

## 1. RATIONALE

According to figures provided by the California Department of Education, 1,323,767 limited-English proficient (LEP) students--nearly one in four (23.58%) of all enrolled students--attended California public schools in 1996. Spanish speakers made up the vast majority of these children, constituting about 79.4% of the total LEP population in California. The U.S. Census Bureau reported an increase of nearly 100% over the past decade in enrollment of LEP students nationwide (an annual growth of about 9.2%), with a total of 2.8 million such students reported in the *1993-1994 Schools and Staffing Survey*.<sup>3</sup> Given the composition of the student population in California and in the U.S., continued research on the nature of bilingualism and the conditions for academic success for bilingual children is a matter of great importance.

As one factor, teachers' attitudes about children's abilities are known to strongly impact upon their success or failure in school. In a comprehensive summary of research on "teachers' thought processes," for instance, Clark and Peterson (1986) point to ongoing psychological research which suggests that "the most important beliefs that teachers have about students are those that deal with teachers' perceptions of the causes of students' behavior or, in other words, teachers' attributions for the causes of students' performance."

---

<sup>3</sup>Figures reported here are from the U.S. Census 1980 and 1990, cited in Macías (1993), California Department of Education (1996), and the National Center for Education Statistics (1997).

Specifically with respect to code switching, Ramirez and Milk (1986) found that teachers differentiate “standard American English” from three marked varieties, with “Hispanicized English” rated more favorably than ungrammatical English constructions and code switching. Of the four varieties of language differentiated in Ramirez and Milk’s (1986) study, code switching was consistently ranked “least acceptable” by teachers. The need for a better understanding of code switching phenomena among classroom teachers is emphasized by Valdes-Fallis (1978):

An understanding of code switching is especially important for those classroom teachers whose students include Spanish/English bilinguals. While a great deal has already been said concerning the importance of acceptance of the child’s home language by the teacher, such discussions have generally involved the different varieties or dialects of both English and Spanish that children bring with them to the classroom. Very little has been said about the characteristics of bilingual speakers who habitually alternate between two languages in their communities. Moreover, bilingualism itself is very poorly understood by most educators, and, for that reason, much of the literature available to the classroom teacher misrepresents language processes that are normal for bilingual speakers of every linguistic community. A typical instance is the labeling of the alternating use of English and Spanish in this country as “Spanglish,” “Mex-Tex,” or “Pocho,” and the common belief, held by many teachers, that children who code-switch really speak neither English nor Spanish [2].

If teachers believe that children who code switch have low language ability in both languages, as Valdes-Fallis (1978) suggests, then this belief may strongly influence their expectations for these children and determine curricular content and teaching practices students receive. (Also see Attinasi (1982) on teachers’ attitudes about code switching.) Thus, research which aims to change teachers’ beliefs about stigmatized language varieties in general, and about code switching in particular, is an important contribution to the fields of education and educational research.

I argue in this chapter that a climate of school failure for language-minority children arises from two intellectual traditions, prescriptivism and “semilingualism.” Rather than challenge these unfounded notions, which have the potential to harm children through tracking mechanisms, some researchers in bilingual education have at times played a role in *promoting* these false and potentially damaging ideas. Below, I present a theoretical framework in which the sociopolitical role of teachers’ beliefs about students may be assessed, and then I sketch some of the ways in which varieties of language and their associations with particular social classes have served as a basis for constructing social hierarchies around myths of “intelligence” and “cognitive skills.” Following this, I discuss the special manifestation of this dogma in work on bilingualism, focusing on code switching, and then I address the implications of the historic Ann Arbor decision for language education for bilinguals. Finally, I comment on the potential misplacement of bilingual children in special education programs as a result of a poor understanding of code switching behavior.

In closing, I propose that fluent bilinguals who code switch have the same rich linguistic competence as monolinguals for the languages they use, a notion I set out to show in the remainder of this volume. If correct, then the proposition that bilingual code switching reflects a linguistic “deficit” of some kind may be dismissed, together with tacit tracking practices associated with it.

### *1.1 Schooling, Propaganda, and Social Class*

The institutional role of teachers’ beliefs may be analyzed in sociopolitical terms, following a recent approach pursued by a number of educational researchers who study

curriculum from the perspective that schools, as the result of many social and historical forces, serve primarily to reproduce an existing social order in which people are divided, often ruthlessly, along lines of class, race and gender (Parsons, 1959; Bourdieu and Passeron, 1977; Willis, 1981; Giroux, 1983; Oakes, 1985; McNeil, 1988; Macedo, 1994; McLaren, 1994). According to Gramsci (1971) and Takaki (1979), modes of discrimination based on race and gender derive from a deeper socio-economic need in capitalist societies, namely, the need to create social classes. From this perspective, schools “process children into roles for economic production” (McNeil, 1988) and sustain class structure by using, among other devices, ideological constructs regarding the status of languages and language varieties which mark disenfranchised groups as inadequate or inferior to the dominant social class. Thus, language attitudes may be a factor in the construction of a social arrangement of the sort the anarchist Mikhail Bakunin (1970 [1883]) described long ago as promoting “the advantage of a dominant minority of exploiters against the interests of the immense majority in subjection to them.”

This view of the role of schools in democratic societies is analogous to Chomsky’s view of the media. Within Chomsky’s (1989) Propaganda Model, the media systematically distorts the news in favor of ruling elites in the U.S. This claim is empirical in nature; it can be tested by looking at the facts. Herman and Chomsky (1988) suggest three techniques which can be used to test it. First, in their work, cases which reputedly support the claim that the media is independent of corporate and elite interests are scrutinized in terms of the Propaganda Model. In a second, far more persuasive technique, paired examples of historical events are studied for discrepancies in

news coverage. Third, an exploration of the range of opinion permitted on a given topic is shown to define the boundaries of acceptable discourse in mainstream media.

