INDICATORS OF BILINGUAL PROFICIENCY IN NEPALI AMONG TIBETO-BURMAN PEOPLES OF NEPAL

Jeffrey D. Webster
University of Texas at Arlington

The myth of bilingualism suggests that everyone in Nepal is fluent in Nepali; however, this study shows large segments of the Tibeto-Burman population are only minimally proficient in speaking and understanding Nepali.

This paper presents data on bilingual proficiency collected in 1992 and 1993 from speakers in seven different Tibeto-Burman ethnolinguistic groups in rural Nepal. Over five hundred subjects were tested, using sentence repetition tests. These data are correlated with demographic variables to show the significant indicators of Nepali proficiency among the target populations.

The results show that only well-educated subjects are comparably proficient with mother-tongue Nepali speakers. Because the same demographic factors in different areas produce parallel patterns of second-language proficiency, the most significant factors can comprise a model to substantially predict Nepali bilingualism. Based on extensive field data, these results provide a powerful tool for language policy and planning in Nepal.

Rationale

The democratic revolution in early 1990 ushered in a whole new era for language policy and planning in Nepal. Nepal's one-language policy was radically changed by the new constitution which now "guarantees the fundamental right of the individual to receive primary education in his/her mother tongue" (Yadav 1992:178).

Practically speaking, though, the difficulties of language development and the huge expense of producing basic literacy materials in dozens of minority languages make the implementation of this right a formidable task.

Only with a concerted collaborative effort between Nepalese and expatriate linguists and literacy specialists will Nepal's minority-language groups gain the recognition and advantage they are seeking with primary, transitional, mother-tongue education (meaning a bilingual education program which introduces literacy and primary subjects through the medium of the mother tongue, then gradually transitions to monolingual Nepali instruction as students' command of Nepali allows).

Since 1990 many minority-language groups have organized their own non-governmental organizations (NGOs). These NGOs are petitioning the government for assistance in establishing transitional mother-tongue education programs in their languages. These NGOs are also attempting to revitalize their cultures by promoting in their communities their own indigenous language, culture, and religion. Ethnic groups such as Magar, Gurung, Chepang, Sherpa, Tharu, Rai, Limbu and others have all formed such groups. Of these, the Magar, Gurung, and Limbu are represented in this research.

The myth of national bilingualism

There is a common misconception in Nepal—throughout all of South Asia in my experience—that "virtually everyone speaks the national language." To be Nepalese is to speak the Nepali language. This myth of bilingualism is partly semantic in nature, ignoring the issue of bilingual proficiency, and obscuring the fact that not everyone who speaks a second language does so as fluently as a mother-tongue speaker.

As used in this paper bilingualism is the ability to speak two (or more) languages, without reference to the degree of fluency in the second language. Most Nepalese citizens have been exposed to Nepali and have at least some passive ability in the language, some bilingual ability. But being able to correctly respond to the question, "Where are you from?" is entirely different from being able to discuss religion, philosophy, or highly emotional issues. To call both abilities "bilingual" is a gross oversimplification which obscures reality. Only those whose Nepali ability (i.e., bilingual proficiency) allows them to control deeply felt and abstract material are linguistically able to be full participants in the nation-state.

Research questions

A fundamental question which Nepal's aspiring ethnic groups need to address is this: do they have a legitimate functional need for mother-tongue education? Any group can claim need based on ethnic identity and the desire for cultural promotion. Functional need, however, refers to the inability, on the part of a significant portion of the language group, to

adequately control oral and written Nepali. The main question underlying any discussion of a group's need for mother-tongue education is the primary question of this study: how proficient in the Nepali language are the Tibeto-Burman ethnolinguistic groups of Nepal?

Definitively answering the question of "how proficient" would require detailed proficiency testing of millions of people—an impossible feat! The assessment of demographic factors, however, is something routinely accomplished in censuses. The detailed correlation of demographic factors and proficiency test results has the potential to provide a tool which will give us broad insights into community-wide Nepali proficiency. This research seeks to provide the basis for such a tool by testing a broad sample of subjects from seven different ethno-linguistic groups and comparing their Nepali proficiency with demographic factors. The driving question is this: can we gain an adequate understanding of Nepali proficiency by just looking at relevant demographic factors?

Choice of test instrument

Answering these research questions requires a test instrument appropriate for testing large samples of people in the field, including uneducated and illiterate people.

The test instrument must be quick to administer and easy for subjects to understand what is required of them; it must be easily adaptable for use in multiple language groups; it should not require administrators to know the subject's mother tongue; and it should discriminate a wide range of second-language proficiencies. Only the sentence repetition test (SRT) as developed by Radloff (1991) meets all of these criteria and is therefore suitable for this kind of testing.

Though an SRT is quite time consuming and difficult to develop, once developed it can be administered in less than five minutes per subject. Radloff's contribution to bilingualism studies has now made it feasible to evaluate large samples in a very short time and to address questions that are the focus of this study.

Nepali variation and standardization

Nepali, spoken as it is over a large, mountainous area where illiteracy predominates, has a number of different dialects. Largely due to the widespread increase in education and communication in Nepal since 1950, the speech of the educated Brahmins and Chetris of Kathmandu is becoming the standard variety of Nepali (Matthews 1992:ii). The publication of the Gorkha patra, previously a weekly and now a daily newspaper, and standardization efforts of the Royal Nepal Academy, have

also led to a large measure of standardization based on the dialect spoken by educated residents of Kathmandu (Verma 1992:77).

The varieties of Nepali, however, are still very much in a state of flux. Michael Hutt (1988:51-52) documents the development of a standard form of Nepali, but also cites the growing divergence between the standard spoken variety and the variety used in literature:

> Due to the often formal contexts of usage of written Nepali, however, a 'high' variety has evolved which is becoming increasingly distant from its spoken counterpart. In fact, there could be said to be three forms of standard Nepali: the spoken norm, based on an ideal version of the Valley dialect, the ordinary written language of everyday literature and the formal 'high' variety of scholars and administrators. Trends of change in the latter two varieties are proceeding in almost exactly opposite directions: in popular literature, colloquialism is becoming more highly prized, in scholarly works more and more words are being borrowed from Sanskrit or English. . . . Scholarly written Nepali and its high-grade spoken counterpart have diverged from the colloquial speech of Kathmandu to the extent that the situation has become almost diglossic.

Matthews (1992:ii), with a more hopeful view of standardization, summarizes this issue which is of great consequence for this study of second-language proficiency in Nepal:

> [There is a] fairly big difference which exists between the spoken and written language. The latter, which is employed in most printed works, newspapers and broadcasts of Radio Nepal, is characterised by the large number of words taken over from Sanskrit, a more or less consistent use of grammatical gender and certain verbal forms which feature only sporadically in everyday speech. Although in the first half of this century, the literary language would have been regarded as artificial, and understood only by the educated elite, an increase in literacy and the growing use of the transistor and in some areas the television have radically altered the situation. The effect of the written style on speech has given rise to a certain inconsistency, and it is no longer uncommon to hear both literary and colloquial forms of the same word in the space of a few sentences.

I would agree that recent changes have "radically altered the situation" regarding the spread of what I refer to here as "Educated Nepali"; however, as I show, the majority of the population of Nepal still do not control Educated Nepali speech near as well as do educated mother-tongue speakers of Nepali in Kathmandu.

