National Development and Language Diversity

Jonathan Pool

Abstract

The joint distribution of domestic language diversity and GDP per capita is highly nonlinear, with almost no high-diversity prosperous countries. Does this support the hypothesis that language diversity impedes economic development?

Data on the size, as a fraction of the population, of the largest native-language group, and the gross domestic product per person per year, were collected for 133 countries. The static relationship between these variables was represented graphically with a scatterplot. The distribution of bivariate values was found to be highly nonlinear, and describable as a dispersion of countries throughout three quadrants of the bivariate space. The quadrant in which almost no countries were located was that with the upper half of GDP and the lower half of largest-language share. In other words, there were almost no highly linguistically diverse, prosperous countries. Although this result is sensitive to the definitions of the variables, some plausible redefinitions would apparently strengthen rather than weaken it. The result appears to support assertions that linguistic diversity is a barrier to economic development, but such a conclusion would be premature, because the distribution of values does not demonstrate an effect, nor, if there is an effect, its directionality.

Publication Information

Jonathan Pool

NATIONAL DEVELOPMENT AND LANGUAGE
DIVERSITY *

I

New, non-Western, and/or developing nations are often said to have
language problems different from,[1] or more frequent or important than,[2]
those of old, Western, and/or developed nations. Reasons given for

[1] Neustupný; Passin; Ferguson (b) 13. But see also Rustow 97.
[2] Passin 453-4; Ornstein. See also Fishman (a) for a distinction between new
and old developing nations' problems.
[3] E.g. Deutsch 2.

these contrasts include the effects of social and political variables on language problems[3] and vice versa.[4]

A much discussed language problem is diversity, which may mean the number of different languages spoken in a given area but may have other meanings which will be mentioned below. Language diversity of one sort or another is held to cause the retardation of development, both political and economic. Language diversity, it is claimed, aggravates political sectionalism;[5] hinders inter-group coöperation,[6] national unity,[7] and regional multinational coöperation;[8] impedes political enculturation,[9] political support for the authorities and the regime,[10] and political participation;[11] and holds down governmental effectiveness[12] and political stability.[13] Similarly it is said that language diversity slows economic development, by, for example, braking occupational mobility,[14] reducing the number of people available for mobilization into the modern sector of the economy,[15] decreasing efficiency,[16] and preventing the diffusion of innovative techniques.[17]

The literature also contains assertions that political or economic underdevelopment, in addition to or instead of being a result of language diversity, is one of its causes.[18] Various kinds of development, it is claimed, give prestige and mobility to certain arenas of life and certain social groups, and thus to the languages prevalent in these arenas and groups. Outsiders then learn these high-status languages, and their spread reduces the level of language diversity (if suitably defined). A society not undergoing much development is thus largely without this cause of the decrease of language diversity.[19] Underdevelopment is also claimed to maintain language diversity by isolating

[4] Grimshaw 197-8.
[5] Fishman (c) 63-4; Sutherlin 66. But for a warning against attributing sectional language conflict to linguistic differences entirely see Pfeffer.
[6] Kloss (a) 75.
[7] Haugen (a) 928; Emerson 133-4; Hertzler 179-81; Deutsch 129-30; Friedrich 559, 572; Richter 10. But see also *ibid.* 12; Deutsch 18-9, 97.
[8] Harries 428.
[9] Fishman (c) 63-4; Verba 532.
[10] Deutsch 4; Machiavelli 10-3.
[11] Stewart (a) 40; Sutherlin 65-6; Valdman 314.
[12] Sutherlin 65.
[13] *Ibid.*; Kloss (b) 8; Rustow 87. But see also *ibid.* 90-1.
[14] Das Gupta & Gumperz 154-6. But see also Deutsch 101ff.
[15] Valdman 314. But see also Deutsch 118.
[16] Fishman (c) 61; Sadler (a) 3-4.
[17] Gumperz 88. Several of these and other arguments can be found in Kelman.
[18] E.g. Hertzler 178-9; Fishman (b) 46-7.
[19] Tauli 20; Kloss (b) 15-6; Stewart (a) 41; Fishman (b) 46; Kloss (a) 77; Wurm 348; Prator 474; Deutsch 158-9.

members of different language groups from communication with each other; when economic development (or any cause) brings them into contact, for example in cities and work places, they tend to learn a common language readily.[20]

The true relation between language diversity and development (if these terms can be defined so as to give rise to a relation) has evident and possible important implications for developmental and linguistic planning.[21] If the assertions outlined above are false, i.e. if there is no relation between development and language diversity, then goals in each domain can be pursued independently. But if language diversity contributes to underdevelopment (or vice versa), then language unification may be a necessary part (or an inevitable result) of successful development planning in a linguistically heterogeneous society. In this case there arises the problem that, in contrast with the notion of 'development', which by connotation if not by definition is accepted as a desirable and plannable goal, language unification is opposed by many;[22] *planned* language unification is further subject to numerous doubts about the empirical practicability, on the one hand, and the morality, on the other, of language planning itself.[23]

II

Regardless of how the planner answers questions about the value or the difficulty of planned language unification, he must estimate the relevance, if any, of language diversity to development. His current ability to make such an estimate, however, is almost nil. Hypotheses about the relations between these two phenomena, such as the assertions summarized in Section I, are not always formulated with enough precision to be tested, and are seldom subjected to thorough testing even when this is in principle possible. The major problems with these hypotheses can be summarized under the headings of concept definition, relation specification, information acquisition, information loss, and causal inference.