It may be helpful to look at a concrete example, offered in Herman and Chomsky (1988). Consider the media's treatment of the Watergate Affair, presented as an embarrassing "domestic scandal." Seen as the zenith of investigative journalism throughout the world, the mainstream press portrayed the Nixon administration's crime as one of using a group of petty criminals to break into a room at the Watergate hotel for reasons that remain obscure. At the height of passion over Watergate, it was discovered that the FBI had been disrupting the activities of the Socialist Workers Party, a legal U.S. political party, for more than a decade. In contrast to the Watergate Affair, this event received virtually no media coverage and is today scarcely known. Similarly, after Nixon's "enemies list" was exposed in the press, it was discovered that the FBI directed the assassination of Fred Hampton, an influential leader of the Black Panthers. This incident alone completely overshadows in significance all the reported crimes of Nixon, but it again received very little media coverage. According to Herman and Chomsky, exposing Nixon's offenses falls within the range of acceptable news coverage because it constitutes no threat to the general social order. In fact, because it lends credence to the notion that the media is made up of independent news organizations, coverage of this sort only serves to further reinforce the general public view that we live in a free and just society. Many, many examples of this sort may be given.

In school curriculum, too, a particular view of the role of the U.S. in world affairs is constructed, one which favors the position of ruling elites. In history classes, for

instance, Columbus is portrayed as an adventurous explorer in search of new lands, while a look at his own notebooks reveals him to have been a murderous mercenary in search of gold and capital to repay the investment of the Queen of Spain (Zinn, 1980). Scores of examples of this sort may also be given, historical portraits that have been wildly reconstructed to conceal relationships between capitalist ventures and social injustices, or which represent U.S. interventionism as heroic self-sacrifice in the interest of spreading democracy.

This perspective on curriculum entails that the teaching function of schools is, in general, highly constrained by their control function, as McNeil (1988) and Macedo (1994) have also argued. Indeed, Chomsky (1988) reports that schools are referred to by the Trilateral Commission<sup>4</sup> as “institutions responsible for the indoctrination of the young,” places which are ideal, the Commission argues, “for imposing obedience, for blocking the possibility of independent thought,” and which should “play an institutional role in a system of control and coercion” (671).<sup>5</sup>

Prescriptivist values and negative views of particular language varieties may be viewed as serving a control function in schools, by raising expectations for children viewed in a positive light (who have speech characteristics of the privileged classes) and lowering them for those viewed negatively (who have speech characteristics of the lower

---

<sup>4</sup>The Trilateral Commission is a think-tank of liberal elites formed in 1973 by Columbia University Professor Zbigniew Brzezinski with the support of David Rockefeller. Jimmy Carter, among many others, is a member. For detailed discussion, see Sklar (1980).

<sup>5</sup>See Rai (1994) for an excellent introduction to Chomsky’s political thought. For Chomsky’s particular thoughts on schooling, see Chomsky (1975), Corson (1994), and the interview with Chomsky in Olson and Faigley (1991).

classes), thus placing children of elites in a position to succeed in school. Oakes (1985) analyzes similar tacit mechanisms as leading to informal ability groupings, documenting the well-studied effects which ability labels have on teachers' expectations.

Language mixture, or code switching, is a speech style of bilinguals which has been especially stigmatized in schools, as Ramirez and Milk (1986) and Valdes-Fallis (1978) have pointed out. However, before looking more closely at some of the forces that lead to this stigma for bilinguals, a general consideration of prescriptivism may be useful.

### *1.2 Prescriptivism and the Status of Languages*

Prescriptivism, in its most general sense, is the view that one variety of language has an inherently higher value than others, and that it ought to be imposed on the whole of the speech community to maintain "standards of communication" (Crystal, 1987). Language academies employed with the task of "purifying" the regional linguistic descendants of Latin were set up as early as 1582 in Italy, 1635 in France, and 1713 in Spain. Proposals for a language academy in England were also popular in the seventeenth century (Jonathan Swift's, among them), but the suggestion lost support as it became evident that the European academies could not halt the tide of language change. (See Crystal (1987) for further discussion.)

The prohibitions on English usage which are most familiar from U.S. high school curricula, picked out of influential prescriptive grammars, typically turn on Latinate analyses advanced in the late nineteenth and early twentieth centuries and used to validate varieties of speech associated with the educated classes in England and the U.S.

(Baugh and Cable, 1978). In the thick of this tradition, which took literary and classical languages to be of greater expressive capacity and complexity, the structuralists in the U.S. sought to carry out a program of research which assumed that all languages of the world are cast from the same die. As Newmeyer (1986: 42) puts it,

As long as American structuralists confined their campaign to the languages of remote tribes, they did little to upset their colleagues in departments of modern and classical languages--in which almost all linguists were situated in the interwar years. But such was certainly not the case when they began crusading for the linguistic equality of *all* dialects of English and other literary languages, no matter how “substandard” they were regarded. This egalitarian view came in direct conflict with the long-seated tradition in the humanities that values a language variety in direct proportion to its literary output.

Indeed, as recently as 1964, the Linguistic Society of America reported to the National Commission on the Humanities that “a fair proportion of highly educated laymen see in linguistics the great enemy of all that they hold dear” (American Council of Learned Societies, 1964).

While much of seventeenth-century Europe was preoccupied with “elite language,” the *Port Royale Grammar* of 1660 advanced a very different view of language and of the human condition.<sup>6</sup> Written in French, the *Port Royale Grammar* formed part of the movement to displace Latin as an outdated mode of academic discourse. However, what marked the *Port Royale Grammar* as deeply distinct from contemporaneous approaches was its devotion to philosophical and universal properties of human language

---

<sup>6</sup>Actually, many of the early notions regarding “good usage” were strongly associated with classical liberal thought, but these ideas were lost as prescriptive grammar forged an alliance with elitism. “Inevitably, schoolteachers found themselves spending more time teaching rules of usage that had no basis in the rational program of traditional grammar, but only in invidious distinctions of class and race” (Nunberg (1983), cited in Newmeyer (1986: 45)).

in descriptive terms, even developing notions of “surface structure” and “deep structure” which would not reappear until Chomsky (1964) (Robins, 1967; Chomsky, 1968; Newmeyer, 1988).<sup>7</sup> Like Chomsky, the *Port Royale* grammarians worked on the Cartesian assumption that normal human intelligence is capable of acquiring knowledge through its own internal resources, making use of the data of experience but moving on to construct a cognitive system in terms of concepts and principles that are developed on independent grounds. The fear that languages might “decay” in the process of change, or the notion that groups from different cultural backgrounds speak “diminished” or “simplified” languages when compared to Europeans, is incompatible with Cartesian assumptions since languages are held to “grow” by virtue of common human resources. (See Bracken (1984) for further discussion.)