Diglossia

The concept of diglossia should be mentioned here because one distinctive of diglossia is that it refers to a significant difference between written and spoken varieties. The important question for this study is not whether or not Nepali is a case of diglossia, but in what ways does Nepali have some of the characteristics of diglossia.

The term diglossia was first used by Charles Ferguson in 1959 to describe a language situation in which two varieties of the same language exist side by side and are used for different functions (1959:328). A more prestigious, standardized, and formal variety is termed the High dialect (H); its informally learned and used counterpart is the Low dialect (L). Fishman revised and extended the concept in 1967 to include unrelated languages as well as dialects of the same language (Fishman 1967:74). Ferguson explained diglossia with reference to nine categories: function, prestige, literary heritage, acquisition, standardization, stability, grammar, lexicon, and phonology. The varieties of Nepali fit his concept in many ways, though not entirely. Ferguson's nine categories are summarized in Table 1 (following Britto 1986:8-9), with reference to how well Nepali fits into this model. Written Nepali is here labeled the H variety; colloquial Nepali is labeled L.

Though Nepali satisfies many of the features of Ferguson's diglossia, and likewise for Fishman's revision of the term which explicitly includes stylistic variations of the same language, it is still unclear whether or not it satisfies a fundamental requirement of either Ferguson's or Fishman's definition of diglossia, i.e., that no segment of the speech community uses H in normal conversation. I think that this requirement does hold true for spoken Nepali: educated Brahmins (upon whose speech Educated Nepali is based) speak a variety of Nepali substantially different from what has become the standard written form. Control test results in Table 5 give credence to this conclusion. Uneducated mother-tongue speakers of Nepali do not score as highly on the Nepali SRT as do those who are educated.

Table 1: Ferguson's diglossia, with rerefernce to Nepali

Category	Characteristic of H:	Characteristic of L:	Nepali situation
Function	Used in formal writing, speeches, and such functions	Used in everyday conversations	H used almost exclusively in formal functions
Prestige	More prestigious	Less prestigious	H has more prestige
Acquisition	Formally, in school	Informally	With few exceptions, H is learned in school
Standardi-zation	Highly standardized	Poorly standardized	H is rapidly being standardized; L is not
Literary heritage	Large, esteemed literature	Small, less highly esteemed literature	H has fairly small, but esteemed, literature
Stability	Autonomous and stable with some interference from L	Autonomous and stable with some interference from H	H is growing more stable as it gets more standardized. H and L are more autonomous as education spreads
Lexicon	Mostly shared with L except for technical words; some paired with L	Mostly shared with H, but with some paired with H	Basic lexicon is the same, but with large technical vocabulary lacking in L. Absence of paired items
Phonology	A shared phonological system with L	A shared phonological system with H	Shared system
Grammar	More complex	Simpler	H is more complex: e.g. H has several register levels and full agreement system

Choice of test language for Nepali SRT

The diglossic characteristics of Nepali pose a problem for a study of bilingual proficiency in Nepali, if by definition one must be educated to be well-versed in the test language; uneducated mother-tongue speakers of Nepali will not necessarily be as proficient as those who are educated.

It is in this context that this study was conducted, and the choice of the best variety of Nepali to use for proficiency testing needed to be made. For the results of this study to have application to non-formal education and language policy and planning, the test language must be the one which predominates in the media and education of Nepal. What I call Educated Nepali is the only logical choice.

Research scope: Tibeto-Burman languages of Nepal

Second language (L2) proficiency is of two types: inherent and learned (Grimes 1987:7). Inherent ability to understand a related speech variety is relatively uniform throughout a speech community because the nature of the relationship between the two speech varieties is essentially constant if various contact variables are controlled (Milroy 1987:21). Learned ability varies widely among individuals and subsegments of a community according to the nature and degree of their contact with the second language. This contact varies on an individual basis, and is closely related to demographic factors like education, sex, age, and degree of travel.

The type of language proficiency studied in this research is learned as opposed to inherent ability. All of the test subjects in this research speak as their first language (L1, or mother tongue) a language in the Tibeto-Burman language family. Nepali, the second language in question here, is an Indo-Aryan language. One would not expect any inherent understanding between speakers of the two language families, nor would one expect one of these languages to be linguistically closer to Nepali than any of the other languages.

The results of this research should be applicable to other Tibeto-Burman languages throughout Nepal. Where the first language is not Tibeto-Burman, however, care must be exercised in applying these results. Languages more closely related to Nepali, or communities with a significantly different set of language-contact factors, will likely have significantly different profiles of Nepali proficiency (Webster 1993a:40-45). However, the principles of this study should be widely applicable to most multilingual situations.

Research areas

Seven ethnolinguistic groups in five research areas were visited for this research. The first four areas are in central Nepal, and the fifth area is in eastern Nepal (see Map 1). These research areas are intended to provide a sample of the Tibeto-Burman groups in Nepal which are representative in terms of their isolation, population, and degree of interethnic contact.

Map 1: Nepal showing of research areas
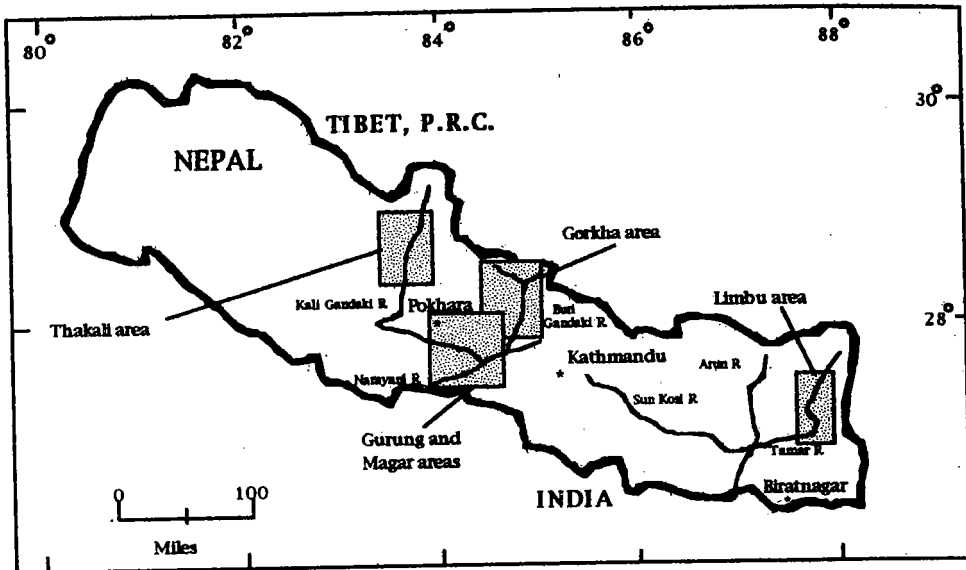
Summary of research areas

This section summarizes the test points in the five research areas in a quantifiable way to facilitate cross-village comparison. In studying interethnic contact in Burkina Faso, Showalter (1991:102) marks specific characteristics which are concrete indicators of relative village isolation and likelihood of interethnic contact. I assume isolation to be inversely related to opportunities for contact with Nepali speakers. I have followed his model here (though the indicators of isolation and opportunity to use Nepali are much different) to arrive at a "L2 exposure" factor for each village.

