ROCKWOOL Foundation Berlin
Institute for the Economy and the Future of Work (RFBerlin)

DISCUSSION PAPER SERIES 137/25

# Out-of-School Learning: Subtitling vs. Dubbing and the Acquisition of Foreign-Language Skills

Frauke Baumeister, Eric A. Hanushek, Ludger Woessmann

www.rfberlin.com

DECEMBER 2025

# Out-of-School Learning: Subtitling vs. Dubbing and the Acquisition of Foreign-Language Skills


## Authors

Frauke Baumeister, Eric A. Hanushek, Ludger Woessmann



## Reference

**JEL Codes:** I21, Z13, L82

**Keywords:** foreign-language skills, English, TV, movies, dubbing, subtitles

**Recommended Citation:** Frauke Baumeister, Eric A. Hanushek, Ludger Woessmann (2025): Out-of-School Learning: Subtitling vs. Dubbing and the Acquisition of Foreign-Language Skills. RFBerlin Discussion Paper No. 137/25





**Imprint**

**RFBerlin**
ROCKWOOL Foundation Berlin –
Institute for the Economy
and the Future of Work

Gormannstrasse 22, 10119 Berlin
Tel: +49 (0) 151 143 444 67
E-mail: info@rfberlin.com
Web: www.rfberlin.com

# Out-of-School Learning: Subtitling vs. Dubbing and the Acquisition of Foreign-Language Skills*

Frauke Baumeister, Eric A. Hanushek, and Ludger Woessmann†

**Abstract**

English-language skills are a near necessity in today's global economy. While prior studies of language acquisition have focused on schools, we show the overwhelming influence of out-of-school learning stemming from historically rooted differences in whether countries subtitle or dub foreign TV content. We identify the causal effect of subtitling in a cross-country between-subject approach that compares English to math skills in European countries that do and do not use subtitles. We find a large positive effect of subtitling on English-language skills of over one standard deviation. Consistent with oral TV transmission, the effect is larger for listening and speaking skills than for reading. Placebo tests do not show effects on native-language reading or science skills. Results are robust to accounting for linguistic similarity, economic incentives to learn English, and cultural protectiveness.

Keywords: foreign-language skills, English, TV, movies, dubbing, subtitles

JEL classification: I21, Z13, L82

November 26, 2025

* We gratefully acknowledge comments from José Montalban and seminar participants at the CESifo Area Conference on Economics of Education, the German Economic Association, the Helsinki Economics of Education Workshop, the Workshop on Education Economics and Policy (WEEP) in Oslo, the International Workshop on Applied Economics of Education in Catanzaro, and the ifo Center for the Economics of Education. We are very grateful to Lavinia Kinne who provided important input in the early stages of this project. This work was supported by the Smith Richardson Foundation. The contribution of Woessmann is also part of German Science Foundation project CRC TRR 190.

† Baumeister: ifo Institute at the University of Munich, baumeister@ifo.de. Hanushek: Hoover Institution, Stanford University; CESifo, IZA, and NBER, hanushek@stanford.edu. Woessmann: University of Munich, ifo Institute; Hoover Institution, Stanford University; CESifo, IZA, and RFBerlin, woessmann@ifo.de.

# 1. Introduction

Globalization has increased the value of proficiency in the English language as international transactions are increasingly negotiated in English. Non-English-speaking countries have reacted by putting greater emphasis on English-language instruction in their schools. But schools are not the only source of potential development of English-language skills. Western entertainment media—movies, TV shows, and the like—are disproportionately produced in English, and consuming these offerings provides another avenue for developing English skills. The alternative treatment modes of either subtitling or dubbing imported media into the local language do, however, differ across countries. This provides insight into learning opportunities outside of schools since subtitling exposes viewers to the English voice track while dubbing does not. In this paper, we use cross-country variation in a between-subject model to estimate the causal effect of subtitling versus dubbing on the acquisition of English-language skills.

The global importance of English is well-documented (de Swaan (2001), Crystal (2003)).[1] Various studies have shown that the extent of English-language proficiency affects the level and pattern of international trade (Fidrmuc and Fidrmuc (2016)), patterns of migration (Adserà and Pytliková (2015)), and payoffs in the labor market of non-English-speaking countries (Ginsburgh and Prieto-Rodriguez (2011), Piopiunik et al. (2020), Hahm and Gazzola (2022)). As a result, in almost all European countries over 90 percent of students study English in school, and 10-20 percent of instruction time in secondary schools across European countries is dedicated to compulsory foreign-language instruction (EACEA (2023)).

Discussions of foreign-language acquisition focus primarily on teaching in schools (e.g., Strazzeri, Oggenfuss, and Wolter (2022)). But this ignores the fact that much language learning

[1] English is the language with the largest number of speakers worldwide (Ethnologue (2023)). At over one billion, the number of second-language speakers is nearly four times that of the next-largest language.

happens outside of the classroom. Moreover, formal instruction proves most effective when combined with outside language exposure (Ellis (2008), Reinders and Benson (2017)). We study the impact of this outside-of-school exposure on language learning.

In many developed countries, time spent watching TV is as large as time spent in school. The average person in France spends 23.2 hours per week watching TV, 21.2 in Germany, 26.5 in Italy, and 27.0 in Poland (RTL AdAlliance (2024)). The European average is 22.5 hours, more than in the US (17.5 hours). Much of the content is imported from the US and thus originally produced in English. Importantly, the approach to translating English-language programming differs sharply across countries where English is not the native language. One subset of countries translates the English dialogue for TV and movies into the local language through the addition of subtitles that leave the oral portion in English, while another subset dubs the dialogue into the local language so that the English dialogue effectively disappears.

A simple aggregate comparison of language outcomes across translation modes may, however, be misleading. The basic choice whether countries subtitle or dub is historical, generally made soon after sound film emerged in the 1920s (Chaume (2012), Rupérez Micola et al. (2019)). The decision to subtitle or dub may nonetheless be related to other country characteristics correlated with English skills. If subtitling is endogenous to foreign-language skills because of reverse causation or omitted variables that are associated with countries' general skill levels, such simple comparisons could be very misleading.

We pursue a cross-country between-subject approach to identify the effect of subtitling on foreign-language skills. Under the identifying assumption that math skills are not affected by TV translation mode, difference in math skills between subtitling and dubbing countries can serve as a counterfactual for the differences in foreign-language skills. This gives rise to an estimation

strategy in the difference-in-differences spirit where countries (that subtitle or dub) constitute the first difference and subjects (English or math) the second difference.

