The Quarterly Review of Economics and Finance, Vol. 37, Special Issue, 1997, pages 245–262

ISSN 1062-9769

The Effects of Education, Parental Background and Ethnic Concentration on Language

CHRISTIAN DUSTMANN
University College of London

This paper studies the acquisition of language capital of immigrants to Germany. The analysis is based on a detailed survey data set. Speaking abilities and writing abilities are investigated for both male and female migrants. The study has three main objectives. Firstly, to investigate in detail the effect of education on language acquisition. For this purpose, extensive information on pre- and post migration education is utilized, and education is broken down to its different components. Secondly, to investigate the effect of parental education on language acquisition. And thirdly, to analyze the importance of living in ethnic enclaves on language ability. The analysis reveals that parental education has a surprisingly strong effect on both components of language capital, while the effect of ethnic concentration is quite moderate.

I. INTRODUCTION

Ever since Chiswick's (1978) seminal article on the earnings assimilation of immigrants to the US labor market, the economic adjustment of migrants has been a major research area in the economic literature on migration. In recent years a specific component of human capital has found a lot of interest: the ability of the migrant to communicate in the language of the host country. Language capital is likely to play a crucial role in the migrant's process of economic assimilation. Language abilities are not only important as a key requirement for certain jobs; they are also likely to serve as a screening device on the side of the employer. Furthermore, language may help the migrant to acquire information about the host country labor market, and to improve job and career opportunities.

A variety of studies have analyzed the process of language acquisition for countries like the US, Australia, Germany, and Israel.[1] Chiswick and Miller (1995) define language fluency as a function of three key variables: *economic incentives, efficiency, and exposure.*

Incentives have two dimensions. Firstly, a static dimension, reflected by the premium the labor market associates with language fluency at any point in time. Incentives to learn a foreign language are certainly low if job opportunities for migrants are restricted to certain occupations, where productivity is not enhanced by language abilities. Secondly, an intertemporal dimension, reflected by the time period over which this premium is captured. The longer the time the migrant remains in the host country labor market, the larger will be his incentive to invest in language skills for a given wage premium. This second point is especially important, since migrants are likely to be quite heterogeneous with respect to the time they remain in the foreign labor market. This is true not only for temporary migrations, but also for permanent migrations, where immigrants enter the country at different ages.[2]

Efficiency relates directly to the technology of human capital production. The efficiency of language production is certainly affected by the individual's general level of knowledge.[3] The technology of language acquisition of illiterate individuals, for instance, is lower in comparison with well educated individuals. At the extreme, an illiterate individual's only option is to learn a new language by communication, while a literate individual can undergo a systematic learning process (for instance, taking advantage of language courses). It is not surprising therefore that the literature establishes unequivocally a positive effect of educational variables on language capital. Education should, however, affect the acquisition of different components of the host country language in different ways. While it is more difficult, but possible to learn to speak a foreign language purely by communication, this is not possible for reading and writing. Here systematic learning, based on written language, is required. Comparisons between the effect of education on different language components should therefore be interesting. Besides speaking fluency, writing abilities are also analyzed below.

To date, most studies do not distinguish between the effect of education on language capital which is acquired in the host country, and education which is acquired in the home country. It is therefore difficult to establish whether the positive effect of education is indeed due to knowledge being self-productive in the acquisition of language capital, or whether the positive effect is mainly caused by an active learning of the host country language while attending host country schools. Furthermore, education is usually measured in years. This imposes the assumption that educational achievements of equal length have the same effects on language acquisition. However, even if a linear effect of education has sometimes been found for some countries with respect to productivity (or wages), it is not clear altogether that this is also true for language capital. In this study, detailed information on different types of education are used, and it is distinguished between education acquired in the host country, and education acquired in the home countries.

Many studies have found that language ability increases significantly the migrant's earnings position in the host country. It seems established that language is an important, and maybe the most important, component of host

country specific human capital. However, different from the ability to run a particular machine, for instance, language acquisition is also determined by factors other than pure productivity considerations. Language allows social communication, understanding of the host country society, and supports the acquisition of cultural knowledge, or leisure activities.

It is very likely that attitudes which value language for its own sake, and the sake of non-labor market activities, are shaped at young ages and inside the family. For instance, the interest of the father in cultural or political issues may well inseminate similar interests in the offspring. In turn, the father's intellectual broadness may largely be determined by his education. Therefore, individuals from families with higher educational background may tend to be more likely to develop an interest for all those other things language can buy. In other words, there are other incentives for language acquisition than just those related to the labor market, and these may be positively related with parental education. This hypothesis may be tested by including parental background variables in a language equation, conditioning on the individuals' educational background additionally.

In this study, some unique information on the father's educational background is used for this purpose.

Exposure is a most important factor for language acquisition, and it may take many forms. Chiswick and Miller (1995) distinguish between units of exposure and intensity of exposure. One may as well say that exposure has a time dimension, and an intensity dimension. Given intensity, a most important variable in all language studies is the number of years of residence. Variables which reflect the intensity of exposure are, for instance, the family background, and partner characteristics. In a recent paper, Chiswick and Miller (1994) emphasize the importance of ethnic networks on immigrants' language proficiency. Chiswick and Miller use what they call a *minority-language concentration* variable to determine the intensity effect of the ethnic composition of the migrant's immediate environment on language proficiency.[4] However, although the immigrant may live in an area with a high percentage of individuals who speak, for instance, the host country language, the immediate neighborhood, and the social neighborhood network, may consist of co-patriots, and vice versa. One may therefore argue that an appropriate measure is the subjectively perceived concentration of ethnic co-patriots in the migrants' environment and neighborhood. Here I use some unique information on the ethnic minority environment, as seen from the migrant's perspective.