Table 2 presents the criteria used in assigning point values to the seven attributes of village isolation. The higher the L2 exposure factor, the less isolated and the more opportunities to use Nepali. Table 3 summarizes the key attributes of each research area and quantifies these attributes according to their contribution to the L2 exposure factor. Relative isolation as measured by distance from the main road has the greatest weight: Gorkha villages all score a "0" on this factor showing greater isolation, whereas

Magar villages score a "4" or "5." The Gorkha Group, with an average of only 1.3, is substantially different from other areas. The Gurung area, on average, is slightly less isolated than the other research areas, having the highest average score. Because of small samples in each Gurung village they are combined here.

Table 2: Criteria for quantifying relative opportunity for L2 exposure in research areas

Attribute	Point Value	Criteria for assigning point value
Access to roads	0-5	Based on travel time to the nearest road (5=a road to the village; 0=six to seven days' walk to the nearest road.
to airport	0-1	1= airport within one day's walk; 0= more than one day's walk.
to trade route	0-2	Distance to a major trade route where Nepali would be needed: 2= on route; 1= easy access to route; 0=not near route.
Availability of high school	0-2	2= in village; 1= within a day's walk; 0= outside walking distance.
markets	0-2	2= in village; 1= within a day's walk; 0= more than one day's walk.
government services	0-2	2= many services (police, bank, clinic, etc.); 1= minimal services; 0= no services.
Ethnic mix	0-2	2= 20% or more mixed; 1=1-19% mixed; 0= ethnically homogenous.

Table 3: L2 exposure factors for each research site.

Village	Road	Air-port	Trade	School	Ethnic mix	Market	Gov. serv.	Village Total
Pragatinag	5	0	0	2	2	2	2	13
Arkhala	4	0	0	0	0	0	0	4
Magar (430,000)							Magar average 8.5	
Marpha	1	1	2	2	1	2	2	11

Chhairo	1	1	1	1	1	1	1	7
Chimang	1	1	1	1	1	1	1	7
Thakali (7000)							Thakali average 8.3	
avg village	4	0	0	1.5	2	1.5	1	10
Gurung (228,000)							Gurung average 10.0	
Lho	0	0	0	0	0	0	0	0
Bihi	0	0	0	0	0	0	0	0
Phidim	0	0	2	0	0	0	2	4
Gorkha (15,000)							Gorkha average 1.3	
Gurung (228,000)							Gurung average 10.0	
Hangdewa	4	1	2	1	2	1	0	11
Okre	4	0	0	2	2	1	0	9
Worden	4	0	0	1	2	1	0	8
Limbu (250,000)							Limbu average 9.3	

METHODOLOGY

The sentence repetition test

The sentence repetition test (SRT) used in this study was developed according to the test development methodology presented by Radloff (1991, Chapter 3). An SRT consists of a set of carefully selected and tape-recorded sentences. Subjects listen to each sentence one by one, on headphones, and are evaluated according to a four-point scale (0-3) on their ability to accurately repeat each sentence. Essentially any deviation from the recorded sentences is counted as an error. The sentences are not related in meaning, but gradually increase in length and complexity. In general, and this is true for the Nepali SRT, there are 18 sentences—three practice sentences followed by 15 which count toward the subject's score. The test administrator allows replays of the practice sentences as many times as is necessary to ensure that each subject understands the methodology before proceeding past the practice sentences.

The SRT is an indirect, correlated test. Scores on an SRT have validity by virtue of being calibrated with an external proficiency standard with some validity. This proficiency standard, the reported proficiency evaluation (RPE), consists of the carefully controlled evaluations of mother-tongue speakers.

SRT results are expressed as a point total out of 45 possible points and correspond to RPE levels. These RPE levels range from 0+ (very minimal proficiency) to 4+ (approaching the proficiency of a native speaker).

Probably at least level-3 proficiency is required to adequately understand most philosophical or religious material (Kindell 1991:28). An adult education specialist in Nepal estimates that, because of the higher register language commonly used in written communication, new literates who are not at least RPE level 3 in Nepali have considerable difficulty with Books 3 and 4 of the Nepalese government's Naya Goreto adult literacy program (Glover and Glover 1993:4).

Radloff (1991:8) is careful to emphasize that the SRT is a screening as opposed to a diagnostic test. It provides a general measure of a subject's proficiency rather than a detailed analysis of a subject's strengths and weaknesses. It takes less than five minutes to completely test one subject, including the time for explanation and practice.

Sentence repetition: what does it test?

Sentence repetition, or elicited imitation as it has been called, has been used in language testing since 1963 (Slobin and Welsh 1968). Many early studies were in the fields of childhood language acquisition and language development and disorders. Subsequent applications of the methodology were in studies of second language acquisition, nonstandard dialects, acculturation, and deviant speech (see Gallimore and Tharp 1981, for review of the literature).

What exactly does sentence repetition test? This question has been debated by many researchers—Oller (1979) provides a good summary of this. Those in favor of the technique argue that as long as mimicry material is beyond the grasp of short-term memory it is a valid test of both comprehension and production. Grimes (1992b:58) is specific in her critique of the SRT, arguing that it is "incapable of evaluating lexical discrimination, discourse comprehension, productive use, and situational appropriateness, at the least." Much of the criticism of sentence repetition as a measure of language competence centers on its weakness as a diagnostic measure of specific language capabilities. None of the researchers who question the methodology on these grounds, says Radloff (1991:78), deny that sentence repetition tests assess the general ability to control the grammar, vocabulary, and phonology of a language.

Hamayan et al. (1978) showed that a repetition task does indeed test more than comprehension and production. They concluded (1978:331):

> The validity of the [repetition] technique rests on the assumption that the child, or the second-language learner, when presented with a sentence longer than his immediate memory span, will pass it through a type of filter—his interim

grammar—before repeating it. If a specific syntactic feature
is not part of an individual's grammar, that element will be
distorted during production.

Test-retest reliability

One measure of the reliability of the SRT is the replicability of test results.
A sample of twenty-four Gurung subjects was tested two times, eighteen
months apart.[3] Average scores on the pre- and post-tests were identical, as
shown in Table 4.

This retest shows not only that the same subjects when tested a second
time will score similarly, but that the same test administered by different
researchers will produce reliable results.

Table 4: Replicability of the SRT: an 18-month retest.

	SRT 1	SRT 2
SRT average score (N=24)	28.0	28.0

SRT development

A detailed test development methodology is found in Radloff (1991) and is
only briefly discussed here. The most important part of test development is
the calibration of the SRT with the reported proficiency evaluation (RPE).
In the RPE subjects who have taken a preliminary version of the SRT are
also evaluated by close acquaintances who themselves are educated mother-
tongue speakers of the test language. In the development of the Nepali SRT
approximately 60 second-language speakers of Nepali, from a variety of
mother tongues, took a preliminary form of the Nepali SRT, and were also
evaluated on the RPE. Procedures require that the RPE evaluators be close
enough acquaintances with subjects to have spoken together in the test
language over a period of time and on a wide range of topics. The Nepali
evaluators were carefully guided as they ranked, and then rated, three to
five L2 friends in each of several areas: vocabulary, fluency, accent,
comprehension, and grammar. Each area of proficiency is weighted,
producing a point total for each subject that is then calibrated with the raw
score from the SRT.

Language proficiency criteria used in the RPE evaluation are the original
Foreign Service Institute skill area descriptions that formed the basis for
early versions of FSI oral proficiency tests. Radloff (1991:130) explains
that the advantages to using the RPE technique are that proficiency

evaluations come from the test language community, are based on that community's norms, and rely on the strength of extended personal relationships with that community.