However, early work in the sociology of language followed in the tradition which viewed culturally distinct languages as related hierarchically, with the languages of the dominant social classes at the top of the “intellectual” scale. According to Dittmar (1976), Schatzmann and Strauss (1955) were the first to formulate what he terms “the Deficit Hypothesis,” the view that the linguistic abilities of particular social groups are deficient or restricted in some way. Schatzmann and Strauss (1955) interviewed members of the lower and middle class about their impressions and experiences after the

---

<sup>7</sup>Technical terms like “deep structure” and “surface structure” must be understood only within the theory in which they are embedded. What the *Port Royale* grammarians meant by a phrase which is often translated as “deep structure” was completely semantic in nature. Their theory of language was certainly not sufficiently developed to include anything like autonomous syntax, but the important distinction between surface forms and their multiple readings was clear. (Frederick J. Newmeyer, personal communication.)

occurrence of a disaster and found that the former used lots of emotional language which reputedly gave rise to “elliptical syntax.” Accordingly, Schatzmann and Strauss (1955) concluded that the lower classes only conveyed their meaning “implicitly,” while the educated classes conveyed their meaning “explicitly.”

This and other work led Basil Bernstein (1971) in 1958 to formulate a distinction between “public language” and “formal language,” later termed “restricted code” and “elaborated code.” Bernstein studied speakers of a non-standard dialect in London and characterized their speech as accessing “restricted code” but not “elaborated code.” According to Bernstein (1971), public language is characterized by “fragmentation and logical simplicity.” By contrast, “formal language” or “elaborated code” may be used to express “universal meaning.” For Bernstein, the restricted code expresses meanings which form a proper subset of the range of meanings expressed in the elaborated code. The appropriate remediation, then, “. . . would seem to be to preserve *public* language usage but also to create for the individual the possibility of utilizing a *formal* language” (1971: 54).<sup>8</sup> Since “restricted” speakers are reputedly limited to logically simple

---

<sup>8</sup>Numerous commentators have portrayed Bernstein as positioned squarely within the camp of the “deficit” theorists, as I do here (Trudgill, 1974; Dittmar, 1976; Boocock, 1980; Bennett and LeCompte, 1990), while others have come to his defense (Halliday, 1995; Danzig, 1995; other papers in Sadovnik, 1995, and Atkinson, Davies, and Delamont, 1995). Bernstein himself explicitly took differences in code to be a property of performance, not competence (1971: 173), but added a decidedly Whorfian qualification: “Clearly, specific linguistic rule systems are part of the cultural system, but it has been argued that the linguistic rule system in various ways shapes the cultural system. . . . [W]hich speech codes are realized is a function of the culture acting through social relationships in specific contexts” (173-174). As Dittmar (1976) points out, what makes Bernstein’s view a species of the Deficit Hypothesis is his perspective that the speech of the educated classes is in some way *greater* (more expressive, less elliptical, so on) than the speech of poor people; that is, the characteristics of “better speech” are taken to be precisely those characteristics which poor people *lack*. In my view, Bernstein’s (1971) work is filled with many contradictory statements which leave much room for diverging interpretations of his ideas. Compare Stubbs (1980) and Hurn (1990) on Bernstein and deficit hypotheses.

expressions, many have taken Bernstein's ideas to suggest that the inability on the part of members of the lower classes to use "elaborated code" corresponds to "cognitive deficiencies," deficiencies which schools must "remedy."<sup>9</sup> However, from the perspective outlined in section 1.1, Bernstein's remarks may be taken to assign schools the task of indoctrinating members of the lower classes to believe that their speech is deficient and manifests intellectual limitations, thereby serving a control function and tacitly justifying the role of the working class as ancillary tools in the hands of corporate elites. From the perspective of the poor and working class, this should come as a very peculiar kind of remedy.

The "deficit" approach to the sociology of language was vehemently challenged in the 1970s by numerous educational researchers and sociolinguists, most notably by Walt Wolfram (1969) and William Labov (1970) in their excellent work on the logic and structure of nonstandard English. Despite insights gained in these studies, however, much of the current work in bilingual education proceeds in the spirit of "deficit" views about low-achieving children, confusing differences of form with differences of degree and competence.

---

<sup>9</sup>Trudgill (1976) reviews studies which show that poor children are in fact *not* limited to a "restricted code." Also see Dittmar (1976). Moreover, none of the statistical data which Bernstein (1971: 5-117) uses to distinguish working class speech from middle class speech ("restricted" from "elaborated" code) is linguistically interesting. The frequency with which constructions (can) occur, apart from zero, says nothing about the underlying system of linguistic competence. More importantly, no current work in cognitive science supports the view that one group's language can express a greater range of meanings than another's.

### 1.3 Code Switching and the Ideology of “Cognitive Deficits”

Valdes-Fallis (1978), Commins and Miramontes (1989), Grosjean (1982) and Ramirez and Milk (1986) report that a widely held belief about code switching is that bilinguals engage in it as a coping strategy, a way of dealing with specific deficiencies in both linguistic systems. Indeed, Lili Rabel-Heymann, from the Department of Linguistics at the University of Calgary, contrasts her own German-English bilingualism with the “morphologically and lexically garbled language many half-educated German immigrants practice” (1978: 222). Although her criticisms of code switching are selective and naive,<sup>10</sup> Rabel-Heymann’s comments reflect a widespread and quite mistaken view, perhaps held nearly as commonly today as in the early twentieth century, that education has the effect of making (native) language “better,” with its potential for making poor people sound like members of elite groups and making ethnic minorities sound like whites.

#### 1.3.1 “Semilingualism” and Linguistic Competence

Commins and Miramontes (1989) have suggested that “a popular belief is that children who code-switch (alternately use two languages) do so because they do not command enough pieces in either language to form a complete code; thus, they are considered semilingual” (445). While some version of the semilingualism thesis has

---

<sup>10</sup>Rabel-Heymann (1978) actually appears to take issue with the particular boundaries used by her “half-educated” countrymen. She gives examples of single-word borrowings and code switches at affixation boundaries as “garbled,” while her own mixtures involve switches at major phrase boundaries (salad -- *mit einem Ei* -- why not? / “salad -- *with an egg* -- why not?”). In any case, the overt prescriptivist values are quite clear and very unexpected of a linguist trained in the late twentieth century.

been endorsed by a number of educational researchers (Ringbom, 1962; Hansegård, 1968; Cummins, 1976a; Toukoma and Skutnabb-Kangas, 1977; Skutnabb-Kangas, 1981; Dunn, 1986), there remains no cogent reason to believe that any such state exists for bilinguals.