Factors which affect language acquisition are likely to have different effects on males than on females.[5] One reason may be that males and females tend to have different roles in professional and in private life. Furthermore, if language is predominantly determined by labor market considerations, then different life cycle plans of males and females may change the effect of the same variable on language capital. The analysis below therefore distinguishes between males and females. The difference in the stock of language capital between males and

females may be due to differences in endowments, or due to differences in the coefficients. A decomposition analysis is performed which isolates these two components.

The main results emerging from the analysis are as follows. Firstly, education in both home and host country affects language positively, with the latter having slightly stronger effects. Furthermore, education has a stronger effect on writing abilities than on speaking abilities. Secondly, the education level of the father has a strikingly strong effect on language, conditional on education achieved by the individual. Thirdly, the effect of ethnic concentration variables is small. And lastly, the decomposition analysis shows that the largest part of the differential in speaking proficiency between males and females is due to differences in coefficients. For writing abilities, the larger part of the difference is accounted for by differences in endowments.

The structure of the paper is as follows: the Section II explains the data and the variables used for the analysis, and discusses briefly the estimation techniques. Section III discusses the results, and Section IV summarizes.

II. DATA AND VARIABLES

The data used for this study are from the *German Socio-Economic Panel* (SOEP). The structure of the SOEP is similar to that of the American *Panel Study of Income Dynamics* (PSID).

One distinctive difference, however, is that it over-samples individuals with non-German nationality. It therefore solves a problem survey data sets usually suffer from, namely, too small numbers of observations to study minorities. On the other hand, it allows exploiting the advantage of survey data, which is detailed information on the individual and the household.

Individuals of foreign nationality in the data are nearly exclusively migrants who came to Germany between 1955 and 1973, and who are of Turkish, Italian, Yugoslavian, Greek and Spanish nationality.[6] It is important to note that in none of these countries is German the spoken language. Furthermore, German is also usually not the first language to be taught at school (which is English). It is therefore unlikely that immigrants, who for the major part are labor migrant who are relatively low educated, had knowledge of the German language at arrival.

Before 1973, severe shortage of labor in the heavily expanding German industry and, at the same time, excess supply of labor in most Southern European countries and Turkey, led to mutually beneficial migration movements. Workers were actively recruited in their home countries. After 1973, active recruitment came to a halt, and in subsequent years the inflow of migrants from these countries consisted mainly of family members. Furthermore, large numbers returned back to their home countries. Accordingly, the data used in this analysis are self selected stock data.

Table 1. **Language Proficiency, Males and Females, 1987**

Variable	Speaking German		Writing German	
	Freq.	Perc.	Freq.	Perc.
A. Males				
very well	122	10.65	72	6.28
well	412	35.95	196	17.10
intermediate	425	37.09	312	27.23
badly	176	15.36	312	27.23
very badly	11	0.96	254	22.16
B. Females				
very well	86	9.48	62	6.84
well	217	23.93	111	12.24
intermediate	301	33.19	136	14.99
badly	252	27.78	227	25.03
very badly	51	5.62	371	40.90

Source: SOEP, wave 3, 1987.

For the present analysis, the third wave of the SOEP is used because it contains some unique information on parental background variables, and on ethnic composition of the neighborhood.

Language proficiency is self reported on a scale with 5 categories. Two components of language are reported: writing abilities, and speaking abilities. Table 1 shows frequencies and percentages for both components, and for males and females. It is obvious that females are less proficient in speaking and writing the host country language than males. Furthermore, while speaking abilities are quite strongly developed (83% of male immigrants claim to speak the German language at least on an intermediate level), this is not the case for writing abilities. Here 33% of all males, and nearly 65% of all females claim to have bad or very bad command of this language component. For the analysis below, the information in Table 1 is condensed into two (0-1) variables, which take the value one if the respondent reports good or very good command of the respective language component.

Table 2 contains variable descriptions and summary statistics for the explanatory variables, separated for males and females. The number of observations of the male sample is 1146, and of the female sample is 907. Only individuals who were older than 18 years in 1987 are included in the samples.

At the time of the interview, male and female migrants resided in Germany for 17.4 and 15.7 years respectively. The difference in the years of residence reflects a typical pattern of labor migration: the wife follows after the husband has established himself in the host country. Males are also slightly older than females, and the overwhelming part of the sample population is married.

The variables SCHG(i), i = 1,2,3 are dummy variables which indicate whether the individual received any schooling in Germany. After primary school