The preliminary form of the SRT had 40-60 sentences. These were reduced to the final form of 15, plus three practice sentences, by choosing those sentences which discriminated subjects' proficiency the best (as measured by the RPE), and which represented a wide range of difficulty.

Control testing

One measure of the validity of the SRT is the control test conducted with mother-tongue speakers of the test language. The control test should demonstrate that mother-tongue subjects perform at the highest levels of proficiency; the RPE levels, after all, are defined in relation to native-speaker proficiency. It should also demonstrate that there is not a significant difference among old and young, educated and uneducated, male and female: all mother-tongue speakers should have comparable proficiency in their first language.

These criteria are very difficult to apply in the case of Nepali because of the diglossic nature of the language. By definition, those who are educated in Nepali are those who control Educated Nepali; those who are uneducated do not control it as well. Age and sex distinctions within an educational category, however, should show little difference in SRT scores on the control test. The degree to which SRT scores are different between educated and uneducated groups is one measure of the degree to which Educated Nepali is different from colloquial Nepali. Table 5 presents the control test results of 54 mother-tongue speakers.

These data show that educated mother-tongue Nepali women performed significantly better (p=.008) than uneducated women. Average scores for women spanned the entire range for RPE level 3+, i.e., though average scores show a sig-nificant difference based on education, they are both within the range of RPE 3+. Data for men are insufficient to determine if education has a significant effect.

All subcategories for the control test averaged RPE 3+, suggesting that the Nepali SRT does not discriminate above level 3+. Put another way it means that for the sake of interpreting SRT results, subjects scoring thirty-four or above on the SRT are indistinguishable from mother-tongue speakers of Nepali.

Table 5: Nepali SRT control test results, showing average score, equivalent RPE level, and sample size.

	Male	Female	Totals
Educated	37.8	39.5	38.5
Uneducated		33.9	33.9
		3+	3+
		22	22
Totals	37.8	35.9	36.6
	3+	3+	3+
	19	35	54

SRT and RPE calibration

As developed by Varenkamp (1994), the Nepali SRT shows a strong positive correlation (r=0.83) with the RPE. The linear relationship between the two variables is quite marked for subjects up to RPE level 3+; the linear relationship for subjects above RPE level 3+ is less clear.

Table 6 presents the calibration scale for converting SRT raw scores to RPE levels, and gives a brief summary description of proficiency levels (Varenkamp 1993), Radloff 1991:242).

Table 6: Nepali SRT raw scores and equivalent RPE levels

SRT raw score range	Equivalent RPE level	Summary description of proficiency level
0-4	0+	Very minimal proficiency
5-8	1	Minimal, limited proficiency
9-14	1+	Limited, basic proficiency
15-20	2	Adequate, basic proficiency
21-27	2+	Good, basic proficiency
28-33	3	Good, general proficiency
34-45	3+ & above	Very good, general proficiency to excellent proficiency

RESULTS

Overview of results

Test sample: The demographic profile of the test sample, given in Table 7, describes the 552 subjects tested from whom complete demographic information was elicited. ...

Table 7.: Demographic profile of overall test sample, showing number and percentage of sample in each sub-category (N=552).

| Sex	Age Group	Education Level Uneducated 0 Years	Primary 1-5 years	Higher 6+ years	Total
	Youth (15-30)	40 (7%)	32 (6%)	68 (12%)	140 (25%)
Male	Middle (31-45)	49 (9%)	18 (3%)	17 (3%)	84 (15%)
	Older (46+)	71 (13%)	3 (1%)	12 (2%)	86 (16%)
Male Total		160 (29%)	53 (10%)	97 (17%)	310 (56%)
	Youth (15-30)	69 (13)	20 (4)	34 (6)	123 (22)
Female	Middle (31-45)	69 (13%)	4 (1%)	1 (0%)	74 (14%)
	Older (46+)	44 (8%)	1 (0%)	0 (0%)	45 (8%)
Female Total		182 (34%)	25 (5%)	35 (6%)	242 (45%)
Overall Sample Total		342 (63%)	78 (15%)	132 (23%)	552 (100%)

That 56% of the sample is male suggests some bias in the sample: men are slightly overrepresented in my sample. Census data from Thakali, Magar, and Limbu areas show two other main areas in which the test sample is biased. Educated men are overrepresented, and uneducated women are underrepresented. This discussion of results focuses on a factor by factor analysis to investigate if the effects of the same factor in different areas are the same or different. For this reason the overall sample in each test area may over- or underrepresent some part of the population as the purpose was not to get the most accurate average but to test enough people

in each sub-category to give a clear picture of proficiency in that sub-category.

Overall totals

A total of 558 subjects were tested on the Nepali SRT. Figure 1 gives the profile of proficiency levels for the total sample, showing decreasing numbers of people at the higher levels of proficiency. Forty-eight percent of subjects scored at RPE levels 2 and below, corresponding with proficiency levels inadequate to satisfy more than basic, routine requirements; only sixteen percent scored at level 3+ and above, comparable to mother-tongue proficiency.

Figure 1: Profile of RPE proficiency levels for entire sample.
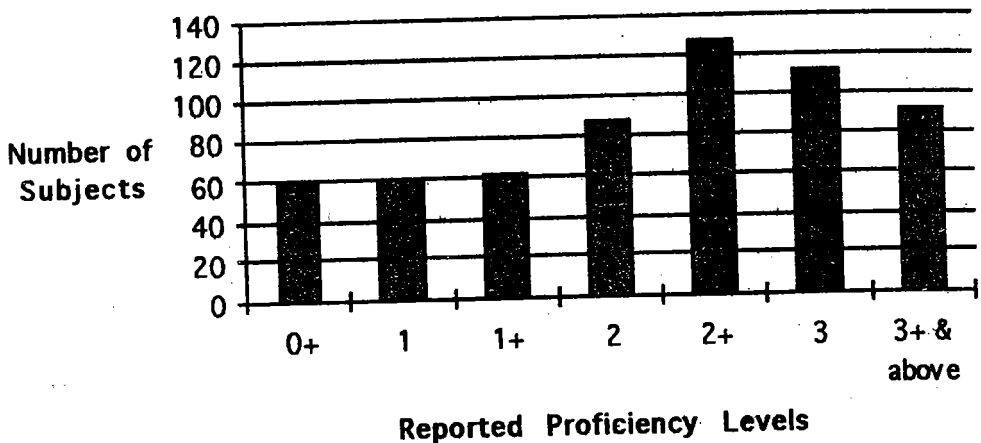
Description of low and high proficiency subjects

Looking in more detail at that part of the population with the lowest ability in Nepali, and that part with the best ability, highlights the great discrepancy between these two groups. Table 8 contrasts the SRT results of 273 subjects in five speech communities. Travel is defined here as having lived more than 6 months outside the mother-tongue area.

Table 8: Comparison of most and least proficient segments of five speech communities (n=273).

	Rpe Levels	
Category	Below Level 2 N = 181	Levels 3+ & above N = 92
Average Age	36 Yerars	26 Years
Sex	56% Female	35% Female
Education	93% Uneducated	12% Uneducated
Travel	86% Untraveled	44% Untraveled

Table 8 shows that those least proficient in Nepali are generally female, older, uneducated, and untraveled. Those who are comparable to the proficiency of mother-tongue Nepali speakers are generally male, younger, educated, and traveled.