Concept definition – The two concepts under examination are 'lan-

[20] Diebold 30; Tocqueville II: 71-2; Stewart (a) 38, 47; Ferguson (a) 6; Togan 39, 47-8, 59, 62-3, 71; Deutsch 41-4, 118-20. But see also *ibid.* 125.
[21] Cf. Neustupný; Kazlas 2-3.
[22] See Haugen (c) 52, 59.
[23] Commonly mentioned problems include the difficulty or impossibility of finding a policy for which enough mass or elite support can be gathered, the intractability of language to manipulation, the scarcity of needed resources and expertise, conflicting goals, undesirable side-effects, and moral objections to the use of coercion to change language.

guage diversity' and 'development'. To enter into hypotheses these (as any) concepts must be defined such that their presence or absence, rank order, amount of change, or absolute value can be established by observation. To enter into *successful* hypotheses, the concepts (as defined) must denote properties which are in fact associated with each other. Different scholars have proposed that language diversity be defined in terms of the number of languages (varieties, mutually unintelligible varieties, dialects, etc.) spoken in a given area (by more than $x\%$ of the population),[24] in terms of the percentage of the population not speaking the most widely spoken language (natively, in the home, as a second language, etc.),[25] in terms of the official (regional, educational, etc.) language(s) (number of them, percentage of the population speaking none of them, etc.),[26] in terms of the distances (linguistic, attitudinal, ideological, etc.), between one language or variety and another,[27] or as a function of several of these or other variables.[28] Development has similarly been defined economically, politically, or with other emphases; in terms of gross output (gross national product [GNP], per capita GNP, political capability, channel capacity, etc.)[29] or its pattern of allocation (income distribution, differentiation, secularization, participation, etc.);[30] and as a state (level of attainment), a rate (of change), or a change of rate.

Relation specification – If the two properties, once defined, are found to be statistically associated, we do not thereby know that one causes the other. The relation between the two (let us call them A and B) can be genuine (A affects B or B affects A), but may also be spurious (S affects A and B, i.e. A and B have one or more common causes). If genuine, the relation can be in one, the other, or both of the two directions,[31] and can have different chains of intervening variables (e.g. A affects I, I affects J, J affects B). In the case at hand, hypotheses have been proposed that link development with language unification through intervening variables of communication, education, cleavage, etc.[32] Spurious relations have also been asserted, with the common cause(s) being forms of nationalism, democratization, revolution, independence,

[24] Fishman (c) 55; Ferguson (b) 11; Richter 5-6.
[25] E.g. Russett *et al.* 132-7; Fishman (c) 67-8.
[26] Stewart (b) 20-1; Passin 449-50; Das Gupta & Gumperz 155-6.
[27] Stewart (b) 22; Haugen (c) 55; Fishman (b) 44-5.
[28] Fishman (c) 67-8; Ferguson (a) 1-2; Kloss (a) 72-7; Rustow 97-2, 102. See Greenberg, also Sadler (b), for a variety of definitions.
[29] Almond & Powell 190-212; Easton 119-27; Eisenstadt 43.
[30] *Ibid.*; Lerner 50-1.
[31] Neustupný 288-9.
[32] Das Gupta & Gumperz 152-3; Stewart (b) 15; Heyd 14; Coleman 36-7; McDavid 17; Sutherlin 66.

education, social mobilization, etc.[33] In addition to telling us the kind of relation it is asserting, a hypothesis should specify the boundary conditions within which the relation is claimed to hold (age of nation, degree of coercion employed in assimilation, etc.).[34]

Information acquisition – Choosing definitions and specifying relations are often impossible or useless, however, because of the paucity of existing information. While there are certain variables in the realm of economic development for which data allow comparisons across nations and across time, the political data are less complete and the linguistic ones still less so. Many countries' censuses do not ask questions on language, and those that do so ask different questions from each other, almost always omit such obviously important information as second-language knowledge, and sometimes change their definitions from one census to the next. Even if the questions are well designed, the responses may fail to reflect real language use or knowledges.[35]

Information loss – While some students of sociolinguistic problems admirably employ demographers' techniques to extract more information from censuses than they appear on the surface to contain,[36] the more common pattern is to waste what data we do have by failing to examine them until they have been compressed into a single index, most often a correlation coefficient, or until the entire range of variation along each variable has been reduced to a few categories or even a dichotomy.[37]

Causal inference – In view of the aforementioned problems, statements of causal relations between language diversity and national development run the risk of being meaningless (concepts not defined), unsatisfying (relations and conditions not specified), or unsupported (data not adequate). Often however, it is recognized that causal hypotheses can not be proposed except as tentative guesses lacking convincing confirmation. Attention is then devoted to careful factual description and generalization which will hopefully provide the data with which hypotheses will be generated and tested in the future.