Our main analysis focuses on non-English-speaking countries in Europe, where exposure to media originally in English language is prevalent. Our analysis sample is evenly divided between countries where subtitling is common on TV such as the Nordic countries and the Netherlands and countries where dubbing is common such as France, Germany, Italy, and Spain. While dubbing countries tend to be larger, allowing them to more readily absorb the added cost of dubbing, we provide evidence that results are not driven by population or language size.

Testing the impact of subtitling on language outcomes requires clear measures of English proficiency, which are difficult to find for representative samples of country populations. We use three different measures. The first is the English Proficiency Index (EPI), the world's largest ranking of countries by adult English skills. EPI is based on the performance of over two million voluntary test takers annually of standardized adaptive tests of English skills available online free of charge. The second measure is the Test of English as a Foreign Language (TOEFL), another standardized test that is widely used for college applications. TOEFL allows analysis by separate domains such as speaking and reading. Due to costly participation, the annual number of TOEFL test takers is only roughly one third of EPI. A downside of these two data sources is that sampling is not representative of the entire population, an issue that we address in our robustness analysis. By contrast, the third measure, the Adult Education Survey (AES), draws on representative samples, but proficiency is self-rated rather than test elicited. Math achievement comes from the Programme for International Student Assessment (PISA), a standardized test of representative samples of 15-year-olds.

Wide variations in English-language skills are found across countries, regardless of the specific measure. For example, over 75 percent of 18-to-24-year-olds in Iceland as well as 68 percent in Denmark and 67 percent in Finland self-report proficient levels of English in the most recent AES, while only 16 percent in Italy and 21 percent in France do so (Eurostat (2025)). The Netherlands, Norway, and Sweden are the top European performers in the most recent EPI report (EF Education First (2024)), placing them in the category of "very high" proficiency (C2 on the Common European Framework of Reference (CEFR)). At this level, participants can on average use nuanced and appropriate language in social situations, read advanced texts with ease, and negotiate a contract with a fluent English speaker. At the other end of the European spectrum, France and Italy fall in the "moderate" proficiency category (lower range of B2 on CEFR). They can on average participate in meetings in one's area of expertise, understand song lyrics, and write professional emails on familiar subjects, but they would not on average understand TV shows, read a newspaper, or make a presentation at work in English.

The results of our analysis show that subtitling has a large positive effect on country populations' proficiency in English that is consistent across all three performance measures. Our baseline estimate of the subtitling effect on EPI scores is 1.5 standard deviations (SD). Average estimates for the representative AES are similar in size, indicating that results are not driven by selective test-taking. The AES effect is strongest at 1.5 SD in the youngest group aged 16 to 34 years but also exceeds 1.2 SD in the age groups 35-55 and 56-75. While smaller on average, the effect on TOEFL scores differs by domain, reaching 1.2 SD in speaking but being insignificant in reading—in line with oral learning from subtitled TV. Treatment effects do not differ by age or gender but increase with parental education.

A series of robustness tests supports our conclusions. Placebo tests show no effects of subtitling on native-language reading and science, addressing concerns of other country-specific factors including families and schools. Qualitative results are unaffected when accounting for the similarity of local languages with English, for the size of country populations and language communities, and for the starting age and instruction time of foreign-language teaching in schools. Results are also robust to controlling for the number of TOEFL test takers, speaking against bias from differential selectivity of test taking. Results similarly hold when using PIAAC numeracy scores or PISA reading or science scores instead of PISA math scores to produce alternative counterfactuals.

Our analysis contributes to several strands of literature. First, while studies of educational production always note the importance of learning factors outside of schools (e.g., Coleman et al. (1966), Woessmann (2016), Hanushek et al. (2022)), little analysis actually delves into the nature and strength of this. Our analysis shows that non-school factors can have significant and important impacts on learning, here on language learning. Second, while there is rich analysis of the economics of language (e.g., Gazzola, Grin, and Wickström (2016), Ginsburgh and Weber (2020), Aparicio-Fenoll and Di Paolo (2023)), it is more focused on the economic impacts of language and linguistic diversity with particular emphasis on languages of minorities and migrants (e.g., Chiswick and Miller (2015)). Our analysis highlights the importance of the larger and more common national policies that foster learning of second languages by the native population. Third, prior studies of TV viewing have studied effects on economic, social, and political outcomes (e.g., Bursztyn and Cantoni (2016), Durante, Pinotti, and Tesei (2019), Campante, Durante, and Tesei (2022)) and on child development (e.g., Gentzkow and Shapiro

(2008), Kearney and Levine (2019)). Our analysis extends the focus to the impacts of TV translation mode on foreign-language skills.

Fourth, this work expands on prior work on the impact of subtitling on language acquisition. Prior studies have primarily focused on experimental settings where children are shown short clips in different translation modes (e.g., Koolstra and Beentjes (1999), d'Ydewalle and Van de Poel (1999), Van Lommel, Laenen, and d'Ydewalle (2006)). One prior study, Rupérez Micola et al. (2019), has looked at English language skills across countries as measured by TOEFL scores. They focus on the impact of subtitling on English skills and introduce the possibility that subtitling may be related to other factors that influence English language acquisition. Their design considers cross-sectional differences in skills, and they instrument subtitling by language size at the time of the dubbing-subtitling decision for movies. They find that subtitling has a strong influence on TOEFL English scores, particularly in the listening domain. Their instrumental variable is appropriate when historical language size is uncorrelated with unmeasured family and school factors that enter into the production function for English skills, a condition that is potentially problematic. Further, the nearly complete correlation (0.95) of historical and contemporary language size raises identification concerns. Fifth, on the methodological side, we contribute to the analysis of international differences in learning (Hanushek and Woessmann (2011)). Our analysis employs a between-subject approach as an identification strategy to extract causal impacts of national policy differences in the cross-country setting.

## 2. Empirical Model

Countries' choices of translation mode are rooted in history. Shortly after the emergence of sound film in the 1920s, countries adopted either subtitling or dubbing and have stuck to their

translation modes ever since (Chaume (2012), Rupérez Micola et al. (2019)). Whether countries subtitle or dub English films is not randomly distributed across countries. Both economic and political reasons have been put forward for the historical emergence. Dubbing has been estimated to be around twenty times as expensive as subtitling (Modot et al. (2007), Ivarsson (2009)), making it profitable only if the target language community is sufficiently large. Furthermore, in the early years of translation mode adoption, fascist regimes tended to prefer dubbing, because this obscured the foreign nature of films and made censorship easier, whereas subtitles were suspected of promoting foreign languages (Danan (1991)). By contrast, choice of translation mode does not appear to be related to the distribution of English language skills at the time. Language instruction across European schools was largely confined to academic secondary schools that enrolled a small minority of the population, and English was not even the primary foreign language choice.

