream students. Emphasis should be placed on the integration of theory, standards, instruction, and assessment.

It is important to remember when setting standards for ELL students that what ultimately matters most is their long-term performance and success. It is unreasonable to expect ELLs to perform comparably to their native English-speaking peers in their initial years of schooling and holding them to this expectation too early in their educational careers can be detrimental to their academic progress, not to mention their self-esteem. The problem enters when students are not pushed to go beyond this stage over time, are presumed to be at an elementary level, or are misdiagnosed as having educational disabilities by teachers unfamiliar with the needs of ELLs.

I very much applaud the recent efforts by the Board in conducting systematic studies of ELL students as they progress through the system. Such longitudinal databases are highly unusual, and much more informative than the cross-sectional snapshots that we have in places such as California. I would recommend that this work be continued, and augmented with additional information, especially in the academic progress of these students in the content areas.

When ELLs are involved in high-stakes testing, consideration must be given to the appropriateness of accommodations (e.g., testing through their native language or through alternative administrations, awarded additional time, etc.) that might be made for this particular group of students. When reviewing the New York Regents Examination, I noted that the level of English vocabulary reflected in both the reading and listening passages is quite advanced and well beyond the beginning English vocabulary that would typically be employed in materials in ESL classes for students learning English as a new language. While the advanced English vocabulary of the Regents examination might reflect English language arts as taught to native English speakers from kindergarten through high school, that vocabulary would be much more difficult, indeed, literally and figuratively foreign to ELL students. As I have indicated, it is my opinion that up to seven years would be required before it could be said that a majority of ELL students would have enough years of exposure to such vocabulary to be able to fully understand the language used on the Regents test not be disadvantaged in their chance of passing it. It is easy to see what a disadvantage ELL students would be operating from without testing accommodations.

#### *Recommendations for professional development of ELL instructors*

The success of students is of course highly dependent upon their teachers' professional preparation and certification. This is particularly important when considering ELL students. Currently, the majority of ELLs are enrolled in mainstream classes, taught by teachers who quite possibly are not prepared to offer the specialized instruction these students necessitate. As such, additional support is needed for professional development aimed at training instructors to better manage the changing student body population.

Both current and future teachers should be provided with the resources necessary to assist them in understanding the methodological and affective accommodations ELL students may need. The recently completed efforts of the National Board for Professional Teaching Standards to develop standards for bilingual and ESL teachers should be applauded as the "deluxe" model, but the magnitude of the problem is staggering when we look at the other elements of professional preparation such as schools of education, state certificate requirements, professional development models, and Title VII incentives. In addition, current knowledge about the effectiveness of strategies for teacher education and the assessment of teacher knowledge and skills is very limited. Lawmakers should demand a systematic inquiry into ways to understand, support, and coordinate all of these efforts.

### *A Final Note on Proposition 227 in California*

As I write this from California, where Proposition 227 originated, it is imperative that I conclude with a note on the much-touted evidence of its effectiveness. Much of the excitement about California comes from a New York Times article that reported on the supposed miracle of a school district in Oceanside, California. I reproduce here some simple points about the California story from my website, which contains much more detail ([http://www.stanford.edu/~hakuta/SAT9/SAT9\\_2000/index.htm](http://www.stanford.edu/~hakuta/SAT9/SAT9_2000/index.htm)):

1. Any given school district's pattern of performance by LEP students should be considered in light of statewide patterns of performance by LEP *and* by native English speakers; our analysis shows that there have been statewide increases in SAT-9 scores for both LEP and native English speakers, following patterns that are virtually identical -- large increases in the early grades, and then tapering off in the fourth grade and beyond. This is not a Proposition 227 effect, but something much more specific to SAT-9.
2. The increases are due to a number of possible causes. Advocates of reforms such as Proposition 227, class size reduction, and increased school accountability would certainly like to give credit to their own individual causes, but there are other explanations that must also be considered. For example, schools and districts have taken the SAT-9 much more seriously this past year, and have taught to the test. Younger children's scores are probably more likely to benefit from increased attention by teachers and school officials to the importance of the test. Also, districts seem to vary considerably in who they included as LEP or as non-LEP, and in percentages of the LEP students that they tested. Of course, the results of a school or district's LEP students would depend a great deal on who they count as LEP and which LEP students were tested. Each claim about "success" for LEP students would need to be scrutinized. It is certainly premature to claim any sort of victory for Proposition 227.
3. SAT-9 is a poor excuse of a measure of English development and academic achievement for LEP students. The test was developed to give normative data in reading and math for native English speakers. The test measures things that are qualitatively different from what would be expected of students learning English. Consider an analogy. Imagine if you had just finished a first set of golf lessons in a driving range, and then you

were taken out to a golf course, asked to play a full 18 holes, and kept score. Unless you were a prodigy, your score would be virtually meaningless, measuring luck much more than it would your ability. The golf score is very meaningful for those who have played for a while (Tiger Woods), but not for beginners (being one, I can testify that I never keep score -- I keep score in a different way, which is the percentage of solid contacts I make per swing). Given that SAT-9 is a weak measure of English for LEP students, we can only expect it to tell us very gross information. It is certainly not refined enough to tell us about differences between program labels, such as bilingual vs. English immersion. (Would I really be able to tell the difference between the effectiveness of different golf instructional approaches based on golf scores for beginners?).

4. The data from 1998 to 2000 show that all districts show rises, pretty much following statewide patterns. There are increases in school districts that have retained bilingual education, in school districts that had English immersion even before Proposition 227 (and therefore were not impacted by the policy), and in Oceanside, which has been acknowledged by the press for having switched faithfully from bilingual to English-only. Because SAT-9 is a bad measure for LEP students (golf scores), the scores for schools and districts are characterized by a lot of random noise, but they did rise in a rough way. That is, all the scores are rising, but the margins of errors are so large that it is not possible to distinguish between different types of language programs.

5. Why did Oceanside LEP students show such big gains from 1998 to 2000? Partly, one has to wonder how it managed to be so low in 1998 -- the average LEP 2nd grader at the 12th percentile (compared to LEP at the 19th percentile statewide), and the average 3rd grader at the 9th percentile (compared to 14th percentile statewide). So, they started out among the lowest in a group of students who score low to begin with. One of the laws of statistics is that the lower the beginning score, the more it will be expected to rise upon retesting. Also, an important perspective is that one can pretty easily find schools report having well-run bilingual education programs, that have equally dramatic gains as did Oceanside.

Thank you for your attention.