Table 2. **Variable Description and Summary Statistics**

code	Description	Males		Females	
		Mean	StdD	Mean	StdD
YSM	years since migration	17.389	5.53	15.706	5.80
AGEENTRY	age at entry	23.290	9.60	23.583	10.57
M	dummy; 1 if married	0.783		0.826	
SCHG1	dummy; 1 if secondary general school in Germany	0.077		0.057	
SCHG2	dummy; 1 if intermediate school in Germany	0.109		0.077	
SCHG3	dummy; 1 if grammar school in Germany	0.018		0.020	
SCHG3	dummy; 1 if university in Germany	0.007		0.003	
SCHH1	dummy; 1 if basic school without degree in home country	0.282		0.353	
SCHH2	dummy; 1 if basic school with degree in home country	0.458		0.374	
SCHH3	dummy; 1 if advanced school in home country	0.123		0.070	
SCHH4	dummy; 1 if university in home country	0.022		0.014	
TRAING1	dummy; 1 if apprenticeship in Germany	0.099		0.055	
TRAINH1	dummy; 1 if job--specific training in home country	0.104		0.060	
NEIGH1	dummy; 1 if all aliens in neighborhood same nationality	0.194		0.205	
NEIGH2	dummy; 1 if some aliens in neighborhood same nationality	0.534		0.547	
SCHF1	dummy; 1 if father acquired no schooling	0.208		0.173	
SCHF2	dummy; 1 if father acquired basic schooling w/o degree	0.349		0.372	
SCHF3	dummy; 1 if father acquired basic schooling with degree	0.306		0.299	
SCHF4	dummy; 1 if father acquired advanced schooling or university	0.041		0.033	
SCHF5	dummy; 1 if no information on father's education	0.094		0.121	
TUR	dummy; 1 if Turkish	0.338		0.342	
JUG	dummy; 1 if Yugoslavian	0.184		0.211	
GR	dummy; 1 if Greek	0.144		0.156	
SPA	dummy; 1 if Spanish	0.132		0.118	
ITA	dummy; 1 if Italian	0.202		0.174	
N. Obs.		1146		907	

Source: SOEP, wave 3, 1987.

(4 years), the individual has three choices in the German school system. To attend a secondary general school (SCHG1; further 5-6 years), to attend an intermediate school (SCHG2; further 6 years), or to attend a high school (SCHG3; further 9 years). Only those who attend high schools are entitled to enroll in university. General and intermediate schools are usually followed by

vocational training in the apprenticeship system, which provides qualification for many jobs which require college education in the U.S. (for instance, medical services like nurses). Whether the individual took part in an apprenticeship scheme in Germany is indicated by the variable TRAING1.

The numbers in Table 2 indicate that 19% of the males, and 15% of the females attended schools in Germany. Furthermore, 10% of males and 6% of females went through an apprenticeship system. The dummy SCHG4 is one if, additionally, the individual attended a university in Germany. This is the case only for a very small fraction of the sample.

The variables schh(i), i = 1, 2, 3 are indicators for schooling received in the home countries. Here the data allows distinguishing between the following types of school education: basic schooling without degree (SCHH1), basic schooling with degree (SCHH2), and advanced schooling (SCHH3). The variable SCHH4 is equal to one if, additionally, the individual attended a university in the home country. Finally, the variable TRAINH1 indicates whether the individual has received some job-specific training in the home country.

The variables SCHF(i), i = 2, 3, 4, 5, measure the father's educational achievements, where the variable SCHF5 is one if no information was available. Here the numbers for males and females are quite similar, which reflects that both come from similar educational backgrounds.

The variables NEIGH1 and NEIGH2 are based on the individual's opinion about the ethnic composition of the neighborhood. Details on the questionnaire questions are given in Table A1. The variable NEIGH1 takes the value one if the individual reports that many or some aliens live in the neighborhood, and they are *all* of the same origin. The variable NEIGH2 takes the value one if the individual reports that many or some aliens live in the neighborhood, and *some* are of the same origin. According to the numbers in Table 2, more than 70% of the sample individuals live in areas where at least some of the neighbors are copatriates.

The last set of variables indicates the country of birth. The largest group in the sample are Turkish born individuals, followed by individuals from Yugoslavia.

III. ESTIMATION

Probit models are estimated, where the dependent variable is equal to one if the individual reports good or very good command of the respective language component. Estimations are performed separately for males and females, and results are reported in Tables 3-7.[7] To ease interpretation, marginal effects are sometimes reported, which are calculated for an individual with the following characteristics: 23 years old at entry, 16 years of residence, and married. All other indicators are set equal to zero. Respective standard errors are calculated by the Delta-method (see, for instance, Greene, 1991).[8]

The Pseudo-R^2 reported in the Tables is that of McKelvey and Zavoina (1975). Veal and Zimmermann (1992) recommend this measure because it mimics the R^2 of the underlying latent variable model.

A. Language and Education

Results on the basic specification on speaking and writing for males are presented in columns 1 and 3 of Table 3. Table 4 reports results for females. As also observed in other studies, the age at entry has a strong negative effect on the probability of the individual of being fluent in the host country language. For males, being 10 years older at entry reduces the probability of speaking or writing the German language well or very well by 21% and 10% respectively; the respective numbers for females are 19% and 20%. While the years of residence variable YSM affects speaking fluency for both genders positively (10 years in the country increase male fluency by 10%, keeping age at entry constant), the effect on writing abilities is insignificant. This reflects the fact that the acquisition of writing skills requires a more systematic approach than a simple exposure to the host country, as indicated by the residence variable.

Various schooling indicators are included in the estimation equation. The reference category is individuals who received no schooling at all. Indicators for schooling acquired in Germany have the expected signs, and are mostly significant. The effect on both components of language skills is quite considerable, and slightly larger for writing than for speaking. This is true for both males and females. For instance, having attended an intermediate school in Germany (SCHG2) increases the probability of being fluent in German by 18%, and it increases the probability to write the language at least well by 23%.

Compared to schooling received in Germany, schooling in the home country seems to have a weaker effect on language. Only individuals with at least advanced schooling at home (SCHH3) do significantly better in terms of language skills. This is true for both males and females. The effect of advanced schooling on written language is quite substantial in size: Calculating the marginal effects for the reference individual shows that those with advanced schooling acquired in the home country have a 17% higher probability to write in German at least well, compared to the reference individual (no schooling). The additional effect of university education in the home country (SCHH4) is strong and significant throughout.