The historical roots may be associated with other sources of language acquisition, leading to bias in cross-sectional estimates of the effect of subtitling on English skills. For example, countries choosing subtitling might have had higher literacy levels or might not have been able to afford the higher cost of dubbing (Chaume (2012)), factors also potentially related to features of families and school systems that affect levels of all skill dimensions. More generally, any country characteristics associated both with populations' overall skill levels and with translation modes would bias cross-sectional estimates.

To identify the effect of subtitling on English-language skills in the presence of such potential biasing factors, we adopt a cross-country between-subject approach. We consider two subjects *s*, English and math, that are each produced in a standard education production function

where subject-specific test scores *T* of country *c* are a function of standard inputs in the areas of families and schools plus an indicator for subtitles:

$$T_{cs} = \alpha + \beta Families_c + \gamma Schools_c + \delta Subtitles_c + \mu_c + \varepsilon_{cs} \quad \forall\, s \in (e, m) \qquad (1)$$

The error term has a country-specific component $\mu_c$ to capture other cultural and institutional aspects and a country-subject-specific component $\varepsilon_{cs.}$ The importance of families and schools has been firmly established since the earliest estimates of education production functions (Coleman et al. (1966), Hanushek (1986)). At the same time, obtaining good measures of these factors has proved elusive, making direct estimates of equation 1 impossible (see, for example, Woessmann (2016)).

Assuming that math skills are not causally affected by whether countries subtitle or dub English films, we can estimate the effect of subtitling on English skills by differencing the two production functions. In such a between-subject model, the inputs from families, schools, and any unobserved country factors drop out as long as they have the same parameters in both production functions. Thus, the between-subject estimate will not be biased from any (observed or unobserved) characteristics of countries' families, schools, or other environmental factors whose impact does not vary by subject.

We estimate the between-subject model by stacking the two subjects and including country fixed effects $\mu_c$ and a subject fixed effect $\theta_s$:

$$T_{cs} = \vartheta + \rho(Subtitles_c * English_s) + \theta_s + \mu_c + \varepsilon_{cs} \qquad (2)$$

where the coefficient $\rho$ on the interaction between subtitling and English identifies the effect of subtitling. The baseline analysis does not include additional control variables, as overall country

characteristics are subsumed in the country fixed effects. Standard errors are clustered at the country level.

There are two key identifying assumptions underlying the between-subject model. First, subtitling does not affect math skills. Second, the impacts of other inputs such as families ($\beta$) and schools ($\gamma$) do not differ between subtitling and dubbing countries. Put differently (and in the spirit of a difference-in-differences approach), the identifying assumption is that the difference between English and math skills in subtitling countries would be the same as in non-subtitling (dubbing) countries in the absence of the subtitling treatment.

A central concern about this identifying assumption relates to issues surrounding the profitability of subtitling and the cultural protectiveness of a country, each of which may enter both into the historical adoption of subtitling and into other sources of English-language skills. We turn to these concerns in our robustness analyses, where we explicitly test for potential bias emanating from linguistic similarity, the size of country population and language communities, and the intensity of foreign language teaching in schools. For this, we extend the analysis to allow the between-subject skill difference to differ with various other country characteristics by including interaction terms with country controls:

$$T_{cs} = \gamma_0 + \gamma_1(Subtitles_c * English_s) + \gamma_2(Control_c * English_s) + \theta_s + \mu_c + \varepsilon_{cs} \qquad (3)$$

In addition, a placebo test provides direct information about the key identifying assumptions. Given the same assumption underlying our model, skills in other subjects such as native-language reading and science should not be causally affected by subtitling. Replacing English tests by native-language reading or science assessments in equation 2 thus provides a placebo test of the underlying model assumptions: subtitling should not have an impact on the difference in achievement between, e.g., math and native-language reading. Estimation of a

significant “effect” would speak against the underlying assumptions of unaffectedness of other subjects and of same impacts of other inputs on different subjects. Most importantly, any inputs that would affect language acquisition in general (as opposed to foreign-language skills) should give rise to an “effect” of subtitling on native-language reading skills compared to math skills in the placebo test.

Given the historical persistence of translation modes, the effect of subtitling identified in our analysis is best interpreted as the aggregate effect of having employed subtitling for two to three generations. That is, the identified effect not only captures impacts running through the TV consumption of the tested individuals but also impacts running through prior effects of subtitling on the individuals’ parents, their teachers, and any other impacts of subtitling on societies at large that may affect individuals’ language acquisition. We thus expect this overall effect of subtitling to be larger than treatment effects stemming only from individual TV viewing.

## 3. Data

We construct a country-level database of translation modes and of different assessments of English proficiency. Each of the three alternative measures of English skills—which we describe in greater detail in Appendix A—has both advantages and disadvantages. In the next section, we perform parallel analyses to understand the potential impact of disadvantages of each assessment.

### 3.1 Classification of TV Translation Modes: Subtitling and Dubbing

Subtitling is the written translation of the original voice track offered in one or two lines on screen while maintaining the original voice track. Dubbing, on the other hand, is the replacement of the original voice track with a voice track performed in the target language. In contrast to

subtitling, it is an invisible translation that adds no visual input for the audience and completely removes the original language (Gottlieb (1994)).

Countries' TV translation modes are extracted from several studies that classify countries into subtitling and dubbing (Chaume (2012), Almeida and Costa (2014), Modot et al. (2007), and Safar et al. (2011)). Appendix Table A1 and the map in Appendix Figure A1 show the classification of the 36 non-English-speaking European countries included in our analysis.[2] The Scandinavian countries, the Netherlands, and Portugal stand out for using subtitles, while neighboring countries such as Austria, France, Germany, Italy, and Spain use dubbing.[3]

We focus our analysis on European countries for several reasons. First, the import of English-language content is very common on TV in European countries. Second, European countries' TV translation mode traditions have been virtually constant historically. This likely reflects specialization of the established translation industry (Modot et al. (2007)) and that people tend to prefer the translation mode they are used to even if they have a choice (European Union (2024)). Third, there is substantial variation in translation modes, with the 36 European countries split evenly between subtitling and dubbing.[4]

[2] We exclude Estonia and Luxembourg because both subtitling and voice-over are common in Estonia and because various products including movies are shown in original version without subtitles in Luxembourg.

[3] In some Eastern European countries, the practice of voice-over—a voice track in the target language by one or two narrators over the original voice track—is common, which we classify as dubbing because the original voices are generally not readily audible. There is no significant treatment effect of voice-over vs. dubbing when separated in our empirical analysis.

[4] These prerequisites for our analysis are not found in other non-English-speaking regions of the world. For example, in Latin America virtually all countries use dubbing. In many countries in the Middle East and North Africa, the primary foreign content to be subtitled or dubbed is not in English. Many countries in Asia cannot be cleanly characterized as subtitling or dubbing because they transitioned from dubbing to subtitling (China) or differentiate by genre (India) or by original language (Indonesia, Malaysia, Singapore) (Chaume (2012)).