III

The frontier of current speculation about the relations between language diversity and national development is probably to be found in

[33] Haugen (a) 928-9; Haugen (c) 63; Das Gupta 17-8; Bowers 396; Tauli 122; Haugen (b) 115; Passin 453.
[34] Fishman (a); Passin 451; Easton 249-50; Deutsch 118-20.
[35] Lieberson (a) 139-40; Bose.
[36] *Ibid.* 136-8, 144-50.
[37] E.g. Kloss (a) 81; Banks & Textor.

the work of Joshua A. Fishman. In a recent article[38] he has compiled descriptive generalizations and offered tentative hypotheses about these relations, based on two catalogs of national-level aggregate and global indicators, the *World Handbook of Political and Social Indicators*[39] and *A Cross-Polity Survey*.[40]

Fishman gets "the decided impression that linguistic homogeneity is currently related to many more of the 'good' and 'desirable' characteristics of polities than is linguistic heterogeneity. Linguistically homogeneous polities are usually economically more developed, educationally more advanced, politically more modernized, and ideologically-politically more tranquil and stable."[41]

He also notes that "many of the reported differences between linguistically homogeneous and heterogeneous polities also appear to be differences between rich and poor polities . . ." (and he controls for one and then the other of these kinds of variables to see which accounts for more of the other variations associated with both).[42] Concerning the causal relation between language diversity and development, he says that the usual explanation gives developmental processes as causes of increased linguistic (and other) homogenization, but that language diversity may also hinder (while language unity helps) development.[43]

Fishman rightly criticizes the definitions underlying many of the (especially linguistic) data available for the two source volumes, regrets the incompleteness of the data even where definitions are good, challenges the use of dichotomization in the *Survey* and of correlation coefficients in the *Handbook*, and suggests remedies for these defects. He accordingly remains tentative in his inference of causal relations between language diversity and development, confining himself mainly to description. Yet what descriptions and inferences he does make are unfortunately based largely on the less reliable and less salvageable data in the *Cross-Polity Survey*.[44] These data are presented in irreversibly categorized (grouped) form, while *Handbook* data are presented as

[38] Fishman (c).
[39] Russett *et al.*
[40] Banks & Textor.
[41] Fishman (c) 60.
[42] *Ibid.* 61-4.
[43] *Ibid.* 60-1. The authors of the *World Handbook*, taking the opposite view of which direction of causation is more obvious, say, "Very possibly this is in part a causal relationship – countries of diverse linguistic composition face a special hurdle in development – but the relationship between linguistic diversity and development is so complex, including the power of economic development to force assimilation to the dominant (or even sometimes a minority) language, that the question demands further inquiry", Russett *et al.* 290.
[44] Perhaps because the *Handbook* was not yet published in final form: Fishman (c) 54.

ratio scales, i.e. with a particular value for each country on each variable. The *Handbook* makes considerable use of rate-of-change data, and even its static figures would be useful in creating time-series files, while neither of these statements can be made about the *Survey*.

IV

Thus, although Fishman's complaints, which are indeed just, mostly require the generation or collection of new data for their redress, one improvement that can already be made is to rescue existing data of the *Handbook* type from information loss. To see how this might be done, let us consider two of the most widely measured language-diversity and development variables: the size of the largest native-language community in a country as a proportion of the population, on the one hand, and the gross domestic product (GDP) per capita, on the other. Rather than aiming directly at a hypothesis that would link these two variables causally, let us first seek only to describe their joint distribution, but to do so in the way most useful for later formulation of hypotheses. We are resigned for now to the fact that varying census questions, artificial exchange rates, and missing hard data make figures for both variables quite low in reliability. What we wish to do is to correct information loss that occurs *beyond* the point of initial data collection.

In the *Handbook* this loss takes two forms: the omission of units (countries) and the summarization of the association between the variables. In the first type of loss, the number of countries is limited to 58, less than half of the universe, because hard data on one or another of the two variables in question are unavailable for the remaining countries.[45] Yet the inclusion of the best available estimates for the latter group of countries would probably not increase the error margin above where it has already been raised by the conflicting definitions on which even the 'hard' data are based.

The second type of information loss occurs when the joint distribution of the two variables is described,[46] not in detail, but in the summary form of a correlation coefficient (0.47). An infinite number of different distributions could have given rise to this same index. Indeed, it gives even less predictive power than generalizations arising from the *Survey*.[47] For even if we treat each variable as having only two values,

[45] Russett *et al.* 133.
[46] *Ibid.* 275, 277, 290.
[47] "Banks and Textor report that linguistically homogeneous polities tend to have at least a 'medium' per capita gross national product (at least 300 U.S. dollars per year) .... Linguistically heterogeneous polities ... tend to have 'low' or 'very low' per capita gross national product ...." Fishman (c) 56.
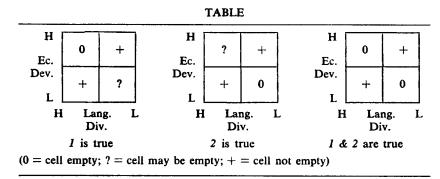
'high' and 'low', there are still four extreme generalizations any one or more of which could be true compatibly with the given correlation:

1a If high language diversity, then low economic development;
1b If high economic development, then low language diversity;
2a If low language diversity, then high economic development;
2b If low economic development, then high language diversity.