Apprenticeship training received in Germany (TRAING1) has a positive, but hardly significant effect for males. For females, the effect of this variable is stronger, and significant. For this group, apprenticeship training received in Germany increases the probabilities of both speaking and writing by 29%. Job specific training acquired in at home (TRAINH1) has no significant effect on language.

The results indicate that education acquired in Germany is more important for language abilities than education received at home. Furthermore, formal

education seems to play a bigger role for writing abilities than for speaking abilities. This indicates that knowledge increases the efficiency of speaking capital production. Another explanation, however, is that the better educated select into occupations where language is acquired more easily at the work place (because of German colleagues), or where the pay off for language capital is higher, therefore creating larger incentives.

B. Parental Education

Columns 2 and 4 in Tables 3 and 4 report results when including educational background variables for the father. The reference group is those individuals whose fathers have received no education, and the category SCHF5 indicates that no information on the father's educational background is given. To ease the discussion, marginal effects for the reference individual are reported in Table 5.

The effects of the father's education on language capacity are surprisingly strong, and most estimates are quite precise. Notice that all these effects are

Table 3. **Education and Language, Males**

	Speaking				Writing			
Variable	**Coef**	**t**	**Coef**	**t**	**Coef**	**t**	**Coef**	**t**
Constant	0.4091	1.45	0.1731	0.57	−0.7376	2.32	−1.0431	2.92
YSM	0.0274	3.26	0.0266	3.15	0.0133	1.37	0.0114	1.16
AGEENTRY	−0.0531	8.64	−0.0523	8.39	−0.0354	4.90	−0.0358	4.81
M	−0.3218	2.58	−0.2957	2.35	−0.5325	4.17	−0.5040	3.88
SCHG1	0.3712	1.97	0.3196	1.67	0.5713	3.10	0.5170	2.77
SCHG2	0.4721	2.61	0.4262	2.33	0.7839	4.34	0.7513	4.09
SCHG3	0.7674	2.05	0.7117	1.92	0.7337	2.34	0.6869	2.17
SCHG4	0.4521	0.88	0.4240	0.82	1.6600	2.85	1.5477	2.71
SCHH1	−0.0468	0.31	−0.0581	0.38	−0.0219	0.13	0.0042	0.02
SCHH2	0.1395	0.96	0.1038	0.70	0.2359	1.46	0.2311	1.39
SCHH3	0.4436	2.48	0.3453	1.89	0.5849	3.06	0.5015	2.56
SCHH4	0.9017	3.03	0.7971	2.63	1.4621	4.85	1.3192	4.22
TRAIND1	0.2803	1.75	0.2422	1.50	0.3425	2.24	0.2866	1.85
TRAINH1	0.0370	0.27	0.0322	0.23	−0.0700	0.39	−0.0622	0.34
SCHF2			0.2384	1.99			0.2341	1.43
SCHF3			0.3690	2.98			0.5105	3.17
SCHF4			0.5941	2.64			0.9090	3.59
SCHF5			0.1946	1.18			0.3571	1.67
TUR	−0.0272	0.18	0.0362	0.02	0.3032	1.71	0.3315	1.85
JUG	0.8796	5.72	0.8679	5.61	0.9849	5.38	0.9601	5.19
GR	0.3567	2.25	0.3595	2.26	0.5209	2.65	0.5354	2.70
ITA	0.2643	1.79	0.2895	1.95	0.3753	2.06	0.4026	2.18
Log-Lik.	−629.6		−623.5		−662.4		−452.4	
R^2_{MZ}	0.19		0.19		0.12		0.12	

Source: SOEP, wave 3, 1987.

Table 4. **Education and Language, Females**

	Speaking				Writing			
Variable	Coef	t	Coef	t	Coef	t	Coef	
Constant	0.3961	1.17	−0.0938	0.24	0.1513	0.39	−0.7873	1.55
YSM	0.0238	2.57	0.0269	2.84	0.0118	1.07	0.0157	1.37
AGEENTRY	−0.0498	7.26	−0.0487	6.93	−0.0516	5.86	−0.0498	5.47
M	−0.6547	4.42	−0.6554	4.38	−0.4864	3.00	−0.4730	2.87
SCHG1	0.5010	2.02	0.4400	1.75	0.6817	2.73	0.6049	2.39
SCHG2	0.6244	2.74	0.5288	2.29	0.6899	2.97	0.5899	2.49
SCHG3	1.6027	2.58	1.3911	2.38	2.0532	3.47	1.8905	3.26
SCHG4	4.8570	0.08	4.0429	0.11	5.2313	0.09	4.8963	0.08
SCHH1	−0.0771	0.47	−0.0823	0.50	−0.4302	2.35	−0.4472	2.39
SCHH2	0.1691	1.04	0.0919	0.55	−0.2668	1.45	−0.3449	1.82
SCHH3	0.6317	2.67	0.4764	1.97	0.6280	2.59	0.4874	1.95
SCHH4	0.8158	1.85	0.7859	1.77	1.4828	3.21	1.5000	3.16
TRAIND1	0.7786	2.79	0.7614	2.72	0.7447	3.07	0.7316	3.00
TRAINH1	0.2144	1.04	0.1575	0.76	0.0553	0.20	−0.0060	0.02
SCHF2			0.4313	2.42			0.9220	2.85
SCHF3			0.6413	3.49			1.0162	3.11
SCHF4			1.0390	3.35			1.1582	2.68
SCHF5			0.3245	1.47			0.7529	2.08
TUR	−0.2240	1.19	−0.1911	1.01	−0.0871	0.38	−0.0411	0.18
JUG	0.8929	4.78	0.9256	4.89	0.7360	3.26	0.7711	3.38
GR	0.1508	0.76	0.1671	0.83	0.1313	0.54	0.1476	0.59
ITA	−0.0526	0.27	−0.0271	0.13	−0.6906	2.57	−0.6841	2.52
Log-Lik.	−396.0		−386.8		−262.4		−255.05	
R^2_{MZ}	0.16		0.16		0.10		0.11	