Analysis of factors

The selection of factors which may have a significant relationship to Nepali proficiency has been motivated by two things: 1) ease of measuring the factor in future field studies, and 2) the need to account for as much variation as possible in average proficiency scores. Some factors, though not causally related to Nepali proficiency, are easily identified and quantified on a large community scale and are related in some indirect way to L2 proficiency. They are indicators of second-language proficiency.

The first two factors examined here are factors which may have an indirect effect on L2 proficiency. "Research areas" and "village" are analyzed because of the observation that similar demographic subgroups in different areas are significantly different in average SRT score. Since all groups speak Tibeto-Burman languages and are thus linguistically approximately equidistant from the test language Nepali, language should have little effect on the factors of research area and village. First I discuss the variation in average score observed in the different test sites, then I propose a factor to account for some of this variation.

Research area

Proficiency varied widely among the five research areas. Comparing samples of uneducated subjects shows the difference most clearly. Figure 2 shows that average SRT score for uneducated subjects varies from a high of 26 in the Gurung area to a low of 10 in the Gorkha area ($\chi 2$=150.9, p<.0001).

Figure 2: Average SRT score for uneducated samples in five research areas.

Average SRT Score

Research Areas
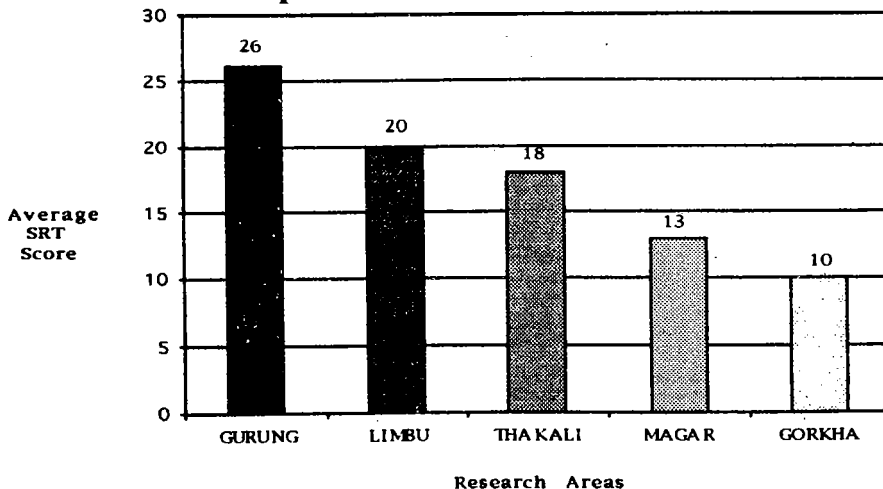
Village

In Table 9 villages are listed by rank order of average score to show the relative strength of a village's effect on subjects' Nepali proficiency.

Table 9: SRT average by village (listed in rank order), for uneducated subjects (χ^2=167.5, p<.0001).

Village	Survey (mother tongue	SRT average for uneducated subjects
Gurung Village	Gurung	26
Hangdewa	Limbu	24
Worden	Limbu	20
Okre	Limbu	19
Marpha	Thakali	19
Chhairo	Thakali	19
Chimang	Thakali	16
Pragatinagar	Magar	16
Philim	Gorkha	12
Arkhala	Magar	11
Lho	Gorkha	10
Bihi	Gorkha	9

Several observations can be made from this rank listing of scores by village. Firstly, villages from the same survey tend to cluster together with little overlap between surveys (e.g., Limbu villages are together towards the top of Table 9, Gorkha villages are together at the bottom). This illustrates, in a different way, what I have already shown in about the significance of research area for Nepali proficiency.

Secondly, there is a wide spread of average scores from village to village. Scores range from a low of nine in Bihi to a high of twenty-six in the Gurung villages. Variation among villages is somewhat more extreme than among research areas, since the latter are simply averages of the villages which comprise them. There is also a significant relationship between research area and average score. It is not surprising, then, that Table 9 shows a similarly significant relationship between village and average score.

Thirdly, even within a particular research area there can be wide variation in average score between villages. Opportunities for speaking Nepali are greater in some villages than in others, depending on access to such things as roads, markets, and ethnic mixing. The two Magar test points, for example, are significantly different (p<.01).

L2 exposure factor

Such differences in average Nepali proficiency among similar demographic subgroups in different language groups and different villages demand a principled explanation. Bilingualism is acquired essentially through contact with L2 speakers: the longer and more intensive the contact, the greater the acquisition. I introduce the concept of the "L2 exposure" factor as an attempt to explain variation in Nepali proficiency among apparently homogenous subgroups in different areas.

The L2 exposure factor contains elements pertaining to entire villages, not just to individuals, and I examined it earlier from a predictive perspective. That is, I try to construct a factor which accurately quantifies the degree of L2 exposure that results from the general character of villages, community-wide factors as opposed to individual ones. The correlation of village L2 exposure scores with individual SRT scores gives one measure of how well this factor helps explain variation unexplained by other factors. Correlation coefficients in Table 10 show a moderate correlation which can account for approximately 25% of the actual variation. This factor is investigated more extensively in a later section.

Table 10: Correlation coefficients for L2 exposure factor and individual SRT score.

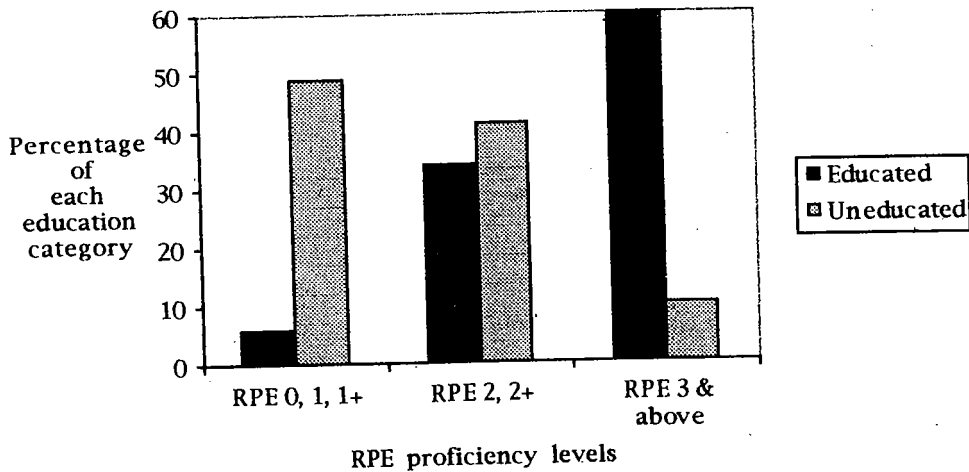| Sample	Correlation Coefficient
Uneducated	0.49
Total sample	0.48

Education

In every community in which I conducted testing, formal public education in Nepali is the factor most highly correlated with acquisition of Nepali, and thus very likely the single most important factor. The more education subjects have, the higher their proficiency in Nepali. Figure 3 summarizes the results by education level for the five speech communities (primary=class 1-5, higher=class 6 and higher).

Figure 3. Graph of average SRT score, by education level, in five speech communities.