It can easily be shown that Generalizations 1a and 1b imply each other, and likewise 2a and 2b. But 1 (a or b) allows low development with low diversity and excludes high development with high diversity, while Generalization 2 (a or b) allows a high-high and excludes a low-low combination. Thus to the extent that we are viewing both variables as dichotomies, we have two pure-type generalizations, *1* and *2*. If the first is true, only high development with high diversity is excluded; if the second is true, only low development with low diversity is excluded;

TABLE

	H	Lang. Div. L			H	Lang. Div. L			H	Lang. Div. L
H Ec. Dev. L	0	+		H Ec. Dev. L	?	+		H Ec. Dev. L	0	+
	+	?			+	0			+	0
	1 is true				*2* is true				*1 & 2* are true	

(0 = cell empty; ? = cell may be empty; + = cell not empty)

and if both are true, both these combinations are excluded (see Table). A choice among these three alternatives would be only a first approximation, of course.

Thus our 'information retrieval' should begin with the addition of data omitted from the *Handbook*, continue by discovering which of our three generalization-alternatives best describes the relationship between the two variables, and conclude with a more precise description of the relationship than any of the three alternatives provides. The first operation, the addition of data, has been facilitated by the publication of a recent article of Dankwart A. Rustow, whose language figures,[48] while questionable in some cases, I have accepted except where more recent

[48] Rustow 94-6. Most of these figures are based on Bruk, and many of the latter's statistics are estimates (because of the absence of census or sample data). Furthermore, Bruk's nationality figures are not all based on native language or second language.

or apparently more reliable ones are available.[49] Data on gross domestic product per capita are less of a problem; they are taken from Deutsch (supplemented in a few cases by other sources)[50] so that they cover approximately the time when the language information was collected
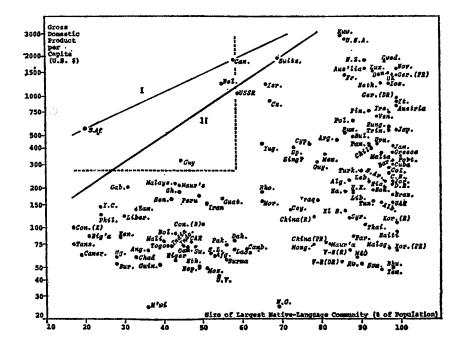

Figure. Gross Domestic Product per Capita (US $) and Size of Largest Language Community (% of Population), c. 1962.

[49] Sources for other language figures: Bel. – Tabouret-Keller 110; Bol. – *Tipologia* 100-1; Burma – Kunstadter 87-9; Bul., Ger. (DR), Ger. (FR), Nor., & Swed. – Tokarev & Čeboksarov I: 34; Camb. – Kunstadter 867; Cey. – Guseva *et al.* 28-30; China (PR) – Roberts 112; Guy. – Nath. 221-2, Smith 6 & 8; Iran – *U.S. Army* 86; Kor. (PR) & Kor. (R) – Čeboksarov *et al.* I: 35; Laos – Kunstadter 255-7; Malays. – Noss 143-4; Maur's – Russett *et al.* 136; Peru – *Tipologia* 109; S. Af. – South Africa (b) and Van den Berghe 222; Thai. – Noss 201; V-N (DR) – Kunstadter 693-4; V-N (R) – *ibid.* 696-700. See Noss for different figures on Burma; Camb., & V-N (R).

[50] The *Handbook* (like the *Survey*) uses gross *national* product, which differs slightly from gross domestic product: see Deutsch 261. GDP/capita values (for 1962) are from Deutsch 262-70, except where missing, i.e.: Kor. (PR) & V-N (DR) – Russett *et al.* 157 (GNP/capita, 1957); Kuw., M'wi, Malays., Mong., N.G., Rho., Tanz., Trin., & Zamb. – Ernst 192-3 (per capita national income, 1962-4).

Gross Domestic Product per Capita (U.S. $)

3000 — Kuw.
• U.S.A.

2000 —
N.Z. •   Swed.
Aus'lia •   Lux. •   • Nor.
Fr. •   Den. •• Ger.(FR)
1500 —   • Can.   • Switz.
Neth. • UK • Ice.

• Bel.
• USSR   • Isr.   Ger.(DR)
1000 —   • It.
Cz. •   Ire.   Austria
I   Fin. • • Ven.
750 —   Pol. •
II   Hung.
Rum. • Trin. • • Jap.
• Bul.
Arg. • Pan. Uru.
500 — • S.Af.   Cyp.   • Jam.
Yug. •   Sp. •   Chile   Greece
400 —   Sing•   Mex. •   Malta Port.
Bar Cuba
• Guy.   Turk. • S.Ar. Col.
300 —   C.R.
Alg. • Leb.   Nic Jor.
Ec. • H.K. Hon. D.R.
Malays.   Rho. •   Lib. • Braz.
Gab. •   • Maur's   Tun. Alb. UAR
200 —   Gh. •   Mor. •   Iraq •
Sen. • Peru   Guat.   Cey. •   Syr.
I.C. • Zam. •   Iran   El S. •   Kor.(R)
150 — Phil. • Liber.   China(R) •
Thai. •

Con.(B) •
100 — • Con.(K)   Bol. •
• Nig'a Ken. • India   CAR   Pak. Dah.   China(PR)   Par.   Haiti
Mali Indo Pak. Laos Camb. Mong. Maur'a Malag. Kor.(PR)
75 — • Tanz. Ang. • Togo Gam. Su. S.L. V-N(R) M&O
Camer. Ug. • Chad Niger Afg. Burma V-N(DR) Rw. Som. Bhu.
Bur. Guin. Eth. Moz.   Yem.
50 — Nep.
U.V.
40 —

30 — M'wi   N.G.