Source: SOEP, wave 3, 1987

conditional on the individual's own educational achievement. For males, the fact that the father acquired basic schooling with a degree (SCHF3) increases the probability of speaking the German language well or very well by 14%, and to write the German language well or very well by 10%. If the father acquired advanced schooling, or a university degree (which is the case for 4.1% of the sample population), these effects are even stronger (22% and 18% respectively).

For females, parental educational background has even stronger effects on both components of language capital. The fact that the father attended a school (without achieving a degree, SCHF2) increases the probability of speaking the language well or very well by 16%. If the father acquired advanced schooling, or a university degree (which is the case for 3.3% of the sample), this probability is raised by as much as 39%.

The father's education interacts with the offspring's education, and inclusion of the father's educational background reduces the size and sign of other educational variables slightly. The size of the conditional effect of the parental variables is surprising. Altogether, the results add nicely to the findings of studies on inter-generational mobility (for instance, Solon, 1992; Zimmerman,

Table 5. **Paental Background, Marginal Effects**

Variable	Speaking German		Writing German	
	ME	t	ME	t
A. Males				
SCHF2	0.091	2.06	0.047	1.59
SCHF3	0.141	3.16	0.103	3.04
SCHF4	0.227	2.67	0.183	2.69
B. Females				
SCHF2	0.162	2.69	0.206	2.99
SCHF3	0.241	3.92	0.227	2.97
SCHF4	0.390	3.42	0.258	2.31

Source: SOEP, wave 3, 1987.

1992; Borjas, 1992) and emphasize the importance of educational parental background not only for the offspring's educational achievement, but also for the acquisition of other factors of human capital. The conclusion that can be drawn from the present analysis is that parental education has a considerable effect on language capital, over and above eventual effects on the offspring's educational achievements.

C. Neighborhood Effects and Country of Birth

Tables 6 and 7 report results of estimations where neighborhood effects are explicitly taken into account. The respective questions in the questionnaire, and the variables constructed from this information, are explained in Table A1. The two variables NEIGH1 and NEIGH2 summarize this information as follows: NEIGH1 equals one if all aliens in the neighborhood are of the same origin (answer 1, question II), and NEIGH2 equals one if some aliens in the neighborhood are of the same origin (answer 2, question II); both are conditional on the fact that many, or some aliens live in the neighborhood (1 and 2, question I). The base category are individuals who report to live in a neighborhood with no aliens, or in a neighborhood where all alien neighbors are of a different origin.[9]

It should be mentioned that ethnic enclave variables of whatever type are potentially endogenous in a language equation—migrants may move into an ethnic community because they speak the host country language quite badly. In fact, Rivera Batiz (1996) argues that immigrants with bad command of the host country language tend to seek employment opportunities in ethnic enclaves. Therefore, estimated coefficients on these variables should be considered with care.

Consider first estimation results on the male sample (Table 6). Columns 1 and 3 report estimates for the basic specification, with nationality dummies excluded. While the coefficients have the expected sign, and are decreasing in magnitude, only the variable NEIGH1 is significant in the spoken language equation. Living in a neighborhood where all alien neighbors are of the same origin reduces the probability of being fluent by 9.8%.

Table 6. **Neighborhood and Language, Males**

	Speaking				Writing			
Variable	Coef	t	Coef	t	Coef	t	Coef	t
Constant	0.7500	2.99	0.5353	1.84	−0.1069	0.39	−0.6044	1.85
YSM	0.0267	3.45	0.2639	3.12	0.0083	0.95	0.0118	1.21
AGEENTRY	−0.0504	8.45	−0.0532	8.63	−0.0328	4.69	−0.0355	4.89
M	−0.3179	2.60	−0.3182	2.54	−0.5263	4.22	−0.5256	4.11
SCHG1	0.2961	1.61	0.3826	2.03	0.4981	2.77	0.5907	3.20
SCHG2	0.3358	1.88	0.4820	2.66	0.6462	3.68	0.8003	4.42
SCHG3	0.6823	1.86	0.7483	1.99	0.6704	2.16	0.7126	2.26
SCHG4	0.3274	0.65	0.3751	0.72	1.4825	2.54	1.5952	2.70
SCHH1	−0.0492	0.33	−0.0408	0.27	−0.0405	0.25	−0.0167	0.10
SCHH2	0.2039	1.44	0.1401	0.96	0.2651	1.68	0.2387	1.47
SCHH3	0.3805	2.19	0.4454	2.49	0.5470	2.93	0.5918	3.08
SCHH4	0.8496	2.94	0.8732	2.91	1.3457	4.64	1.4290	4.71
TRAIND1	0.2605	1.63	0.2608	1.61	0.3070	2.03	0.3199	2.07
TRAINH1	−0.0418	0.31	0.0388	0.28	−0.1331	0.77	−0.0677	0.38
NEIGH1	−0.2589	2.12	−0.1943	1.52	−0.2248	1.58	−0.2113	1.43
NEIGH2	−0.1321	1.38	−0.1793	1.81	−0.1129	1.04	−0.1829	1.63
TUR			0.0162	0.11			0.3419	1.90
JUG			0.9127	5.88			1.0156	5.50
GR			0.3744	2.35			0.5373	2.73
ITA			0.2675	1.81			0.3848	2.10
Log-Lik.	−658.4		−627.8		−481.19		−460.9	
R2	0.19		0.19		0.12		0.12	

Source: SOEP, wave 3, 1987.