Research Areas
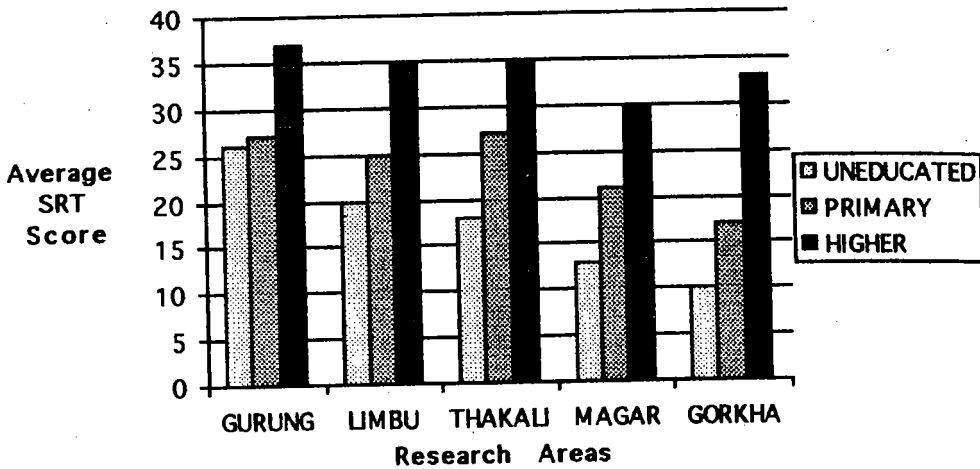
Comparing those who have some education with those who are uneducated reveals a striking difference, further demonstrating the importance of education in Nepali acquisition. Figure 4 compares these two groups for the total sample. The majority of those who have some education test at RPE level 3 or higher; the majority of those who are uneducated are equivalent to RPE levels below 2.

Figure 4. Proficiency level by education level in five research areas (educated, N=210; uneducated, N=344).

RPE proficiency levels

Another way of measuring the effect of education on acquisition of Nepali is by computing the coefficient of correlation between years of education and SRT score. Correlation coefficients in Table 11 show strong correlations between SRT score and years of education for each research

area. The correlation is strongest for Thakali (r=0.78) and the least strong for Gorkha (r=0.62). All correlations can be described as "high correlation; marked relationship" (Fasold 1984:104).

Table 11: Correlation of years of education and SRT score.

Survey	Magar	Gurung	Gorkha	Thakali	Limbu	Total
Correlation coefficients	0.76	0.67	0.62	0.78	0.72	0.70

If, as I have shown, education has such a strong positive correlation with L2 proficiency, how much education is necessary to elevate subjects to a level similar to that of mother-tongue speakers? Comparing average SRT score by grade level completed in school (Figure 5) reveals that those who have completed grade ten or higher score above 36, equivalent to the highest RPE level; those completing grade eight or nine average 33, equivalent to the upper end of RPE level 3, not significantly different from uneducated mother-tongue Nepali speakers. Those completing only grade seven or less on average do not achieve proficiency levels similar to mother-tongue Nepali speakers.

Figure 5. SRT score by grade level completed, for total sample (p<.00001).
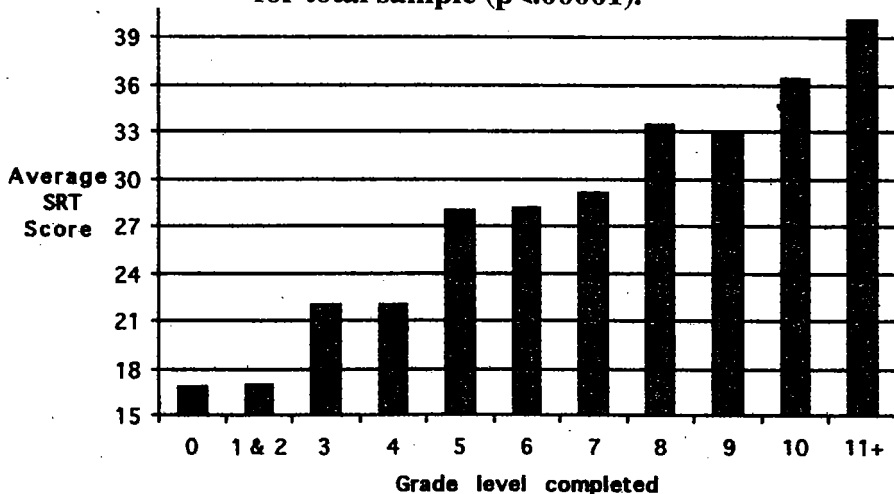
NFE among Gurung of Gorkha District

A sample of 24 Gurung were tested twice on the SRT, about 18 months apart. During the 18-month interval, 17 of the subjects had taken a Nepali non-formal education (NFE) course ranging in length from 6-18 months. Results show two things clearly: 1) the process of becoming literate

through an NFE course does not elevate SRT scores; and 2) results of testing the SRT are replicable—testing a sample of subjects for a second time gives virtually identical results to the first time (see Table 12). Averages for both groups remained virtually constant on pre- and post-tests. NFE participants scored slightly better than non-participants on both the pre- and the post-test.

Table12. Eighteen-month longitudinal study.

	SRT1 average (pre-NFE)	SRT2 average (post-NFE)
NFE participant (N=17)	29.0	28.9
No NFE taken (N=7)	25.7	25.9

Looking at a broader sample of all uneducated subjects, I compared SRT results as a function of NFE experience, regardless of whether they took one or both tests (Table 13). Interestingly, the sample of NFE subjects scores significantly better on the SRT than the non-NFE sample. This says that, in the Gurung area, the kind of person who attends an NFE course is above average in Nepali proficiency irrespective of the NFE course.

Table13. Average SRT score for uneducated sample, by NFE experience.

	SRT Average
NFE Experience (N = 33)	27.6
No NFE Experience (N = 28)	22.9

First language

Almost by definition, the most significant determinant of high proficiency is the subject's mother tongue, defined as the language of the childhood home. Not all members of a minority-language group grow up speaking that minority language. Those raised in homes speaking Nepali should be predictably more proficient in Nepali than those who learn it as a second language.

Results from the Gurung area, shown in Table 14, show this to be true: those raised in Nepali-speaking homes are significantly better in Nepali than those raised in Gurung-speaking or mixed homes (ANOVA p<.01).

**Table14. Average SRT score by first language for
uneducated Gurung subjects.**

First Language	Average SRT / Standard deviation / Sample size
Gurung	24.1 / 5.6 / 40
Nepali-Gurung	23.6 / 1.5 / 5
Nepali	31.4 / 6.7 / 8

Uneducated Gurung raised in Nepali-speaking homes performed similarly to uneducated mother-tongue speakers of Nepali who took the control test as one would expect. Counterintuitively, those from homes where they reported speaking "both" languages did not score significantly better than those reporting only Gurung as the language of the home. This indicates that the mixture of language use in the childhood home does not help Nepali proficiency and may actually hinder it. Data from other areas are needed to confirm this hypothesis.

Towards a predictive model of L2 proficiency
Explaining variation in SRT scores

Thus far I have dealt with one independent variable at a time to determine which of the variables have a significant relationship with SRT score. For those factors which are interval variables, simple regression analysis shows the degree to which an increase or decrease in the independent variable leads to a corresponding change in the dependent variable, the SRT score. Among the variables examined so far, education is by far the most significant factor correlating with SRT. The number of years of education correlates with SRT score at $r=0.70$ ($r^2=0.49$). Thus education accounts for nearly 50% of the variation in SRT score.