20 —
Size of Largest Native-Language Community (% of Population)

10   20   30   40   50   60   70   80   90   100

(early 1960's). Using these sources, it has been possible to expand the population from 58 to 133 countries.

For the second operation, the data thus acquired are presented together in the form of a scatter plot, showing the joint distribution of the two variables without any information loss (see Figure).[51] It is clear from a glance at the Figure (cf. the Table) that the first alternative (the truth of Generalization 1) is the best approximation of the three to the true relationship between GDP per capita and proportional size of the largest native-language community. If we dichotomize each variable at the midpoint of the range of variation actually exhibited, the upper left cell (bounded by dotted lines) is nearly empty.

Having gone beyond a single index to three patterns possibly responsible for it and having discovered which one of them in fact best fits the (expanded) data, let us now go to the third step and see what information would allow us to say more about the joint distribution than we have done with this first approximation. There are three ways in which we could proceed.

First, we could describe the distribution more precisely. We can say, for example, that all the countries are excluded from a triangle (bounded by Line I) in the upper left 12% of the rectangle which is created by the ranges of variation of the two variables. The complete emptiness of this corner means, in a descriptive sense, that a country can have any degree of language uniformity or fragmentation and still be underdeveloped; and a country whose entire population, more or less, speaks the same language can be anywhere from very rich to very poor. But a country that is linguistically highly heterogeneous is always underdeveloped or semideveloped, and a country that is highly developed always has considerable language uniformity. Language uniformity, then, is a necessary but not sufficient condition of economic development, and economic development is a sufficient but not necessary condition of language uniformity (speaking descriptively, not causally). Going beyond this generalization, we can determine the actual rate of occurrence of particular combinations of economic development and

[51] According to Rustow's criteria, every independent country with more than 100,000 people and every dependency with more than a million people is included in the Figure if data are available (N = 133). Variances are approximately standardized in the Figure by subjecting the economic variable to a logarithmic transformation, such that each unit of vertical distance represents an equal proportionate, rather than absolute, change. The *Handbook* presents scatter plots for several other pairs of variables: Russett *et al.* 304-10, 327-31. Cf. also graphic methods for describing a similar kind of relation at the individual (intranational) level, and reasons given, in Deutsch 137, 139-42.

language diversity, by defining Line I and measuring the frequencies in various parts of the space under it.[51]

Secondly, we could examine the effects of changing the definition of one or another of the two variables, along lines suggested in Section II. As one example, it seems that a somewhat stronger association would hold if language diversity were defined as 'percentage of population *able* to speak the most widely spoken language' instead of the size of the largest community of *native* speakers. In this case the excluded triangle would expand at least to Line II,[53] since under this definition of language diversity South Africa would move over to 42%, Canada to 80%, and Belgium to 60%.[54] The excluded triangle would probably be larger still, but the necessary data on Switzerland are unavailable.[55] In fact, so few countries collect statistics on second-language knowledge that it is at present impossible to employ that definition in multinational comparative work.

Thirdly, we could define the distribution more complexly, recognizing that the relation between the two variables under consideration is open to influences of outside variables. We could attempt to locate the most important boundary conditions and specify their effects. The use of a scatter plot makes it possible to compare countries that are the same on one variable but different on the other (e.g. Rwanda and Burundi, Korea and Japan), and to look for culturally, regionally, or otherwise distributed variables that might explain deviant cases and increase the present low predictability of one variable from the other.[56] In view of the limits placed on the first approach by the unreliability of the data, and on the second by the scarcity of more refined data, this third tack may well be the most profitable one for now.

It is important to remember that what has been begun here is only a

[52] It would be possible (and popular) to standardize both variables precisely (cf. note 51), define and apply a measure of deviation from random association (i.e. non-association), and calculate the statistical significance of the observed distribution's deviation. But the imprecision of our raw data, the irrelevance of the degree of deviation from randomness to our aim of predicting a value on one variable from a value on the other, and the doubtful meaningfulness of statistical significance when applied to an entire population (see e.g. Morrison & Henkel) call into question the usefulness of these operations here.

[53] And only one country (South Africa) would remain in the upper-left rectangular cell.

[54] South Africa (a); Russett *et al.* 135. The ability datum for South Africa is from 1951, that for Belgium from 1947.

[55] On the absence of these data for Switzerland, see Meli 19; for a good analysis of what data do exist, see McRae.