It is likely that ethnic neighborhood indicators interact with the country of birth dummies, since the tendency to establish enclaves may be differently developed between individuals from different countries. Columns 2 and 4 therefore include country of birth dummies, where the reference group is immigrants with Spanish nationality. It turns out that this affects only slightly size and sign of the neighborhood variables.

Differences in language abilities between immigrants of different origin, however, remain quite significant, despite inclusion of ethnic neighborhood variables. While, compared to the Spanish base group, immigrants from Turkey are not significantly more fluent, immigrants from Yugoslavia have, on average, a 36% higher probability of speaking the host country language well or very well, and a 32% higher probability to write well or very well. One possible explanation for differences in country of origin with respect to language proficiency, which I gave in another paper (Dustmann, 1994), are enclave effects—country dummies simply reflect different propensities of immigrants to live in ethnic communities. Obviously, however, country of origin effects persist even if controlling for ethnic community effects.[10] In fact, comparing Tables 3 and 6

reveals that the country of origin dummies are hardly affected by the ethnic neighborhood variables.

One possible reason for the differences with respect to country of origin may be language distance. Another reason may be cultural differences, which make the host country language more difficult to access for some immigrant groups. The low performance of Turkish immigrants could be explained by these factors. However, this would not explain the low performance of the Spanish base group. Ethnic concentration may, furthermore, also take place at the work place. Immigrants in Germany are concentrated in certain occupations, for instance, in blue collar jobs in manufacturing. Immigrants of the same origin may choose occupations and firms where co-nationals are already employed, thereby creating ethnic concentrations. Reasons are that employers screen with respect to origin, and favor immigrants from certain countries because of previous experiences. Employees, on the other hand, may find it easier and more appealing to work in firms where a large proportion of the work force are co-nationals. This may create ethnic networks at the work place. If the propensity of ethnic concentration at the work place is differently developed between individuals of different origin, and if ethnic concentration at the work place hinders language acquisition, this should be reflected by country dummies.

Table 7. **Neighborhood and Language, Females**

	Speaking				Writing			
Variable	**Coef**	**t**	**Coef**	**t**	**Coef**	**t**	**Coef**	**t**
Constant	0.5912	2.04	0.5866	1.67	0.2345	0.72	0.2897	0.71
YSM	0.0240	2.80	0.0225	2.41	0.0112	1.11	0.0102	0.92
AGEENTRY	−0.0452	6.93	−0.0500	7.26	−0.0465	5.60	−0.0521	5.90
M	−0.7243	5.00	−0.6462	4.33	−0.5369	3.41	−0.4838	2.98
SCHG1	0.3773	1.56	0.5157	2.06	0.5519	2.34	0.6713	2.68
SCHG2	0.4515	2.02	0.6022	2.63	0.4736	2.12	0.6680	2.86
SCHG3	1.3723	2.37	1.5742	2.57	1.9304	3.41	2.0030	3.42
SCHG4	4.0205	0.10	4.4458	0.12	4.5366	0.12	5.1302	0.09
SCHH1	0.0756	0.48	−0.0781	0.48	−0.2823	1.60	−0.4168	2.26
SCHH2	0.3884	2.49	0.1637	1.00	−0.0303	0.17	−0.2605	1.41
SCHH3	0.7701	3.36	0.6161	2.60	0.7866	3.35	0.6247	2.56
SCHH4	0.7881	1.87	0.8362	1.88	1.3746	3.17	1.5067	3.26
TRAIND1	0.7733	2.85	0.8137	2.88	0.6759	2.88	0.7483	3.07
TRAINH1	0.1875	0.95	0.2090	1.02	0.0040	0.01	0.0617	0.22
NEIGH1	−0.2681	1.81	−0.1420	0.90	−0.3318	1.80	−0.2365	1.20
NEIGH2	−0.2599	2.19	−0.2668	2.15	−0.1357	0.98	−0.1449	0.99
TUR			−0.2083	1.09			−0.0490	0.21
JUG			0.9074	4.84			0.7400	3.27
GR			0.1500	0.75			0.1342	0.55
ITA			−0.0640	0.32			−0.6770	2.51
Log-Lik.	-428.3		−393.5		−284.2		−261.2	
R2MZ	0.14		0.15		0.10		0.11	

Source: SOEP, wave 3, 1987.

A further explanation is selection. It could be that the degree of self selection is different for different countries. One reason may be different institutional settings and reward systems. If ability is rewarded by a low premium in the home economy, but a high premium in the host economy, then the benefits from migration tend to be higher for the more able. The reward on ability tends to be low in planned, egalitarian economies, but quite high in market economies. Yugoslavia had been a planned economy, and, interestingly, immigrants from Yugoslavia tend to out perform all other immigrants with respect to language capital. In fact, in an analysis which draws on the Roy model, Borjas (1987) shows that one condition for a positive selection is that the host country has a more dispersed earnings distribution than the home country. The selection hypothesis is also supported by findings by Schmidt (1992), who analyses differences in male earnings between immigrants from different origins. Schmidt comes to the conclusion that Yugoslavian workers clearly out perform the other nationalities.