### 3.2 Measures of English-Language and Math Skills: EPI, TOEFL, AES, and PISA

**English Proficiency Index (EPI).** EPI is a freely available adaptive online test of English reading and listening skills voluntarily taken by people interested in learning about their English proficiency (EF Education First (2024)). Available annually since 2011 and for 30 non-English-speaking European countries, the number of test takers exceeds 2 million in recent years.

A concern about EPI is the self-selection of its test takers. Test takers are likely positively selected in terms of educational attainment and English knowledge. They are almost all working adults or college students with a median age of 26 in 2024 (EF Education First (2024)). While this selection could introduce bias, our robustness analysis indicates no impact of any selective test taking.

**Test of English as a Foreign Language (TOEFL).** TOEFL is widely accepted by universities worldwide as proof of English proficiency (ETS (2024)), with average country scores based on roughly 600,000-700,000 annual participants available since 1991. The format was paper-based (PBT) until 1997, computer-based (CBT) from 1998-2005, and internet-based (iBT) since 2005. PBT and CBT tested reading comprehension, listening comprehension, structure, and written expression domains, and iBT added a speaking section. TOEFL iBT is available for 34 non-English-speaking European countries.

Because TOEFL is most commonly taken by students applying for English-language university programs, sample selection concerns also apply to TOEFL. However, again, robustness analysis indicates TOEFL selection does not bias our results.

**Adult Education Survey (AES).** AES covers representative country samples and provides self-rated language skills (Eurostat (2024)). Conducted four times between 2007 and 2022, AES covers 28 non-English-speaking European countries and targets random samples of 2000-3000

respondents of a country's resident population. Respondents rate their foreign-language proficiency on a four-point scale from very basic to proficient.

AES' representative samples of adults allow us to assess the relevance of selectivity of the other two English proficiency measures. On the other hand, AES provides only self-rated foreign-language proficiencies, which could introduce bias if people from dubbing or subtitling countries systematically either over- or underrate their English proficiency.

**Programme for International Student Assessment (PISA).** PISA collects internationally comparable data on several skills for representative samples of 15-year-old students. It provides the math performance that is our counterfactual along with native-language reading and science skills that are central to our placebo estimates. PISA is available in eight waves from 2000 to 2022, covering 36 European countries that do not have English among their official languages.

### 3.3 Extracting Average Country Scores

The available country scores provide an unbalanced panel because not all countries participate in all waves of a test. To construct country mean scores that are not affected by differences in the timing of participation, we estimate the following model for each test series:

$$T_{ct} = \gamma + \tau_t + \mu_c + \varepsilon_{ct} \tag{4}$$

where $T_{ct}$ is the test score of country $c$ at time $t$, $\tau_t$ are wave fixed effects, and the $\mu_c$ are country fixed effects that provide the time-adjusted average country scores used in our analysis. We standardize these scores to have mean zero and SD one in the joint sample of the 24 countries for which all tests (including TOEFL PBT) are available.[5]

[5] If a test series changes its scale, we construct country mean scores across all years with the same scale and then take the average of these means, weighted by the number of years in which the respective scale was used. For example, the TOEFL score is a weighted average of the mean TOEFL PBT score (6 out of 32 years), the mean TOEFL CBT score (8 years), and the mean TOEFL iBT score (18 years).

### 3.4 Descriptive Statistics

The top-performing countries on the EPI English test are the Netherlands, Sweden, Denmark, Norway, and Finland (Appendix Table A1). The bottom-performing countries are Albania, Ukraine, Russia, Moldova, and Belarus with only somewhat higher achievement in Italy, France, and Spain.

Scores are highly correlated across our three measures of English proficiency. For example, EPI is correlated 0.856 with TOEFL and 0.846 with AES (Appendix Table A2).

English performance on EPI and TOEFL is also strongly correlated with math performance on PISA (0.590 and 0.790, respectively). These strong correlations clearly indicate that cross-sectional English language patterns capture more than just TV translation modes. The top- and bottom-performing countries—the Netherlands and Albania—are the same in the joint EPI-PISA sample. These consistent country-level score differences across tests underscore the importance of dealing with the wide range of confounding factors including family and school backgrounds in international comparisons. The importance of unmeasured family, school, and institutional factors in the level of achievement presents a problem to the instrumental variable approach in Rupérez Micola et al. (2019). If these factors are correlated with historical language sizes of countries, the exclusion restrictions needed for the IV estimation of subtitling will be violated. Because these factors have proven difficult to measure, they are not readily controlled by including observables.

## 4. Results

The impact of subtitling on English language learning is strongly supported by our results. We first describe baseline estimates across the different language measures and then consider the robustness of results to potential confounding factors.

### 4.1 Baseline Results

Subtitling consistently has a strong positive effect on proficiency in English as a second language across all three outcome measures. Table 1 reports the results of our cross-country between-subject analysis for EPI, TOEFL, and AES scores. The EPI measure of English-language skills increases by 1.45 standard deviations (SD) if a country has subtitling rather than dubbing as the TV translation mode (column 1). The point estimate on the TOEFL test is lower at 0.91 SD (column 2), likely reflecting the fact that TOEFL is a high-stakes test for anybody wanting to attend an English-language college. These incentive effects will lead TOEFL test takers to be particularly positively selected in dubbing countries if people tend to perceive a certain threshold below which test participation is not expedient and if the achievement distribution in dubbing countries is indeed to the left of subtitling countries. In contrast to these two standardized assessments, AES is self-assessed but has the advantage of representative sampling. Importantly, the effect of subtitling on the average self-rated English proficiency in AES is 1.38 SD, very similar to the EPI estimate. This similarity supports the interpretation that the test-based results are unlikely to be driven by non-representative sampling.[6]

[6] Results are similar when estimating the model on the same sample of 25 countries for which all three measures are available; the estimates are 1.38 for EPI, 0.91 for TOEFL, and 1.27 for AES (Appendix Table A3).