In seeking to explain L2 proficiency, or better yet, in predicting it, one must try to account for as much as possible of the total variation in SRT scores. The problem of explaining variation is as follows:

total variation in L2 proficiency =   variation explained by education
                                      + unexplained variation.

Unfortunately, the unexplained variation is 50%, a large amount. In the remainder of this section I attempt to construct a model which will account for this unexplained variation.

Multiple regression analysis, though a statistical tool for interval data, can provide a rough idea of the relative strength of ordinal and nominal variables as well. Multiple regression of education, the L2 exposure factor,

age, sex, and travel reveals that only the first two variables are significant predictors of SRT score. The L2 exposure variable, discussed earlier is both a summary of community-wide variables and a first attempt to quantify variables which affect entire communities of people. Using this version of the L2 exposure factor, and education, the equation for predicting SRT scores is:

$$\text{predicted SRT} = 1.5*\text{education} + .75*\text{L2 exposure} + 11.5$$

Thus, given no education and an L2 exposure score of zero yields a predicted SRT of 11.5. This value is one measure of the amount of residual, unexplained variation in SRT scores. To put it another way, it is the amount of ambient proficiency that exists on average in the five research areas.

A fairly intuitive way of measuring the efficacy of a model is to count the percentage of predicted cases which are within a defined distance of the measured RPE level, say a half-level or a full RPE level (following Simons 1979:122). This would be formulated as follows:

$$\text{ratio of prediction accuracy} = \frac{\text{predictions within a half level on RPE}}{\text{total predictions}}$$

This ratio multiplied by 100 is the percentage of cases predicted within a half-level. A final measure of prediction accuracy which accounts for how far off predictions are is based on the concept of deviations. Correct predictions are a deviation of zero; predictions off by one half-level on the RPE are a deviation of one; predictions off two half-levels (one full level on the RPE) are a deviation of two; and so on. The formula for the percentage of prediction accuracy is thus:

$$\frac{\text{percentage of}}{\text{prediction accuracy}} = \frac{\text{total predictions} - \text{total number of deviations}}{\text{total number of predictions}} \times 100$$

With education as the single predictor of RPE level, 82.4% of predicted scores are within a half-level of the measured scores, 95.9% are within one full level. When this L2 exposure factor is added to the equation, the percentage of predicted cases within a half-level barely rises, to 82.5% (see Table 22).

Improving the model

The initial L2 exposure factor is based on my perception of factors that would contribute to L2 exposure and thus to increased proficiency. To determine which of the single components of this L2 exposure factor correlate significantly with SRT score I did simple regression of each variable with the mean of each village's uneducated sample. To do this I first redefined the individual components of the L2 exposure factor so that they were all interval variables. Table 15 presents the values used for each village and the resulting correlation coefficient with the mean for uneducated samples. The three highest correlation coefficients are underlined. (All components except for ethnic mix and government services are measured in days of travel, with ten hours equal to a full day of travel; ethnic mix is percentage of mixing with other language groups; government services is number of government entities).

These correlation coefficients are quite intuitive: the first six factors all deal with distance from a key source of L2 exposure, and so a negative correlation means that as distance increases and exposure decreases, the measured effect of that exposure (the SRT score) is less and less; for percentage of ethnic mix, L2 exposure and thus proficiency increases in direct proportion to amount of ethnic mixing (the presence of speakers of other languages increases the opportunity for using Nepali).

Table 15. Components of L2 exposure factor, with correlation coefficient of each factor with SRT.

Village	Uned Mean	Rd	Air	Trade	High school	Mid school	Mar ket	Ethnic mix	Gov. serv.
Pragatinagar	16	0.1	1.0	1.0	0.1	0.1	0.1	30	5
Arkhala	11	1.0	1.8	2.0	1.0	0.1	1.0	5	1
Marpha	19	4.0	0.1	0.1	0.1	0.1	0.1	20	3
Chhairo	19	4.0	0.1	0.2	0.2	0.2	0.2	10	2
Chimang	16	4.0	0.2	0.2	0.2	0.2	0.2	10	2
Gurung	26	0.7	1.2	1.0	0.1	0.1	0.2	30	3
Lho	10	5.5	5.5	1.0	3.0	2.0	3.0	1	1
Bhi	9	5.0	5.0	1.0	2.8	2.0	2.5	1	1
Philim	12	4.5	4.0	0.1	2.0	0.1	2.0	5	5
Hangdewa	24	1.0	0.2	0.2	0.2	0.2	0.2	20	3
Okre	19	0.7	1.0	0.5	0.1	0.1	0.2	20	4
Worder	20	1.0	2.0	1.0	0.2	0.2	0.2	20	2
Correlation Coefficient		−.57	−.72	−.34	−.80	−.61	−.79	−.81	.31

Only three components of the L2 exposure factor are greater than 0.75, the arbitrary level I chose. Adding these factors into the model one at a time shows the degree to which they improve upon the simple model that has only education as predictor. Table 16 summarizes the percentage of explained variation in the model as additional factors are added to the model.

Among the five versions of the model there is little difference in the percentage of cases predicted within a half-level RPE, all of the models being close to 85% accurate. All of the models are better than 95% accurate in predicting proficiency to within one RPE level. The right column of Table 16, showing the percentage of prediction accuracy, shows model three to be the best overall model.

Table 16. Predictive ability of different models

Compoents in modeal	Percentage of cases in which predicted RPE is			Percentage of prediction accuracy
	Exactly predicted	within one half-level	within two half-levels	
1. education	28.4%	32.4%	95.9%	6%
2. education + 12 exposure factor	36.9%	82.5%	96.4%	16%
3. education + ethnic mix	44.2%	84.2%	97.5%	25%
4. education + ethnic mix + high school distance	40.8%	85.1%	97.3%	23%
5. education + ethnic mix + high school distance + market distance	40.8%	85.1%	97.3%	23%

The most predictive model of L2 proficiency includes just two factors: years of education and percentage of ethnic mix. The equation for this model is as follows:

Predicted score on Nepali SRT = 1.7 *Years of education + 0.35 *Percentage ethnic mix + 10.9

For subjects with no education living in a perfectly homogenous speech community, this model would predict an SRT score of 10.9. Each year of education increases SRT score by 1.7 points, and each one percent increase

in a community's ethnic heterogeneity increases a person's SRT score by about one-third point.

Theoretically, predicted scores greater than 45 are possible when the ethnic mixture of a community is high enough. Such scores are not actually possible, but what this means practically is that those living in such villages are forced by circumstance to speak Nepali in virtually all domains outside the home, which in most cases will produce proficiency comparable to that of mother-tongue speakers.

I do not think there is anything particularly special about ethnic mixture being the best component of L2 exposure to include in the model as opposed to other components. Different measures of geographic isolation (distance from nearest road, school, or airstrip) will generally correlate quite highly with each other and with the degree of ethnic homogeneity of a village (the more isolated, the more homogenous). Choosing one variable over any other would not have made a substantial difference in the predictive ability of the resulting equation, partly because each adds only a small improvement to the single-factor model.

Summary and conclusions
Summary of results
How proficient in Nepali are the different Tibeto-Burman groups in Nepal? This is the primary question of this study and the one on which language planning issues hinge. To answer this question and others, this study tested a sample of 558 subjects on the Nepali sentence repetition test. These subjects are from seven different ethnolinguistic groups, living in five main areas in the hills of Nepal.