The notion of “semilingualism,” the idea that a bilingual might lack linguistic competence for any language in her repertoire, was first introduced in a 1962 radio talk by the Swedish philologist Nils Erik Hansegård (who called it *halvspråkighet*), and it was later picked up by Håkan Ringbom (1962) who conjectured that “a period of ‘double semilingualism’” occurs when an individual abandons her native language altogether in favor of a second language (267).<sup>11</sup> Throughout this dissertation, the term “second language” refers to a language acquired relatively late in life; a distinction between child second language and adult second language will be drawn when relevant.

Before discussing the details of Hansegård’s ideas, it is worthwhile to consider what sense may be given to the notion “lack of linguistic competence” entailed by the semilingualism thesis. Variations of semilingualism are used extensively in discussions of bilingualism and second language acquisition, but it is sometimes a difficult concept to make sense of.

Chomsky’s earliest publications in generative grammar (1955, 1957), concerned with the development of mathematical models of syntax, were devoted to showing that a

---

<sup>11</sup>The historical sketch given here is due to Skutnabb-Kangas (1981). I am also indebted to Skutnabb-Kangas and to her translators Lars Malmberg and David Crane for the English summary of the Swedish literature which appears in Skutnabb-Kangas (1981).

hybrid generative-transformational model would be required to generate even a small portion of a human language. Given this aim, there was no need to look beyond the language most accessible to his readers (English) for examples and counter-examples. Later work (Chomsky, 1965) focused on refining the generative-transformational model.

Through the mid-1970s, however, Chomsky's (1965) conception of the "ideal speaker-listener in a completely homogeneous speech-community" caused considerable misunderstanding among a number of linguists interested in language variation, as Newmeyer (1983) has pointed out. For example, Cedergren and Sankoff (1974: 335) mocked what they took to be the generativists' futile "search for ideal speakers," and Anshen (1975: 6) complained that "one never seems to find 'the ideal speaker-hearer in the perfectly homogeneous speech community.'" Ringen (1975: 26) pointed out that "there is evidence that actual speech communities are not homogeneous."

However, Chomsky's idealization was nothing new to the study of grammar. Many others, notably Ferdinand de Saussure, Martin Joos and Josef Vachek, had proposed similar notions--never because linguists believed that languages actually do not differ, but simply because idealizations are often methodologically convenient, as in other scientific inquiry. Those who accept Chomsky's idealization no more believe that speaker-hearers are in reality "completely homogeneous" than biologists believe that pulmonary or cardiovascular systems are homogeneous across individuals; similar methodological idealizations are used in the study of these organs as well. Indeed, the same book which posits the speaker-listener idealization puts forth the notion of a "Language Acquisition Device" (LAD), the central task of which is to determine a

grammar compatible with the language data of a particular linguistic community  
(Chomsky, 1965)!

In fact, given the particular formalism of Universal Grammar, it makes sense to think of a language as a state of the language faculty that is “some accidental product of varied experience” (Chomsky, 1995a: 7). Thus, it is conceivable--quite likely, in fact--that each of us differs in *some* respect in terms of our knowledge of language or linguistic competence.<sup>12</sup> As Chomsky (1993a: 20) has said,

If my granddaughter were to say “I brang the book,” we would not hesitate to say she is following the rule for “sing-sang-sung,” contrary to “common agreement.” True, her internal language may change, replacing “brang” with “brought.” If it does not, she’ll be speaking a language that differs from mine in this among many other respects, and speaking it “correctly,” insofar as the word means anything.

In his treatment of creole variation, which I shall return to in chapter 2, Bickerton (1973a, 1973b, 1973c) avoids some of the pitfalls which Ringen and others have fallen into by casting variation in terms of “the dynamic paradigm.” On this view, it is assumed that “polylectal grammars” exist for the community as a whole, but that only a subset of such grammars is instantiated for an individual. Thus, Bickerton’s “dynamic paradigm retains the concept of the autonomous grammar-forming individual at the cost of rejecting the principle (held tacitly or explicitly by virtually all other persuasions) that individuals and community grammars are isomorphic” (1973a: 25). Unlike others interested specifically in language variation, Bickerton affirms “the independence of grammar from context” and rejects the “incorporation of social or contextual factors in

---

<sup>12</sup>For some very interesting experimental work which documents individual differences among English speakers in terms of native intuitions, see Gleitman and Gleitman (1979) and Ross (1979).

grammar” (1975: 184). In addition, after noting that distinctions between realized and unrealized complementizers are overtly marked in Hawaiian Pidgin English but not in the several languages which came into contact with it, and that such markers also appear in numerous other, completely unrelated creoles, Bickerton (1982) appealed to Universal Grammar for an explanation:<sup>13</sup>

There is only one hypothesis that will account for both phenomena: the hypothesis of an innate bioprogram for language, which, instead of imposing outer limits on possible forms of language, specifies a set of highly particularized, substantive structures which are accessible to the child . . . [26-27].

Thus, if the idea that someone “lacks linguistic competence” is simply equated with a *difference* in linguistic competence (different, say, as compared with other members of the community), then it will likely be true that all of us “lack linguistic competence” in some respect. This, then, would not be a useful notion.

Nevertheless, since there are respects in which language learners may be presumed to be very much the same, by virtue of UG or the innate “bioprogram,” studies which have addressed language impairments within the context of a well-articulated theory have led to a meaningful characterization of “impairment.” Curtiss (1977), for instance, assessed the linguistic capabilities of Genie, a child who was tragically isolated in a small bedroom from twenty months of age until the age of thirteen and a half, by testing Genie’s knowledge of Ross’s constraints, *wh*-movement, closed class morphological items, and other aspects of universal grammar. In more recent work,

---

<sup>13</sup>In a recent paper, Luján and Parodi (1996) present a similar example of language change in Andean and Los Angeles Spanish in the context of the minimalist program, also based upon assumptions of a biologically innate Universal Grammar.