**Table 1: The effect of subtitling on English skills**

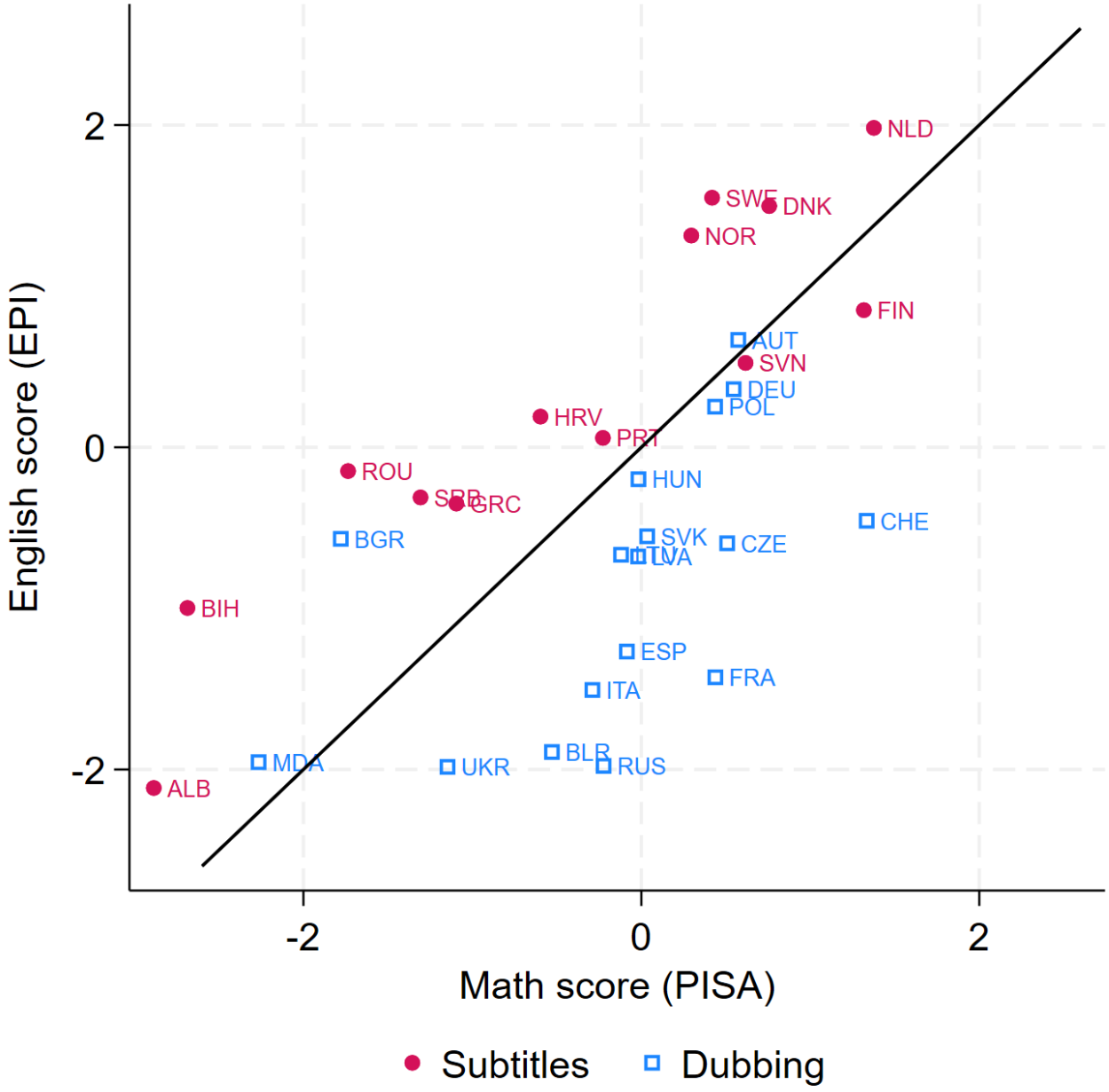
| | EPI | TOEFL | AES |
|---|---|---|---|
| | (1) | (2) | (3) |
| Subtitles * English | 1.450*** | 0.912** | 1.375** |
| | *(0.371)* | *(0.360)* | *(0.543)* |
| Subject fixed effects | Yes | Yes | Yes |
| Country fixed effects | Yes | Yes | Yes |
| Observations | 60 | 68 | 56 |
| Countries | 30 | 34 | 28 |

Notes: Least squares regressions. Dependent variable: test scores in English and math. Unit of observation: subject-specific country score, standardized in joint country sample. Subtitles: translation mode with 1 = subtitling, 0 = dubbing. Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: EPI, TOEFL, AES, PISA.

The estimated effect of over one SD implies a huge impact of dubbing. By these estimates, the English skills of the population in dubbing countries such as Austria, Germany, or Poland would be roughly as high as the English proficiency in the Scandinavian countries or the Netherlands had they adopted subtitling as their translation mode (see Appendix Table A1).

The fundamental idea underlying our identification and the main results themselves can be easily visualized in a simple scatterplot. Figure 1 plots EPI English scores against math scores across countries, indicating whether each country uses subtitles or not. There is a very strong positive association of English and math scores: countries with higher English proficiency such as the Netherlands, Sweden, and Denmark also systematically perform relatively high in math, whereas countries with low English proficiency such as Albania, Moldova, and Ukraine also have relatively low math achievement. This pattern underscores the identification issues of simple cross-sectional associations between translation mode and English skills, where some combination of families, schools, and other country-specific factors have a strong influence on achievement levels.

**Figure 1: English and math achievement across countries by translation mode**




Notes: Test scores in English and math, country average across all waves, standardized in joint country sample. See Appendix Table A1 for country abbreviations. Data sources: EPI, PISA.

A country's position relative to the 45-degree line in Figure 1 reflects whether it performs relatively better in English or in math, and the average difference in the orthogonal distances from the 45-degree line between subtitling and dubbing countries effectively depicts the subtitling effect identified in our analysis. With few and small exceptions, subtitling countries are above the 45-degree line, whereas dubbing countries are below it. That is, nearly all countries that use subtitles have English skills that are better than what would be predicted by their math skills. By contrast, nearly all countries that use dubbing perform worse in English than would be expected based on their math skills. The plot also shows that the estimated treatment effect is systematic across countries and is not driven by particular outliers.

### 4.2 Results by Skill Domain and by Age

Quite intuitively, the effects of subtitling are larger for speaking and listening than for writing and reading. The TOEFL test allows for separate analyses by the four English skill domains of speaking (only in internet version), listening, writing, and reading (Panel A of Table 2). The pattern of results is consistent with the expected effects of TV translation modes: regularly hearing people talk in English on TV should particularly affect listening and speaking skills, whereas reading and writing skills require additional learning that is not transmitted by TV viewing. In fact, the effects on English speaking and listening skills on the TOEFL test are over one SD, putting the TOEFL estimate closer to the estimates of the other two English skills measures. By contrast, the smaller estimate for reading does not differ statistically significantly from zero and is significantly smaller than the speaking and listening estimates.