[56] Within most geographical regions, however, the relationship is substantially weaker. See Kazlas.

limited form of information recovery. Its immediate result has been no more than one or two descriptive generalizations more precise than those offered by the *Handbook* or by Fishman, and the preceding suggestions for further work can at best lead to still more improved descriptions. These, however, are *not* convertible into advice for a development-minded language planner or a language-minded development planner. For our statements are descriptive and static: they describe what is, rather than predicting what would be under other conditions, and they deal with states rather than rates. The planner needs predictive, dynamic hypotheses: good guesses about how a country's value on one variable would change if he changed its value on the other. Knowledge of this kind does not and can not follow logically from static description.

V

In spite of this warning, it may be objected that the Figure does indeed suggest something about the role of language in development, something that has important implications for language policy. Specifically, the relationship pictured in the Figure brings to mind a statement of Deutsch in *Nationalism and Social Communication*:[57]

Assimilation in language or culture involves the learning of many new habits, and the unlearning of many old ones – habits, in both cases, which often interlock and reinforce each other. Such learning as a rule is slow; its changes are counted in decades and generations.

The growth of an economy or a technology, on the other hand, may be much more rapid; transportation systems and markets can grow very quickly; workers or immigrants may be recruited and imported within a few years, or sometimes even months. Much of this economic or technological development may force people into new and inescapable contacts with each other as workers, customers, and neighbours – contacts far narrower, perhaps, than the range of human relations that can be communicated within one culture; but contacts far wider than the relations which can be communicated in the absence of a common culture to outsiders.Linguistically and culturally, then, members of each group are outsiders for the other. Yet technological and economic processes are forcing them together, into acute recognition of their differences and their common, mutual experience of strangeness, and more conspicuous differentiation and conflict may result.

An empirically similar, though normatively different, statement is made

by Fishman,[58] whose analysis of the separate effects of economic and linguistic variables on other characteristics

strongly suggests that the simultaneous pursuit of the advantages of higher economic status coupled with the protection of maintenance of values cultural-linguistic differences is not a wil-o'-the-wisp.

What Deutsch treats as a real danger and Fishman sees as a real promise is that a country might move from the lower left corner to the upper left corner of our Figure and remain there. The relation which we have observed might seem to show that this is impossible, and that a planner who insists on preserving cultural-linguistic pluralism had better be ready to sacrifice economic progress. The Figure may appear to demonstrate that development either requires or brings about second-language learning (Line II) followed closely by native-language change (Line I), so that no country is ever caught at any one time in the upper left corner. Thus the Figure may be viewed as confirming Greenberg's expectation[59] that

the increase of communication that goes with greather economic productivity and more extensive political organization will lead typically to the spread of a lingua franca, whether indigenous or imported, resulting in widespread bilingualism and the ultimate disappearance of all except a single dominant language.

Other possible mechanisms of this effect, universal education and the growth of inequalities in prestige among languages, have been mentioned earlier in this paper. Whatever the intervening variable(s) might be, this relation, which denies even the notion of a stable auxiliary national language,[60] would if true challenge both Deutsch's fear of, and Fishman's hope for, development with diversity.

Our data are not, however, sufficient to demonstrate the truth of this relation, as has been explained at the end of Section IV. Although it is a common practice to infer causal relations from cross-sectional comparative data such as ours, the inference is supported only when there is additional information, or when we make assumptions on which the inference can depend. In this case, we must know or assume that some of the countries currently in the upper right corner were once in the lower left corner.[61] In this event we could say that the Deutsch and Fishman phenomenon has had a change to happen but – in its extreme form of economic development with no homogenization at all – has

---

[57] Deutsch 125-6.

[58] Fishman (c) 64.
[59] Greenberg 110. Cf. the concern about the rapid change of native language of immigrant Canadians: see Canada I: 22-7.
[60] The implications of the statements cited here for the question of an auxiliary international language may also merit consideration.
[61] This is doubted by Fishman (c) 61.

never happened. To conclude that a milder version of the effect, such as the *slow* homogenization described by Deutsch, has never happened, we would further need a series of snaphots of the changing scatter plot at suitable time intervals. Even then, however, the fact that something has never happened does not necessarily support the claim that it will not happen in the future. Any characteristic that distinguishes those countries now in the lower left corner from those that used to be there ᴜ.ay be made a boundary condition in a hypothesis predicting economic development without the need for linguistic unity; then history, while unable to confirm the hypothesis, will be equally unable to provide an exception.[62]

The Figure alone thus does not refute Deutsch and Fishman. But it also offers no support for their view. We have seen, in fact, that at the extremities of linguistic diversity there is at present not a single country able to serve as a model (or living proof of the danger) of economic-development-*sans*-assimilation-in-language. This fact should make us at least skeptical enough of claims for development with diversity that we ask to see the evidence in favor of those claims. Even if the evidence confirms the possibility of this combination, the obverse of what was said above applies: the fact that something has once happened (and is thus possible) does not necessarily mean that it *will* happen in the future. Indeed the absence of a contemporary model may itself discourage planners and politicians from attempting to bring about the combination in question, and thus perpetuate its absence.