Curtiss and Schaeffer (1997) looked for morphological evidence of certain functional categories by examining the speech of children who had undergone left or right hemispherectomies. Studies of children with specific language impairment (SLI) may be reviewed in Watkins and Rice (1994), Rice, Wexler and Cleave (1995), among many others. What these studies share is a reliance upon a precisely formulated theory which makes well-motivated assumptions about universal properties of human language.

In sharp contrast, Hansegård (1968, 1975) characterized “semilingualism” in the absence of any theory of language, as “a confused grab-bag of prescriptive and descriptive components,” as Edelsky *et al.* (1983: 2) have put it. For Hansegård, the term denoted a lack of competence in all languages an individual knows in any of six areas: (a) size of the repertoire of words and phrases which are understood or actively available in speech; (b) linguistic correctness; (c) degree of automatism; (d) ability to create or neologize; (e) mastery of the cognitive, emotive and volitional function of language; (f) richness or poorness in individual meanings (that is, whether reading or listening to a particular linguistic system “evokes lively and reverberating semantic images or not”) (Hansegård, 1975: 8; cited in Skutnabb-Kangas, 1981: 253). Skutnabb-Kangas (1981) reviews a number of attempts to test Hansegård’s definition empirically, concluding that tests for the first three criteria show that “semilingualism” has not been found to hold for any population but that little work has been done regarding the last three of Hansegård’s criteria. In fact, with the possible exceptions of (a) size, (c) automatism and (d) ability to neologize, little sense may be given to Hansegård’s grab-bag of “special linguistic abilities.”

Importantly, little sense may be made of his notion of linguistic correctness (Hansegård's criterion (b)). As I have emphasized above, since grammars may differ across individuals, it is hard to know how to count an utterance as "linguistically correct." If a person's linguistic behavior corresponds to a high degree with the linguistic behavior of a community, his behavior might be described as socially "correct," but it is "linguistically correct" if and only if it corresponds to some structural description of his individual linguistic competence. If linguistic evidence suggests a genuine impairment in a person's language faculty in the context of a precisely formulated theory of linguistic competence, then it might make sense to regard such a person as "semilingual," perhaps even "alingual."<sup>14</sup> But no such evidence is advanced for children labeled "semilingual" by Hansegård, Skutnabb-Kangas or others, and a prevalent condition of this sort would be completely unpredicted by what is known about child language acquisition.<sup>15</sup>

---

<sup>14</sup> "Language attrition," a phenomenon in which an  $L_1$  takes on characteristics of a developing  $L_2$ , is studied by Kaufman (1991) and many others. However, there are very important empirical and conceptual differences between the notions of "language attrition" and "semilingualism." As in other cases, none of Hansegård's six criteria are true of children who undergo first language attrition (metamorphosis). Rather, effects of acquiring a second language include changes in the mental representation of a first language under particular conditions set by the environment (social dominance of the  $L_2$ , parental refusal to use  $L_1$ , and so on).

<sup>15</sup>It has been found, for instance, that children learn approximately 10-12 new words per day, during the most active acquisition period (ages 2 to 6), often on one exposure and in highly ambiguous circumstances, and that direct conversational interaction may be irrelevant to this learning process. In fact, children know things about elementary aspects of sentence structure for which they simply have no evidence at all, regardless of the target language involved. These facts and others have led many cognitive scientists to believe that language acquisition is directed by innate principles of universal grammar (Chomsky, 1959), or an internal "bioprogram" (Bickerton, 1981). On these assumptions, a child's failing to acquire a language should be as unusual as a child's failing to grow legs or recognize faces (compare Curtiss, 1989; Mehler and Dupoux, 1994).

Hansegård's idea that the size of linguistic repertoire<sup>16</sup> ought to factor into a notion of linguistic competence is also problematic. Attempts to calculate how many words a person knows are invariably plagued with individual and cultural differences. Judgments about vocabulary size can often be misguided, as they frequently turn on individual differences in interest and facility in talking about particular topics. That the Masai of modern Tanzania do not have a ready command of the topic of French homelife does not indicate a lack of proficiency in Masai, just as a Parisian's inability to readily discuss Tanzanian cattle herding techniques does not indicate a lack of proficiency in French. We naturally expect this difference in vocabulary, given the differences in experience.

Similarly, in regard to Hansegård's final criterion, it is hard to imagine that a person might not get "lively and reverberating semantic images" for some expressions in her languages, as best I can understand what that means. In any case, such "images" would seem to result from a complex interaction of mental faculties and would least of all be a property of the language faculty.

Valadez, MacSwan and Martínez (1997) studied three low-achieving children assessed as having "clinical disfluency," a variety of "semilingualism," and compared their language with that of a control group in terms of grammatical richness (defined in terms of variety of grammatical structures), error rates, and vocabulary. Vocabulary was

---

<sup>16</sup>Hansegård actually understood "size of the repertoire" to include not just words, but also phrases, and these were apparently taken to be either "understood" or "actively available in speech" or both. In light of developments in linguistics since the 1950s, in which phrases are taken to be built up from words by a system of rules, I will construe his notion of "size of repertoire" to include vocabulary only.

tested using a subtest of the CELF (Clinical Evaluation of Language Functions) (Semel and Wiig, 1986). Using the coding and analysis system detailed in Curtiss, Schaeffer, Sano, MacSwan and Masilon (1996), this study found that the experimental group was empirically indistinguishable from the control group, with all children using a rich variety of grammatical constructions with an error rate below 3% and performing equally well on all vocabulary tasks. The results strongly support the claim that children identified by official school mechanisms as “semilingual” may in fact differ in no linguistically interesting way from other children.

In sum, then, Hansegård’s six criteria for semilingualism are advanced with no theoretical motivation. Why should his six criteria be viewed as symptoms of semilingualism any more than any other criteria (say, that the speakers use non-European languages or code switch)? Given the “bioprogram” alluded to by Bickerton (1982), and the biological basis for human language assumed in Chomsky (1959, 1993a, 1995b), Pinker (1994) and many others, an impairment of the sort Hansegård expects to find could only result from severe pathology. Since the differences which Hansegård attempted to locate for proposed “semilinguals” have not been found (while evidence to the contrary has), and since no empirical or theoretical justification has been put forth for

selecting his criteria in the first place, then the semilingualism thesis may be safely discarded.<sup>17</sup>

Finally, as Skutnabb-Kangas (1981) points out, there is a clear similarity between Hansegård's (e) cognitive aspect of language and Cummins' (1976a, 1976b) Threshold Hypothesis. This leads us to the next section.