**Table 2: The effect of subtitling on English skills by domain and age**

**Panel A: Effects by skill domain (TOEFL)**

| | Total | Speaking | Listening | Writing | Reading |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Subtitles * English | 0.912** | 1.159** | 1.084*** | 0.821* | 0.429 |
| | *(0.360)* | *(0.467)* | *(0.352)* | *(0.429)* | *(0.403)* |
| Subject fixed effects | Yes | Yes | Yes | Yes | Yes |
| Country fixed effects | Yes | Yes | Yes | Yes | Yes |
| Observations | 68 | 68 | 68 | 68 | 68 |
| Countries | 34 | 34 | 34 | 34 | 34 |

**Panel B: Effects by age group (AES)**

| | Age 16-75 | Age 16-34 | Age 35-55 | Age 56-75 |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Subtitles * English | 1.375** | 1.512*** | 1.284** | 1.256** |
| | *(0.543)* | *(0.511)* | *(0.520)* | *(0.602)* |
| Subject fixed effects | Yes | Yes | Yes | Yes |
| Country fixed effects | Yes | Yes | Yes | Yes |
| Observations | 56 | 56 | 56 | 56 |
| Countries | 28 | 28 | 28 | 28 |

Notes: Least squares regressions. Dependent variable: test scores in English and math. Unit of observation: subject-specific country score, standardized in joint country sample. Subtitles: translation mode with 1 = subtitling, 0 = dubbing. Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: TOEFL, AES, PISA.

It is possible that the effect of subtitling on English-language skills may differ by age because of the effective differences in treatment dosages. On the one hand, the accumulation of intergenerational effects may lead to stronger effects in the younger population. On the other hand, the introduction of language choice options on DVDs and Blu-rays and the increasing prevalence of video streaming services such as Netflix and Amazon Prime could lead to weaker effects in the younger population. These options imply that the population in dubbing countries

can increasingly choose to watch movies in the original English language, an option likely more used by the young who tend to prefer subtitles over dubbing (European Union (2024)).[7]

Results suggest that neither of the potential age-differing impacts of subtitling is particularly driving overall English-language learning. The AES enables separate analyses for different age groups. The baseline analysis aggregated the full age range of 16-75 years. Columns 2-4 of Panel B of Table 2 report analyses for the separate age ranges of 16-34, 35-55, and 56-75 years. Results show significant effects of subtitling on the English skills in each of the age groups.[8] While the point estimate is larger in the youngest group, the estimates are not statistically significantly different between the three age groups.[9]

### 4.3 Subgroup Analyses by Gender and Parental Education

We find similar effects of subtitling on the English skills of women and men. Using the micro data of the AES and PISA studies allows us to distinguish subgroups by gender and parental education in each country.[10] Subgroup analyses indicate significant coefficients for females and males with very similar point estimates (Panel A of Table 3).

[7] On average across European countries, 65 percent of the age group 15-24, but only 44 percent of the age group 55 and older say they prefer to watch foreign films with subtitles rather than dubbed (European Union (2024), p. 52). Still, Netflix follows established translation mode traditions, dubbing most of its content in France, Germany, Italy, and Spain, but not in Scandinavia (Gruenwedel (2020)).

[8] Math skills refer to 15-year-olds in PISA in each of these analyses. Using PIAAC math skills instead allows to employ the same age groups for math than for English in the AES but substantially reduces the sample size to 18 countries. Still, qualitative results are the same: point estimates are also somewhat larger in the youngest age group but not significantly different from the older age groups.

[9] Relatedly, reaching back to 1991, the TOEFL tests allow for an analysis of subtitling effects over time which does not show a strong time-varying pattern (Appendix B.1).

[10] To match the age of PISA participants, we restrict the AES sample to 16- to 34-year-olds in these analyses.

**Table 3: The effect of subtitling on English skills by gender and parental education**

**Panel A: Effects by gender (AES)**

| | Female | Male |
|---|---|---|
| | (1) | (2) |
| Subtitles * English | 1.499*** | 1.514*** |
| | *(0.526)* | *(0.507)* |
| Subject fixed effects | Yes | Yes |
| Country fixed effects | Yes | Yes |
| Observations | 56 | 56 |
| Countries | 28 | 28 |

**Panel B: Effects by parental education (AES)**

| | No high school degree | High school degree | Tertiary degree |
|---|---|---|---|
| | (1) | (2) | (3) |
| Subtitles * English | 1.077** | 1.686*** | 2.031*** |
| | *(0.500)* | *(0.493)* | *(0.580)* |
| Subject fixed effects | Yes | Yes | Yes |
| Country fixed effects | Yes | Yes | Yes |
| Observations | 56 | 56 | 56 |
| Countries | 28 | 28 | 28 |

Notes: Least squares regressions. Dependent variable: test scores in English and math. Unit of observation: subject-specific country-average score of the specified subgroup. Subtitles: translation mode with 1 = subtitling, 0 = dubbing. Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: AES, PISA.

By contrast, the effect of subtitling differs strongly by parental education. Point estimates in Panel B of Table 3 increase significantly from participants whose parents have no high school degree (ISCED 1 or 2) over at least one has a high school degree (ISCED 3 or 4) to tertiary degree (ISCED 5 or above).[11,12] Given generally positive associations between parental

[11] Differences in coefficient estimates are statistically significant at the 5 percent level between columns 1 and 3 and between columns 1 and 2 and at the 10 percent level between columns 2 and 3.

[12] Similarly, Anghel, Cabrales, and Carro (2016) and Strazzeri, Oggenfuss, and Wolter (2022) find worse effects of English instruction in school for low-SES and low-track students.

education and children's own education and competencies, this heterogeneity may reflect selective foreign-content viewing by socioeconomic background, complementarities with other skill areas, or complementarities with English learning or usage beyond TV watching.[13]

## 5. Robustness Analysis

Our conclusions about the importance of subtitling prove highly robust to tests of the assumptions underlying our specification and to other variations of the baseline model.

***Effects on native-language reading and science as placebo tests.*** A natural placebo test considers the impact of subtitling on native-language reading (rather than English). The treatment effect of subtitling on PISA reading scores (where PISA math scores provide for the counterfactual) is not significant (Table 4, column 1). Similarly, subtitling does not affect science scores in our between-subject model (column 2).[14] These analyses directly address concerns that our subtitling results reflect general language aptitude or other family or schooling effects that would show up as subject-specific impacts which would compromise our between-subject identification.[15]

[13] In the smaller country sample that matches AES to PIAAC data, we find very similar education gradients by parental education and by individuals' own education. Similarly, point estimates are very similar in the subgroups of lower-educated participants with higher-educated parents and higher-educated participants with lower-educated parents, making it hard to distinguish between socioeconomic background or own education as dominant mechanism.

[14] This contrasts with English-instruction reforms in school for which negative effects on science skills have been shown (Anghel, Cabrales, and Carro (2016), Strazzeri, Oggenfuss, and Wolter (2022)).