REFERENCES

Almond, Gabriel A., and G. Bingham Powell, Jr.
1966   *Comparative Politics: A Developmental Approach* (Boston: Little, Brown).
Banks, Arthur S., and Robert B. Textor
1963   *A Cross-Polity Survey* (Cambridge: MIT).
Bose, Ashish
1969   "Some aspects of the linguistic demography of India", in *Language and Society in India: Proceedings of a Seminar*. Indian Institute of Advanced Study, *Transactions*, Vol. 8 (Simla), pp. 37-57.
Bowers, John
1968   "Language problems and literacy", in Fishman *et al.*, 381-401.
Bruk, S. I. (ed.)
1962   *Čislennost' i rasselenie narodov mira* (Moscow: Akademija Nauk SSSR).

[62] The extent to which hypotheses require no boundary conditions of this kind in order to be confirmed may be a measure of the utility of studying the early development of the developed countries to make predictions about currently developing ones.

Canada
1967   Royal Commission on Bilingualism and Biculturalism. *Report*. Vol. I (Ottawa: Queen's Printer).
Čeboksarov, N. N., S. I. Bruk, R. F. Its and G. G. Stratanovič (eds.)
1965   *Narody Vostočnoj Azii* (Leningrad: Nauka).
Coleman, James S. (ed.)
1965   *Education and Political Development* (Princeton: Princeton).
Das Gupta, Jyotirindra
1968   "Language diversity and national development", in Fishman *et al.*, 17-26.
Das Gupta, Jyotirindra and John J. Gumperz
1968   "Language, communication and control in North India", in Fishman *et al.*, 151-66.
Deutsch, Karl W.
1966   *Nationalism and Social Communication: An Inquiry into the Foundations of Nationality*, 2nd ed. (Cambridge: MIT).
Diebold, A. Richard, Jr.
1962   "Mexican and Guatemalan bilingualism", in Rice, 26-33.
Easton, David
1965   *A Systems Analysis of Political Life* (New York: Wiley).
Eisenstadt, S. N.
1966   *Modernization: Protest and Change* (Englewood Cliffs, N.J.: Prentice-Hall).
Emerson, Rupert
1962   *From Empire to Nation: The Rise to Self-Assertion of Asian and African Peoples* (Boston: Beacon).
Ernst, Morris L.
1967   *The Comparative International Almanac* (New York: Macmillan).
Ferguson, Charles A.
1962a   "Background to second language problems", in Rice, 1-7.
1962b   "The language factor in national development", in Rice, 8-14.
Fishman, Joshua A.
1968a   "Language problems and types of political and sociocultural integration: A conceptual postscript", in Fishman *et al.*, 491-98.
1968b   "Nationality-nationalism and nation-nationism", in Fishman *et al.*, 39-51.
1968c   "Some contrasts between linguistically homogeneous and linguistically heterogeneous polities", in Fishman *et al.*, 53-68.
Fishman, Joshua A., Charles A. Ferguson and Jyotirindra Das Gupta (eds.)
1968   *Language Problems of Developing Nations* (New York: Wiley).
Friedrich, Carl Joachim
1963   *Man and His Government: An Empirical Theory of Politics* (New York: McGraw Hill).
Greenberg, J. H.
1956   "The measurement of linguistic diversity", *Language*, 32: 109-15 (No. 1).
Grimshaw, Allen D.
1966   "Directions for research in sociolinguistics: Suggestions of a non-linguist sociologist", in Lieberson (b), 191-204.
Gumperz, John J.
1962   "Language problems in the rural development of North India", in Rice, 79-90.

Guseva, N. R., A. M. D'jakov, M. T. Levin and N. N. Čeboksarov (eds.)
1963   *Narody Južnoj Azii* (Moscow: Akademija Nauk SSSR).
Harries, Lyndon
1968   "Swahili in modern East Africa", in Fishman *et al.*, 415-29.
Haugen, Einar
1966a   "Dialect, language, nation", *American Anthropologist*, 68: 922-35 (No. 4, Aug.).
1966b   *Language Conflict and Language Planning: The Case of Modern Norwegian* (Cambridge: Harvard).
1966c   "Linguistics and language planning", in William Bright (ed.), *Sociolinguistics: Proceedings of the UCLA Sociolinguistics Conference, 1964* (The Hague: Mouton).
Hertzler, Joyce O.
1966   "Social uniformation and language", in Lieberson (b), 170-84.
Heyd, Uriel
1954   *Language Reform in Modern Turkey*. Oriental notes and studies, no. 5 (Jerusalem: Israel Oriental Society).
Kazlas, Juozas A.
1970   "A study of linguistic diversity", unpublished ms.
Kelman, Herbert C.
1969   "Language as aid and barrier to involvement in the national system". Paper prepared for the Conference on Language Planning Processes, East-West Center, Honolulu. Reprinted in this volume, 185-212.
Kloss, Heinz
1966   "Types of multilingual communities: A discussion of ten variables", in Lieberson (b), 7-17.
1968   "Notes concerning a language-nation typology", in Fishman *et al.*, 69-85.
Kunstadter, Peter (ed.)
1967   *Southeast Asian Tribes, Minorities, and Nations*, 2 vols. (Princeton: Princeton).
Lerner, Daniel
1964   *The Passing of Traditional Society: Modernizing the Middle East* (New York: Free Press).
Lieberson, Stanley
1966a   "Language questions in censuses", in Lieberson (b), 134-51.
1966b   (Ed), *Explorations in Sociolinguistics* (Bloomington, Indiana & The Hague: Mouton).
Machiavelli, Niccolo
1964   *Machiavelli's The Prince*. Edn. and trans. of *Il Principe*, by Mark Musa (New York: St. Martin's).
McDavid, Raven I., Jr.
1969   "The sociology of language", unpublished ms.
McRae, Kenneth D.
1964   *Switzerland: Example of Cultural Coexistence*. Contemporary affairs, no. 33 (Toronto: Canadian Institute of International Affairs).
Meli, M. A.
1963   "Des langues parlees en Suisse", in *Le bilinguisme en Suisse, en Belgique et au Canada*, Etudes et documents, no. 5 (Brussels: Foundation Charles Plisnier).
Morrison, Denton E. and Ramon E. Henkel