### **1.3.2 The Threshold Hypothesis and Language Proficiency**

Cummins suggested that the level of “language proficiency” attained by a bilingual child in first and second language may affect her cognitive growth in other domains. In early work, Cummins believed that there were two thresholds, and that attainment beyond the lower threshold “would be sufficient to avoid retardation, but the attainment of a second, higher level of bilingual competence might be necessary to lead to accelerated cognitive growth” (1976a: 24). For Cummins, a child with low levels of proficiency in L<sub>1</sub> and L<sub>2</sub> may suffer “negative cognitive effects.” Once mastery in one language has been obtained, the child has moved beyond the first threshold and will suffer neither positive nor negative effects. Finally, “positive cognitive effects” obtain when a child develops high proficiency in both languages. Thus, Cummins (1976a,

---

<sup>17</sup>It is important to emphasize that the semilingualism thesis advances the claim that there are (non-pathological) cases of language learners who speak no language “correctly.” To deny that such a population exists entails that all (non-pathological) children acquire at least one language as their native language, and, while there may be some differences of word choice and grammatical form, their native languages will be characteristically human. However, a second or third language, perhaps acquired later in life, may be markedly non-fluent, inexpressive, and may have a relatively impoverished vocabulary. I will return to these issues in section 2.1.

1976b, 1981),<sup>18</sup> like Toukomaa and Skutnabb-Kangas (1977)<sup>19</sup> and many others, incorporates a version of the semilingualism thesis into his model, purporting that a school-age child may lack proficiency for all languages in her repertoire, a view which has not been sustained empirically.<sup>20</sup>

However, Cummins provides a more carefully considered definition of “language proficiency” than did Hansegård. Drawing from Canale and Swain (1980) and Canale (1981), Cummins unpacks language proficiency into the four basic components of grammatical competence, sociolinguistic competence, discourse competence, and strategic competence. Grammatical competence, of course, designates knowledge of syntax, morphology, phonology, and the lexicon. Cummins’ (1981) definition of “sociolinguistic competence” is essentially the same as Chomsky’s definition of “pragmatic competence,” restricted to “knowledge of conditions and manner of

---

<sup>18</sup>Cummins’ (1979, 1994) well-known BICS/CALP distinction has also been criticized for ranking varieties of language into “basic-interpersonal” and “cognitive-academic,” with the underlying assumption that the latter somehow manifests greater cognitive ability. However, Canale and Swain (1980) and others’ criticisms of this distinction led Cummins to posit his famous juxtapositions of “cognitively demanding” with “cognitively undemanding” and of “context embedded” with “context reduced” (see Cummins (1981, 1994) for discussion). While these ideas made some very positive political advancements for bilingual learners who had acquired only enough English to appear fluent on the playground, considerably more work must be done to properly characterize the learning situation of bilinguals.

<sup>19</sup>In fact, Cummins (1981: 39) illustrates his Threshold Hypothesis with a figure adapted from Toukomaa and Skutnabb-Kangas (1977: 29) in which he uses the phrase “limited bilingualism” where “semilingualism” is used in the same figure reproduced in Skutnabb-Kangas (1981: 223).

<sup>20</sup>Cummins (1981: 36) cites three studies which he takes to have reported findings consistent with the original model of Cummins (1976a). However, the studies he cites only find that high levels of bilingualism co-occur with strong performance on a variety of cognitive tasks. These findings do not rule out that some third factor (such as self-esteem, perceived self-efficacy, ethnic identity, personal values, so on) did not independently result in both achievements. In any event, no evidence has been found that children with “limited bilingualism” (“low level in both languages”) actually exist, much less that they suffer “negative cognitive effects.”

appropriate use, in conformity with various purposes” (Chomsky, 1978: 224). For Cummins (1981: 7), discourse competence consists in “knowledge of how to combine meanings and forms to achieve a unified text in different modes.” Finally, Cummins takes strategic competence as the “mastery of verbal and non-verbal strategies” which compensate for breakdowns in other domains of competence and for performance limitations (his examples include use of dictionaries, ability to paraphrase, and gestures). After some discussion of Canale and Swain (1980) and specific developmental issues in second language learning, Cummins (1981) settles on a framework in which “literacy is viewed as one aspect of communicative proficiency” (14).<sup>21</sup>

This view of language “proficiency” is consistent with important legal developments in bilingual education. As Macías (1993) notes,

The LEP (and non-English proficient) category was defined as a subgroup of language minority/non-English language background whose English proficiency was not sufficient for them to participate effectively in an English-only classroom. The term “limited-English speaking” ability was derived from the federal Bilingual Education Act legislation of 1968. The Bilingual Education Act amendments in 1978 added *reading and writing* English to the definition and the term became “limited-English *proficient*.” To be English *proficient*, then, meant a person had to be able to speak, understand, read and write English, not just speak it [232].

When applied to a situation in which English is a second language, it certainly makes sense to require “proficiency,” as defined here, before placing bilingual children in

---

<sup>21</sup>Cummins uses the terms “language proficiency” and “communicative proficiency” synonymously (Cummins, 1981: 7, n2).

English-only classrooms.<sup>22</sup> In this case, saying that a child is proficient in a second language is simply saying that she knows it reasonably well, and that she can read and write in it.

However, a problem emerges when this notion of “proficiency” is used to describe first language ability because it suggests that literacy is a normal aspect of language development. In contrast to the way in which children learn to read and write, however, a native language is acquired effortlessly and without instruction.<sup>23</sup> As Chomsky (1993a: 29) has recently put it, “Language acquisition is something that happens to a child placed in a certain environment, not something the child does.” While considerable work has been advanced within this framework for language acquisition (see Goodluck (1991) and Atkinson (1992) for summaries), nothing suggests that children acquire the ability to read in any similar way.<sup>24</sup> Success in reading is dependent upon direct or tacit instruction, practice, and considerable effort, like success in other academic endeavors (Gough and Hillinger, 1980; Perfetti, 1985; Adams, 1994).