[15] The lack of an effect on native-language reading and science skills also allows us to use them instead of math skills to produce alternative counterfactuals. Estimated effects on English skills are robust in these specifications, as well as to using adult numeracy skills to produce another alternative counterfactual (Appendix B.2).

**Table 4: Placebo tests: Native-language reading and science skills**

| | Reading (native language) | Science |
|---|---|---|
| | (1) | (2) |
| Subtitles * reading/science | -0.019 | -0.113 |
| | *(0.187)* | *(0.147)* |
| Subject fixed effects | Yes | Yes |
| Country fixed effects | Yes | Yes |
| Observations | 72 | 72 |
| Countries | 36 | 36 |

Notes: Least squares regressions. Dependent variable: test scores in reading/science and math. Unit of observation: subject-specific country score, standardized in joint country sample. Subtitles: translation mode with 1 = subtitling, 0 = dubbing. Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data source: PISA.

***Accounting for linguistic similarity, language size, and school instruction.*** The similarity of a country's native language to English does not affect the results. If language similarity related both to the historical decision to subtitle and to overall English skills, a bias would be introduced. The linguistic similarity score from Adserà and Pytliková (2015) is not, however, significantly related to the between-subject achievement difference, and the estimated effect of subtitling remains virtually unchanged when accounting for linguistic similarity (Table 5, column 2).

These inconsequential effects of native language similarity also hold more broadly for language families. Appendix Figure A2 replicates Figure 1 while splitting countries into four language families—Germanic, Romance, Slavic, and other (Katzner (2002)). Average achievement in both math and English clearly differs by language family. However, the practice of subtitling or dubbing systematically varies within each language family, and the treatment effect of our identification strategy—the distance from the 45-degree line—remains clearly visible within each language family.

**Table 5: Accounting for linguistic similarity, language size, and school instruction**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Subtitles * English | 1.450*** | 1.451*** | 1.241*** | 1.233*** | 1.385*** | 1.498*** | 1.153*** | 1.183** |
| | *(0.371)* | *(0.376)* | *(0.413)* | *(0.363)* | *(0.413)* | *(0.379)* | *(0.414)* | *(0.478)* |
| Linguistic similarity * English | | 0.380 | | | | | 0.509 | 2.342 |
| | | *(1.577)* | | | | | *(1.722)* | *(1.372)* |
| Population size * English | | | -0.009 | | | | -0.005 | -0.013 |
| | | | *(0.005)* | | | | *(0.005)* | *(0.010)* |
| Language size * English | | | | -0.003* | | | -0.002 | -0.001 |
| | | | | *(0.001)* | | | *(0.001)* | *(0.001)* |
| Starting age * English | | | | | 0.004 | | | -0.047 |
| | | | | | *(0.107)* | | | *(0.142)* |
| Share instruction time * English | | | | | | 0.063 | | 0.077 |
| | | | | | | *(0.102)* | | *(0.096)* |
| Subject fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 60 | 60 | 60 | 60 | 52 | 46 | 60 | 46 |
| Countries | 30 | 30 | 30 | 30 | 26 | 23 | 30 | 23 |

Notes: Least squares regressions. Dependent variable: test scores in English and math. Unit of observation: subject-specific country score, standardized in joint country sample. Subtitles: translation mode with 1 = subtitling, 0 = dubbing. Linguistic similarity: linguistic similarity score (Adserà and Pytliková (2015)). Population size: millions, averaged over 2010-2023. Language size: estimated number of first- and second-language speakers, upper bound, millions (Ethnologue (2023)). Starting age: age at which foreign languages are first taught at school (EACEA (2023)). Share instruction time: share of instruction time in secondary schooling assigned to foreign-language teaching (EACEA (2023)). Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: EPI, PISA.

Our results are also unaffected by differences in the size of a country's population or the worldwide number of speakers of its language. Population and language size may have affected a country's historical adoption of subtitling or dubbing as its primary translation mode by influencing whether dubbing was economically viable and by providing economic incentives to learn English more generally. Accounting for a country's population size (Table 5, column 3) or for the number of first- and second-language speakers of a country's language worldwide

(column 4) slightly lowers the estimated impact of subtitling but does not affect the overall qualitative conclusions about the mode of translation.[16]

Our results are also impervious to accounting for the intensiveness of foreign-language teaching in school. Bias would be introduced if TV translation modes were correlated with efforts at foreign-language instruction in school. Foreign-language instruction policy in schools could proxy a country's protectiveness of its native language while also being related to the choice of subtitling. The estimated effect of subtitling is qualitatively unaffected when accounting for the age at which foreign languages are first taught at school (Table 5, column 5) or for the share of instruction time in secondary schooling that is assigned to foreign-language teaching (column 6).[17]

Even when combined, these potential threats to identification show no significant impact on our subtitling conclusions. Results are qualitatively unaffected when simultaneously accounting for the language and country characteristics that are available in the full country sample (column 7) or additionally for school policy in the slightly reduced sample (column 8).

The finding that none of the interactions with the other variables is significant may raise concerns about the informational content of these variables. However, linguistic similarity is in fact strongly positively correlated with English skills, and population and language size are negatively correlated with English skills (Appendix Table A4, column 1). But linguistic similarity is not correlated with subtitling (column 2). As expected, population and language size

[16] In fact, the effect of subtitling is highly robust even when dropping all countries with populations of over 20 million (such as Germany, France, Italy, and Spain) from the analysis altogether (coefficient 1.215, std. err. 0.458).

[17] When adding triple interactions between subtitling, English, and the schooling variables, we find insignificant triple interaction terms. Point estimates tend to be larger in countries with longer foreign-language instruction time, though, tentatively indicating complementarity rather than substitutability between out-of-school and school learning. However, the triple-interaction specification is rather demanding in our setting of limited country observations. There are also no significant triple interactions between subtitling and the time spent watching TV in a country as a proxy for treatment intensity.

are indeed negatively correlated with subtitling (column 2), and their interactions with English in the Table 5 regression are indeed significantly negative when excluding the interaction of subtitles with English from the model (column 3 of Table A4). The finding that the estimates turn insignificant when subtitles are included indicates that this result is fully driven by their association with subtitles. Among the schooling variables, only instruction time shows a negative association with subtitles (column 2), suggesting that dubbing countries may partially compensate by additional foreign-language instruction.

***Accounting for selectivity of test taking.*** As discussed, EPI and TOEFL scores do not come from representative country samples, opening the possibility of selectivity of test taking. The concern is that differential selectivity of test taking between subtitling and dubbing countries could introduce bias. For example, if the population in subtitling countries has higher English proficiency, they may feel more confident taking the test. On the other hand, they may also see less need to signal sufficient proficiency, indicating that the bias could go in either direction. The similar impact for the representative AES scores indicates that overall results are not driven by selective test taking, but additional analysis of selectivity of TOEFL is possible.