1969   "Significance tests reconsidered". *The American Sociologist*, 4: 131-40 (No. 2, May).
Nath, Dwarka
1950   *A History of Indians in British Guiana* (London: Nelson).
Neustupny, Jiri V.
1968   "Some general aspects of 'language' problems and 'language' policy in developing societies", in Fishman *et al.*, 285-94.
Noss, Richard
1967   *Language Policy and Higher Education*. Higher education and development in South-East Asia, vol. III, part 2 (Paris: UNESCO & IAU).
Ornstein, Jacob
1964   "Patterns of language planning in the new states", *World Politics*, xvii: 40-49 (No. 1, Oct.).
Passin, Herbert
1968   "Writer and journalist in the transitional society", in Fishman *et al.*, 443-57.
Pfeffer, Karl Heinz
1969   "Sprachenfrage und soziale Unruhe in Pakistan", *La monda lingvo-problemo*, 1: 129-39 (No. 3).
Prator, Clifford H.
1968   "The British heresy in TESL", in Fishman *et al.*, 459-76.
Rice, Frank A. (ed.)
1962   *Study of the Role of Second Languages in Asia, Africa, and Latin America* (Washington: Center for Applied Linguistics).
Richter, William L.
1968   *The Politics of Language in India*. Unpub. Ph.D. dissertation, Univ. of Chicago.
Roberts, Janet
1962   "Sociocultural change and communication problems", in Rice, 105-23.
Russett, Bruce M., Hayward R. Alker, Jr., Karl W. Deutsch and Harold D. Lasswell
1964   *World Handbook of Political and Social Indicators* (New Haven: Yale).
Rustow, Dankwart A.
1968   "Language, modernization, and nationhood – An attempt at typology", in Fishman *et al.*, 87-105.
Sadler, Victor
1962   "A measure of 'internationality' ", *International Associations* (No. 7), 480-81.
1969   Editorial. *La Monda linguo-problemo*, I: 1-4 (Jan.).
Smith, Raymond T.
1962   *British Guiana* (London).
South Africa
1951   Office of Census and Statistics. *Population Census: 8th May*, Vol. 6, "Language and literacy".
1960   Bureau of Statistics. *Population Census, 1960: Sample Tabulation*, population census series, No. 8, "Miscellaneous".
Stewart, William A.
1962a   "Creole languages in the Caribbean", in Rice, 34-53.
1962b   "An outline of linguistic typology for describing multilingualism", in Rice, 15-25.
Sutherlin, Ruth E.
1962   "Language situation in East Africa", in Rice, 65-78.

Tabouret-Keller, A.
   1968 "Sociological factors of language maintenance and language shift: A
         methodological approach based on European and African examples",
         in Fishman *et al.*, 107-18.
Tauli, Valter
   1968 *Introduction to a Theory of Language Planning*. Acta Universitatis
         Upsaliensis, Studia philologiae Scandinavicae Upsaliensia, 6 (Uppsala:
         Almquist & Wiksell).
Tipologia socioeconómica de los países latinoamericanos
   1963 *Revista interamericana de ciencias sociales*, vol. 2, número especial.
Tocqueville, Alexis de
   1954 *Democracy in America*. Trans. of *De la démocratie en Amérique*, by
         Henry Reeve. Ed. Phillips Bradley, 2 vols. (New York: Vintage Books).
Togan, A. Zeki Velidi
   1942-7 *Bugünkü Türkili (Türkistan) ve yakın tarihi*. Vol. I, *Batı ve Kuzey
         Türkistan* (Istanbul: Ibrahim Horoz & Güven).
Tokarev, S. A. and N. N. Čeboksarov (eds.)
   1964 *Narody zarubežnoj Jevropy*. Vol. I (Moscow: Nauka).
U.S. Army Handbook for Iran
   1963 (Washington: Special Operations Research Office).
Valdman, Albert
   1968 "Language standardization in a diglossia situation: Haiti", in Fishman
         et al., 313-26.
Van den Berghe, Pierre L.
   1968 "Language and 'nationalism' in South Africa", in Fishman *et al.*,
         215-24.
Verba, Sidney
   1965 "Comparative political culture", in Lucian W. Pye and Sidney Verba
         (eds.), *Political Culture and Political Development* (Princeton: Prince-
         ton).
Wurm, Stephen A.
   1968 "Papua-New Guinea nationhood: The problem of a national language",
         in Fishman *et al.*, 345-63.