Results for females are reported in Table 7. The effects of enclave variables on language are similar to those for males. Again, the effect of country of origin dummies is only marginally affected by inclusion of enclave variables. As for males, individuals from Yugoslavia have acquired significantly more language capital.

D. Decompositions

It is clearly obvious from the numbers in Table 1 that males are doing far better in terms of language capital than females. While 46% of the male population reports to speak the host country language at least well, this is the case only for 33% of the female sample—a difference of 13%. The difference in writing abilities is 4% in favor for males. But what drives these differences? Are they due to different endowments, or due to differences in the estimated coefficients? To answer this question, a decomposition analysis is performed below.

The difference in probabilities between males and females to master the respective language component well or very well is given by the following expression:

$$\Delta = \sum_{i=1}^{n^M} [\Phi(X_i^M \beta^M)]/n^M - \sum_{i=1}^{n^F} [\Phi(X_i^F \beta^F)]/n^F \tag{1}$$

$$= \Delta E1 + \Delta C1$$

$$= \sum_{i=1}^{n^M} [\Phi(X_i^M \beta^M)]/n^M - \sum_{i=1}^{n^M} [\Phi(X_i^M \beta^F)]/n^M$$

$$+ \sum_{i=1}^{n^M} [\Phi(X_i^M \beta^F)]/n^M - \sum_{i=1}^{n^F} [\Phi(X_i^F \beta^F)]/n^F$$

or by

$$\Delta = \sum_{i=1}^{n^M} [\Phi(X_i^M \beta^M)]/n^M - \sum_{i=1}^{n^F} [\Phi(X_i^F \beta^F)]/n^F \tag{2}$$

$$= \Delta E2 + \Delta C2$$

$$= \sum_{i=1}^{n^F} [\Phi(X_i^F \beta^M)]/n^F - \sum_{i=1}^{n^F} [\Phi(X_i^F \beta^F)]/n^F$$

$$+ \sum_{i=1}^{n^M} [\Phi(X_i^M \beta^M)]/n^M - \sum_{i=1}^{n^F} [\Phi(X_i^F \beta^M)]/n^F,$$

where n^M and n^F are the number of observations in the male and female sample, X_i^j, $j = F, M$ is the vector of regressors for individual i, drawn from sample j, β^j is the corresponding vector of parameter estimates, and $\Phi(.)$ is the distribution function of the standard normal distribution. Δ is the total difference in probabilities, and $\Delta E1$, $\Delta C1$ or $\Delta E2$, $\Delta C2$ are the components attributable to differences in endowments and coefficients respectively.

The difference between the two decompositions is that they use a different base category. The decompositions reported are therefore constructed as the average of these components: $\Delta E = (\Delta E1 + \Delta E2)/2$ and $\Delta C = (\Delta C1 + \Delta C2)/2$. The decomposition is based on estimations reported in columns 2 and 4 of Tables 3 and 4.

For the spoken language, the total difference in the probability of being fluent between males and females is 13.2%, where 9.4% can be attributed to differences in coefficients, and 3.7% to differences in endowments. For the written language, the total difference is 4.3%. Here the major part of the language capital differential is accounted for by differences in endowments (3.1%). Only 1.2% are explained by differences in endowments.

IV. SUMMARY

This paper studies two components of language capital for a sample of migrants to Germany. The analysis has three main objectives: Firstly, to investigate the effect of education on language capital, thereby distinguishing between education received in the home- and in the host country, and between the different components of education. Secondly, to analyze the effect of parental background variables on the individuals' language acquisition. And thirdly, to investigate enclave effects, using self reported indicators about the density of individuals of the same origin in the immediate neighborhood.

The analysis reveals that schooling achieved in the home countries has slightly smaller effects on the migrant's language capital than schooling

achieved in the host country. Most interestingly, the education level of the father has a surprisingly strong and positive effect on language capital. This is true for males and females, as well as for spoken and written language, and the effects are conditional on education of the individual. One explanation may be that factors which increase the motivation to acquire the foreign language are positively related to the father's educational background. The results suggest that parental education affects the acquisition of language capital not only indirectly via the offspring's education, but, additionally, directly. These direct effects appear to be strong.

Enclave effects have a surprisingly small effect on language. Conditional on enclave indicators, differences between individuals of different origin remain significant. This indicates that, as often hypothesized, country differences do not merely reflect differences in the propensity to create enclaves. In fact, at least for Germany, there is hardly any interaction between enclave variables and country of origin variables. This calls for other explanations. One is that these variables reflect ethnic concentration at the work place. Another are differences in the distance between home- and host country language. Finally, it may also be that migrants from different countries of origin undergo different selection processes.

APPENDIX

Table A1. **Neighborhood Variables**

Question	Answer
Question I	
Are there aliens living in your neighborhood?	1. Yes, many
	2. Yes, some
	3. no
	4. I do not know
Question II	
Are the aliens living in your neighborhood of the same nationality?	1. Yes, all
	2. Yes, some
	3. No, nobody

Source: SOEP, wave 3, 1987.

NOTES

*Direct all correspondence to: Christian Dustmann, University College London, Department of Economics, Gower Street, London WC1E 6BT U.K.

1. The acquisition of language capital has been studied in numerous papers, for example, Carliner (1981), McManus, Gould and Welch (1983), Kossoudji (1988), Rivera-Batiz (1990), Chiswick (1991), Beenstock (1992), Chiswick and Miller (1993), Dustmann (1994), and Chiswick and Miller (1995).

2. Dustmann (1996b) finds some evidence that the intended duration in the host country affects language proficiency.

3. This relates to the notion that human capital is self-productive, being a basic assumption in many theoretical models on human capital production—see, for instance, Ben-Porath (1967) and Heckman (1976).