In addition, taking literacy as an aspect of language proficiency suggests that members of non-literate cultures and societies have relatively *low* “language

---

<sup>22</sup>For a detailed discussion of the literature on bilingual education programs and their advantages, see Wong Fillmore and Valadez (1988), Macías (1989, 1993), Krashen (1996), and Rossell and Baker (1997).

<sup>23</sup>Some effort was made in the 1970s to link child language acquisition to “motherese,” a tacit form of instruction, but this research program was not successful. However, see Fernald (1985) for an interesting discussion of prosodic characteristics to which infants appear to attend especially well. See also Wexler and Culicover (1980) and Curtiss (1989) for reviews of research on “motherese.”

<sup>24</sup>Certainly lots of linguistic information is *used* in decoding text, but this fact does not comment on the way in which decoding skills are acquired.

proficiency,” whether they are monolingual or bilingual, in contrast to many “highly proficient” Western cultures. In an interesting study of the Vai, a Mande-speaking people of northwestern Liberia,<sup>25</sup> Scribner and Cole (1981) reported that literacy had very specific effects determined by the purposes to which it was put (memorizing the Qur’an appeared to improve linguistic memory skills, while knowledge of Vai script appeared to improve ability to accurately retell the rules for playing a board game); the various tasks used to study Vai literates and nonliterates found nothing “that would lead us to speak of cognitive consequences of literacy with the notion in mind that such consequences affect intellectual performance in all tasks to which the human mind is put” (70). Similarly, reading ability should no more be taken as an indicator of general “high language proficiency” than should the ability to write rap songs, decode graffiti, or commit long spoken texts to memory for retelling to children and grandchildren for generations to come. So far as anybody knows, there are no general, positive cognitive effects of literacy.

Thus, analyzing language proficiency or linguistic competence to include literacy skills is a poor way to conceptualize the learning situation of schoolchildren. Not only does such a conception lack theoretical or empirical motivation, it also has the effect of

---

<sup>25</sup>The Vai constitute a unique population. As one of the few cultures to have independently invented a phonetic writing system, Vai script is transmitted without formal schooling. In Vai society, literates and non-literates share common material and social conditions, allowing “for a more direct test of the relationship between literacy and thinking than is possible in our own society” (Scribner and Cole, 1981: 62).

incorrectly characterizing the learning situation of pre-literate LEP children.<sup>26</sup> Literacy is an important aspect of school curriculum for all children. Children labeled “semilingual,” but who are in reality improperly assessed, perfectly competent LEP first-language users, are placed at a marked disadvantage when learning to read or when engaging in other academic tasks.

Despite strong criticisms (Edelsky *et al.*, 1983; McGroarty, 1988; Cummins and Miramontes, 1989; Valadez, 1995; Valadez, MacSwan and Martínez, 1997), Cummins’ views remain the dominant influence in bilingual education today.<sup>27</sup> Other species of the semilingualism thesis have also received a great deal of criticism (Labov, 1972; Shuy, 1978; Heath, 1982a), but Cummins and Miramontes (1989) assert that the doctrine is nevertheless very often used by teaching practitioners in elementary and secondary schools today as an explanation for academic failure on the part of bilinguals.

Rather than helping children, these ideas stand to do serious harm; they are ideological constructs, effectively serving to promote the interests of elite groups by promoting tracking and privileged expectations for certain majority children, as outlined in section 1.1. The view that code-switched speech corresponds to a deficit by any name, cognitive or linguistic, can no more be sustained than the nineteenth century view that

---

<sup>26</sup>Given his conception of language proficiency, it may be that Cummins’ (1981: 39) version of the semilingualism thesis (“low level in both languages”) denotes a lack of literacy skills rather than an actual deficit in the language faculty. However, since no distinction is made in Cummins’ work between acquiring a language and learning to read, the Threshold Hypothesis does not guarantee this interpretation.

<sup>27</sup>For evidence of this, peruse any textbook designed for use in a bilingual teacher education program, such as Díaz-Rico and Weed (1995), or any state document regarding bilingual education, such as Leyba (1995) or Holt and Tempes (1982).

African and Native American languages are impoverished in comparison to European languages.

#### 1.4 *The Ann Arbor Decision, Code Switching, and Language Education*

Despite work by Wolfram (1969), Labov (1970, 1975) and many others, schools continue to regard the native dialects of some students as deficient. In July, 1979, a federal suit filed in the Eastern District Court in the state of Michigan provided a significant legal precedent for the determination of school districts' legal and educational responsibilities with regard to the particular issue of Vernacular Black English (also called Ebonics or African American Vernacular English [AAVE]) in public schools. The final disposition of *Martin Luther King Junior Elementary School Children et al. v. Ann Arbor School District Board* has come to be commonly called the "Ann Arbor decision."

In this case, parents charged that the school did not facilitate equal participation in academic programs for their AAVE-speaking children, and that the school had not recognized the widespread use of Black dialect and its implications for the teaching of reading and "standard" English. In court, William Labov and J. L. Dillard discussed the phonological and syntactic characteristics of Black English, in addition to educational barriers confronted by these children, "the most damaging of which is the tendency of teachers to make such speakers ashamed of their native dialect by teaching standard

English without recognizing that the child uses a dialect acceptable to his linguistic community,” according to Bountress (1982: 79-80).<sup>28</sup>

Judge Joiner directed the counsel for the school board to submit a plan for helping teachers work with students who speak AAVE. The Ann Arbor Board of Education submitted a plan that specified in-service training for teachers in the district on topics of language development, language variation, the special characteristics of Black English, the nature of bilingual code switching, and ways of using such information in helping students learn to read.

As the judge himself anticipated, the Ann Arbor decision had important implications for bilingual children as well. Proficient bilinguals<sup>29</sup> who code switch are often viewed as linguistically deficient, just as children who speak AAVE are, as comments by Valdes-Fallis (1978), Commins and Miramontes (1979), Grosjean (1982) and Ramirez and Milk (1986) affirm. The Ann Arbor decision asserts the legal responsibility of schools to understand the language situations of the children they educate, and to treat their home language as equally legitimate as the home language of any other child.

Heath (1982b) has also pointed to specific cultural dimensions of language use in the schools which may lead to inequality. Arguing that language use and socialization among primary school children in some African-American groups differ drastically from

---

<sup>28</sup>Also see Straker (1980).