Overview of Nepali proficiency
The proficiency profile below projects the results of this study to the wider population in each area. These projections are based on census data and demographic profiles constructed for each area. They correct for known over- or underrepresentation of subgroups within the samples. This profile, then, is the best prediction of what the overall profile would look like if all the members of each community were tested.

As can be seen by brief inspection of this proficiency profile, there is substantial variation among the five main communities. Gurung in Gorkha District are the most proficient of the five areas, followed by the samples from the Limbu, Thakali, Magar, and Gorkha areas.

Figure 6: Profile of proficiency for all research areas.
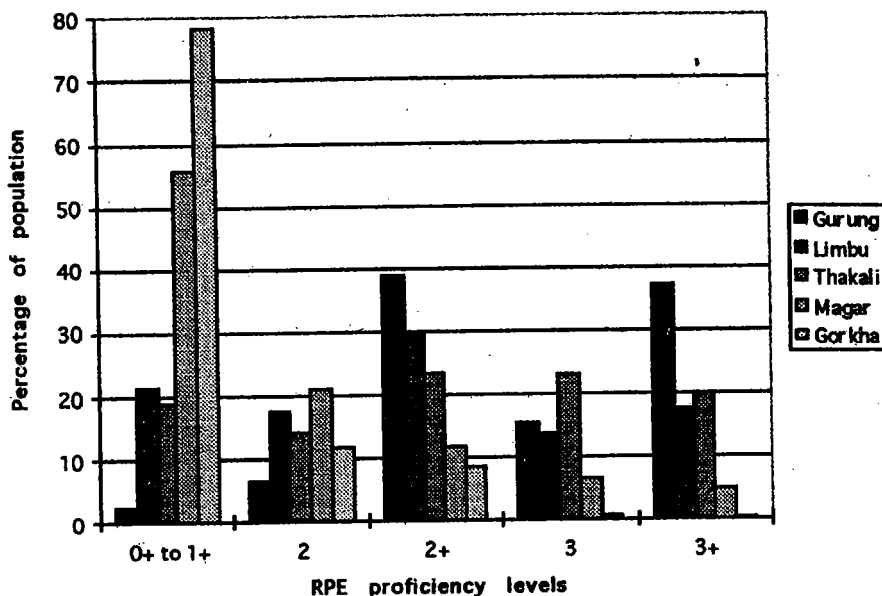
Interpretation of results

The interpretation of these results, and their application to practical language planning issues, depends largely on one's goals and criteria. A population who is RPE level 3+ in Nepali is, at least on the basis of this methodology, indistinguishable from mother-tongue speakers of Nepali. One can argue that because of the diglossic nature of Nepali and the fact that most literature in Nepali has been and continues to be written in "high" Nepali, a population needs to be RPE level 3+ to adequately understand written Nepali. Practical experience from adult non-formal education classes suggests that RPE level 3 is the standard that needs to be reached for successful completion of Books 3 and 4 of the Nepal government's Naya Gareto literacy program. Criteria for other language applications will no doubt require different levels of Nepali proficiency.

I began this paper by defining an ethnic group's "functional need" for mother-tongue education as the "inability, on the part of a significant portion of the language group, to adequately control oral and written Nepal." Given the overall results in this study, as in Figure 6, a strong case can now be made that significant portions of most of the groups in this study do not control Nepali comparably to mother-tongue speakers of Nepali. Perhaps it's time for primary, transitional, mother-tongue education programs to be initiated in Nepal?

Notes

1. This paper is condensed from my M.A. thesis of the same title (Webster 1995), the research for which was carried out under auspices of the Center for Nepal and Asian Studies (CNAS) in 1993.
2. I wish to acknowledge the input of Dr. Clare O'Leary and Bryan Varenkamp concerning this issue.
3. Thanks to Dr. Warren Glover for testing the initial subjects, and to students in the Gorkha NFE program for being willing subjects, twice!

References

Britto, Francis. 1986. Diglossia: a study of the theory with application to Tamil. Washington, D.C.: Georgetown University Press.

Fasold, Ralph. 1984. The sociolinguistics of society. Oxford: Basil Blackwell.

Ferguson, Charles A. 1959. Diglossia. Word 15:325-340.

Fishman, Joshua A. 1967. Bilingualism with and without diglossia; diglossia with and without bilingualism. Journal of social issues, 32:29-38.

Gallimore, Ronald, and Roland G. Tharp. 1981. The interpretation of elicited sentence imitation in a standardized context. Language Learning 31(2):369-392.

Glover, Warren, and Jessie Glover. 1993. Bilingualism testing in Southern Gorkha District. ms.

Grimes, Barbara. 1987. How bilingual is bilingual? Notes on linguistics 40:3-23.

_____. 1992b. Notes on oral proficiency testing (SLOPE). In E. Casad (ed.), Windows on bilingualism, 53-60. Dallas: Summer Institute of Linguistics and University of Texas at Arlington.

Hamayan, Else, Barbara R. Markman, Susanne Pelletier, and G. Richard Tucker. 1978. Differences in performance in elicited imitation between French monolingual and English-speaking bilingual children. International Review of Applied Linguistics 16(4):330-339.

Hutt, Michael J. 1988. Nepali: a language and its literature. London: School of Oriental and African Studies.

Kindell, Gloria E., ed. 1991. Language assessment criteria. In Gloria Kindell (ed.), Proceedings of the international language assessment conference, Horsleys Green, 23-31 May 1989. Dallas: Summer Institute of Linguistics.

Matthews, David. 1992. A course in Nepali, second ed. New Delhi: Heritage Publishers.

Milroy, Leslie. 1987. Observing and analysing natural language. Oxford: Basil Blackwell.

Oller, John W. 1979. Language tests at school—a pragmatic approach. London: Longman Group.

Radloff, Carla F. 1991. Sentence repetition testing for studies of community bilingualism. Dallas: Summer Institute of Linguistics and the Universtiy of Texas at Arlington.

Showalter, Stuart D. 1991. Alternative bilingualism measures for rural African settings. Doctoral dissertation. Georgetown University.

Simons, Gary F. 1979. Language variation and limits to communication. Technical Report No. 3 to the National Science Foundation from the Cornell Department of Modern Languages and Linguistics. Ithaca: Cornell.

Slobin, Dan I., and C. Welsh. 1968. Elicited imitations as a research tool in developmental psycholinguistics. Working paper 10. Berkeley: Language-Behavioral Research Laboratory, University of California, Berkeley.

Varenkamp, Bryan. 1993. "The Nepali sentence repetition test," paper presented at the XIV Annual Conference of the Linguistic Society of Nepal, 25 November 1993.

Varenkamp, Bryan. 1994. The Nepali Sentence Repetition Test: Evaluating Community-wide Bilingualism. Nepalese Linguistics 11:41-43.

Verma, Manindra K. 1992. Nepali. In William Bright (ed.), International encyclopedia of linguistics, volume 3, 76-78. Oxford: Oxford University Press.

Webster, Jeff. 1993a. Indicators of L2 proficiency, revisited. Notes on Scripture in use and language programs 38:37-56.

_____. 1995. Indicators of bilingual proficiency in Nepali among Tibeto-Burman peoples of Nepal. Thesis, University of Texas at Arlington.

Yadav, Ramawatar. 1992. The use of the mother tongue in primary education: the Nepalese context. Contributions to Nepalese Studies 19(2):177-90.