4. This variable is defined as the percentage of the population in the age group 15-64 in the region which speaks the same minority language as the immigrant.

5. Boyd (1992) also emphasizes the distinct roles of men and women in host-country language proficiency.

6. See Dustmann (1996a) for details on after-war migration to Europe.

7. Likelihood ratio tests reject pooled estimation for males and females.

8. Let $\bar{x}$ be the vector of means of the explanatory variables for the respective sample, and $\hat{\beta}$ the estimated parameter vector. Then the marginal effect of a variable x_i on the probability of having a good or very good command over the respective component of the German language is given by $\phi(\bar{x}'\hat{\beta})x_i = \hat{\xi}$, where ø is the density of the standard normal distribution. The asymptotic covariance matrix of the marginal effects is then given by $C = \left[\frac{\partial \hat{\xi}}{\partial \hat{\beta}}\right] V \left[\frac{\partial \hat{\xi}}{\partial \hat{\beta}}\right]'$, where V is the variance-covariance matrix of the estimated parameter vector $\hat{\beta}$. Entries on the main diagonal of C are used to calculate the t-statistics in Table 5.

9. The last category was also added as an additional dummy variable, but it turned out to be insignificant in all estimations.

10. Migrants, of course, may have moved into different communities, and language acquisition at earlier stages of their migration history may well have been affected by different neighborhood compositions.

REFERENCES

Beenstock, M. (1993). *Learning Hebrew and Finding a Job: Econometric Analysis of Immigrant Absorption in Israel.* Paper presented at the CEPR-Conference The Economics of International Migration: Econometric Evidence, Konstanz, 1993.

Ben-Porath, Y. (1967). "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy, 75*: 352-365.

Boyd, M. (1992). "Gender Issues in Immigration and Language Fluency." Pp. 305-372 in *Immigration, Language and Ethnic Issues: Canada and the United States*, edited by B.R. Chiswick. Washington, D.C.: American Enterprise Institute.

Borjas, G.J. (1987). "Self-Selection and the Earnings of Immigrants," *American Economic Review, 77*: 531-553.

_____. (1992). "Ethnic Capital and Intergenerational Mobility," *Quarterly Journal of Economics, 107*: 123-150.

Carliner, G. (1981). "Wage Differences by Language Group and the Market for Language Skills in Canada," *Journal of Human Resources, 16*: 384-399.

Chiswick, B.C. (1978). "The Effect of Americanization on the Earnings of Foreign-Born Men," *Journal of Political Economy, 86*: 897-921.

_____. (1991). "Reading, Speaking, and Earnings among Low-Skilled Immigrants," *Journal of Labor Economics, 9*: 149-170.

Chiswick, B.R. and P. Miller. (1985). "Immigrant Generation and Income in Australia," *Economic Record*: 540-553.

_____. (1993). "Language in the Immigrant Labor Market." Pp. 229-296 in *Immigration, Language and Ethnicity: Canada and the United States*, edited by B.R. Chiswick. Washington, D.C.: AEI Press.

_____. (1994). *Ethnic Networks and Language Fluency among Immigrants*. Paper presented at the ESF conference Migration and Development, Greece, Crete.

_____. (1995). "The Endogeneity Between Language and Earnings: An International Analysis," *Journal of Labor Economics, 13*: 201-246.

Dustmann, C. (1994). "Speaking Fluency, Writing Fluency and Earnings of Migrants," *Journal of Population Economics, 7*: 133-156.

_____. (1996a). "Return Migration—The European Experience," *Economic Policy, 22*: 213-250.

_____. (1996b). *Temporary Migration, Human Capital, and Language Fluency of Migrants*. CEPR discussion paper No. 1376.

Greene, W.H. (1991). *Econometric Analysis*. New York: MacMillan.

Heckman, J.J. (1976). "A Life Cycle Model of Earnings, Learning and Consumption," *Journal of Political Economy, 84*(supplement): S11-S44.

Kossoudji, S.A. (1988). "English Language Ability and the Labor Market Opportunities of Hispanic and East Asian Immigrant Men," *Journal of Labor Economics, 6*: 202-228.

McKelvey, R. and W. Zavoina. (1975). "A Statistical Model for the Analysis of Ordinal Level Dependent Variables," *Journal of Mathematical Sociology, 4*: 103-120.

McManus, W., W. Gould and F. Welch. (1983). "Earnings of Hispanic Men: The Role of English Language Proficiency," *Journal of Labor Economics, 1*: 101-130.

Rivera-Batiz, F. (1990). "English Language Proficiency and the Economic Progress of Immigrants," *Economics Letters, 34*: 295-300.

_____. (1996). "English Language Proficiency, Quantitative Skills, and the Economic Progress of Immigrants." Pp. 60-76 in *Immigrants and Immigration Policy: Individual Skills, Family Ties, and Group Identities*, edited by H. Duleep and P. Wunnava. Greenwich, CT: JAI Press.

Schmidt, C.M. (1992). *Country-of Origin Differences in the Earnings of German Immigrants*. Discussion Paper 92-29, University of Munich.

Solon, G. (1992). "Inter generational Income Mobility in the United States," *American Economic Review, 82*: 393-408.

Veal, M.R. and K.F. Zimmermann. (1992). "Pseudo-R^s in the Ordinal Probit Model," *Journal of Mathematical Sociology, 16*: 333-342.

Zimmerman, D.J. (1992). "Regression Toward Mediocrity in Economic Stature," *American Economic Review, 82*: 409-429.