<sup>29</sup>See section 2.1 for a discussion of the notion of “bilingual proficiency.”

language use and socialization in the dominant culture, Heath (1982b) argued that schools frequently place such children at a marked disadvantage when the curriculum is tailored to the interests of the dominant classes. In this respect, Heath contended, the failure of these children is *anticipated* by the curriculum, and the success of children from the dominant classes is facilitated. Commins and Miramontes (1989) similarly argue that the skills of children who teachers take to be “semilingual” or limited to BICS are ignored by school curricula, and the failure of these children is often similarly constructed by the institution which purports to educate them.

Genuine language education in the schools, in which language is studied as an object of the natural world, could have the long-term effect of changing attitudes among teachers, children, and, conceivably, society at large, as suggested by O’Neil (1971). In addition, if African-Americans, bilingual children and others studied their language situations from a position of linguistic equality in the spirit of intellectual inquiry, the reflective and thoughtful use of language cherished by prescriptivists could emerge as a natural consequence. Finally, O’Neil and Honda (1993, 1994) have shown that the scientific study of language may be used to teach children about principles of science and rational inquiry, and to develop curiosity. When unfounded language attitudes and proscriptions give way to genuine linguistic inquiry in school curricula, then language education will also have the potential to take on a genuine teaching function in schools.

### *1.5 Bilingualism and Placement in Special Education*

Finally, Ortiz and Yates (1983) report that more than three times as many Latino/Latina students were classified as “learning disabled” in Texas than would be

expected from their proportion in the school population. Indeed, as Artiles and Trent (1994) point out in a survey of literature and data extending from 1968 to 1993, the overrepresentation problem has not improved. Furthermore, the misplacement of language-minority children in special education programs has often been attributed to a mistaken understanding of language and its relation to culture and intellectual ability. Cummins (1984) provides an extensive review of teacher and school psychologist referrals for special education in which failure among language-minority children is attributed to cognitive deficiency or a lack of motivation.

In contrast to the attribution of the causes for minority-student failure articulated in these referral reports, Artiles and Trent (1994) attribute this “construction of school failure” in part to values of competition and evaluation, rooted in the dominant American culture and institutionalized in schools. Indeed, it has been reported that the diagnosis of “almost 75% of mild mental retardation is linked to various SES-related environmental contingencies” (Baumeister, Kupstas and Klindworth, 1990, cited in Artiles and Trent, 1994).

Here again, greater insight into the nature of bilingualism may illuminate the concerns of special education teachers and school psychologists. Children with genuine learning handicaps may receive the special attention they need, while language-minority students, properly understood, are set free from the debilitating effects of labeling and of alienating school labor.

## 1.6 Conclusions

An enriched understanding of linguistic aspects of code switching has the potential of changing teachers' beliefs about their students' abilities and accomplishments, just as Wolfram's (1969) and Labov's (1972, 1975) work has successfully done, at least to some degree, with regard to Black English. The treatment of children by adults is, without a doubt, the most crucial variable affecting their success or failure in school (Johnson, 1969; Coates, 1972; Crowl and MacGinitie, 1974; Oakes, 1985; Clark and Peterson, 1986). Because bilinguals who code switch are frequently viewed as deficient in both their languages, they are often viewed as incapable of the sort of progress normally expected of children learning to read, write, and use language creatively. In some cases, these children are even institutionally labeled as "special education students" or "clinically disfluent" and assigned to classrooms reserved for "the learning disabled" or the "mildly retarded."

Artiles and Trent (1994) point out that

The notion of disability is concerned with atypical functioning or educational performance due to biological, psychological, and/or social factors. The level of functioning for individuals with disabilities falls in the lower portion of the normal distribution curve. The notion of disability exists because we have established parameters to judge when a person functions anatomically, physiologically, intellectually, and/or psychosocially within the limits of what is considered typical. On the other hand, cultural diversity is not defined--at least theoretically--by a standard parameter of functioning. Although it is also concerned with the idea of difference, it is not--unlike the disability construct--inherently linked to the notion of deviance [424].

This dissertation pursues the thesis that there are no definable linguistic parameters by virtue of which bilinguals who code switch may be distinguished from

monolinguals or from bilinguals who do not code switch. To demonstrate this thesis, we assume that

- (a) nothing constrains<sup>30</sup> code switching apart from the requirements of the mixed grammars; and
- (b) code switchers have the same grammatical competence<sup>31</sup> as monolinguals for the languages they use.

To the extent that (a) and (b) lead to successful analyses of code switching data in chapter 5, they may be regarded as correct.<sup>32</sup> (For present purposes, and for reasons which will be clarified in section 2.1, a “code switcher” is an individual who alternately uses two or more languages at or below sentential boundaries, and who has had continual, sustained exposure and practice in these languages since infancy.)

A possible criticism of the work on code switching reported in this dissertation is that it is highly theoretical and often quite technical, and hence possibly inaccessible to the teachers and policy makers it is intended to inform. However, in proper historical context, Labov’s work on non-standard English might have been seen as equally

---

<sup>30</sup>As will be made clear in chapter 5, “constrain” is used here in a technical sense: there can be no rule or principle of grammar which refers to code switching. There are obviously pragmatic constraints, but these too may follow from the same principles which determine choice of speech register in monolingual speech.

<sup>31</sup>Here we return to the idealized notion of linguistic competence of Chomsky (1964). Again, while individual differences are expected for all members of a speech community, the assumption that the speech community is fairly homogeneous is reasonable and methodologically appropriate.

<sup>32</sup>Logically the claim is of the form ‘All Ps are Q’ (that is, the grammaticality judgments for all code-switched sentences may be explained in terms of independently motivated principles of the relevant grammars and nothing more). To prove this, we show that in a reasonable sample of Ps, all are Q. Our “reasonable sample” will consist of all the facts of the Spanish-Nahuatl corpus obtained for this dissertation, plus data from other corpora where we expect to find counter-examples. This material is presented in chapter 5.

technical, as it drew upon recent developments in abstract phonology of the sort discussed at length in Chomsky and Halle (1968).

The primary role of such work is to impact on language education research as a whole and ultimately contribute to the formation of new, improved frameworks for language education. The present work, to the extent that it leads to an enriched understanding of bilingualism, may have a positive impact upon policy, teaching and curriculum. I will return to these topics in chapter 5, drawing some connections between the basic research developed here and its impact upon practice.