We can look at differential selectivity by using data on test participation rates by country available for the paper- and computer-based TOEFL tests.[18] We find that the number of test takers per 1000 inhabitants neither significantly affects English performance nor substantially alters our estimates of the impact of subtitling (Table 6, column 1). In the listening domain, the point estimate of 1.36 SD (column 2) is roughly as large as the baseline EPI and AES estimates in Table 1.

[18] Because the speaking domain is only tested in the internet-based version, we cannot implement this specification for the speaking domain. Unfortunately, our request for data on the number of test takers in the internet-based version was declined by Educational Testing Service, the provider of TOEFL. Similarly, Education First did not respond to our request for data on the number of national test takers on EPI.

**Table 6: Accounting for selectivity of test taking (TOEFL)**

| | Total | Listening | Writing | Reading |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Subtitles * English | 1.076** | 1.361** | 0.799 | 0.689 |
| | *(0.471)* | *(0.515)* | *(0.533)* | *(0.501)* |
| Participation * English | -0.320 | -0.288 | -0.100 | -0.668 |
| | *(0.710)* | *(0.708)* | *(0.805)* | *(0.757)* |
| Subject fixed effects | Yes | Yes | Yes | Yes |
| Country fixed effects | Yes | Yes | Yes | Yes |
| Observations | 62 | 62 | 62 | 62 |
| Countries | 31 | 31 | 31 | 31 |

Notes: Least squares regressions. Dependent variable: test scores in English and math. Unit of observation: subject-specific country score, standardized in joint country sample. Subtitles: translation mode with 1 = subtitling, 0 = dubbing. Participation: number of test takers per 1000 inhabitants, averaged across all waves. Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: TOEFL PBT and CBT, PISA.

## 6. Conclusions

The learning of basic skills is generally studied in school contexts even though substantive basic skill acquisition happens outside of schools. Importantly, acquiring foreign-language skills is significantly affected by the extent to which people are exposed to the foreign language in their everyday lives. A primary component of foreign-language exposure is watching TV—at least if foreign-language content is shown in the original language. European TV employs considerable content originally produced in English, but not all viewers will be exposed to English language through this avenue. If the content is dubbed into their native language, the English language component is completely obscured. This contrasts sharply with content broadcast in English with translation provided in subtitles.

We study the effect of subtitling vs. dubbing on the learning of English as a foreign language. Our cross-country between-subject approach compares English to math proficiency in

subtitling and dubbing countries, eliminating potential bias from overall skill and cultural differences across countries.

Results indicate a very strong positive effect of subtitling on the English proficiency of the population. Effects are huge, surpassing one SD in most specifications. Consistent with learning from oral transmission of TV content, effects are strongest for speaking and listening skills and weakest for reading skills.

The estimates are best interpreted as long-run effects that capture not only impacts of individual TV viewing but also intergenerational effects running through improved English skills of parents and teachers. Interestingly, however, our results do not indicate strong differences in patterns across age or time. This might reflect countervailing factors such as the expansion of alternative media transmission channels such as DVDs and streaming services that can increasingly expose people in dubbing countries with material in English.[19] Thus, how our results extrapolate into the future constitutes an interesting question for future research.

The fact that all European countries teach English in schools indicates that English language acquisition is an important policy goal. English has established itself as the lingua franca in large parts of international communication in business, science, and digital technologies. Our results do have direct policy implications for non-English-speaking countries that aim to strengthen their populations' skills in English. Simply put, adopting the practice of movie dubbing has major consequences for children's acquisition of foreign-language skills.

Going beyond the specific issue of language acquisition, this analysis shows the powerful impact of non-school factors on learning. While the very earliest estimates of educational

[19] For example, recent evidence suggests a positive trend in the English proficiency of German ninth-graders over the past decade (Stanat et al. (2023)), against a strongly declining trend in other subjects such as German, math, and science. The diverging trend might reflect increased exposure to English content on social media and other online sources, particularly since the experiences of school closures during Covid-19.

production functions emphasized the importance of families and peers on learning (Coleman et al. (1966)), little subsequent work has delved into mechanisms underlying these effects or has estimated well-identified elements of these impacts. This analysis shows that non-school factors can have significant impacts on the differences in achievement and skills across countries. It also points to the fact that thinking about educational policies might usefully go beyond the schools of a country to consider broader sites of learning.

## References


Adserà, Alícia, Mariola Pytliková (2015). The Role of Language in Shaping International Migration. *Economic Journal* 125 (586): F49-F81.

Almeida, Patrícia Albergaria, Patrícia Dinis Costa (2014). Foreign Language Acquisition: The Role of Subtitling. *Procedia - Social and Behavioral Sciences* 141: 1234-1238.

Anghel, Brindusa, Antonio Cabrales, Jesus M. Carro (2016). Evaluating a Bilingual Education Program in Spain: The Impact Beyond Foreign Language Learning. *Economic Inquiry* 54 (2): 1202-1223.

Aparicio-Fenoll, Ainoa, Antonio Di Paolo (2023). Language Economics. In *Handbook of Labor, Human Resources and Population Economics*, edited by Klaus F. Zimmermann. Cham: Springer.

Bursztyn, Leonardo, Davide Cantoni (2016). A Tear in the Iron Curtain: The Impact of Western Television on Consumption Behavior. *Review of Economics and Statistics* 98 (1): 25-41.

Campante, Filipe, Ruben Durante, Andrea Tesei (2022). Media and Social Capital. *Annual Review of Economics* 14: 69-91.

Chaume, Frederic (2012). *Audiovisual Translation: Dubbing*. London: Routledge.

Chiswick, Barry R., Paul W. Miller (2015). International Migration and the Economics of Language. In *Handbook of the Economics of International Migration*, edited by Barry R. Chiswick, Paul W. Miller. Amsterdam: North-Holland: 211-269.

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, Robert L. York (1966). *Equality of Educational Opportunity*. Washington, D.C.: U.S. Government Printing Office.

Crystal, David (2003). *English as a Global Language*. 2nd ed. Cambridge, UK: Cambridge University Press.

d'Ydewalle, Géry, Marijke Van de Poel (1999). Incidental Foreign-Language Acquisition by Children Watching Subtitled Television Programs. *Journal of Psycholinguistic Research* 28 (3): 227-244.

Danan, Martine (1991). Dubbing as an Expression of Nationalism. *Meta* 36 (4): 606-614.

de Swaan, Abram (2001). *Words of the World: The Global Language System*. Cambridge, UK: Polity Press.

Durante, Ruben, Paolo Pinotti, Andrea Tesei (2019). The Political Legacy of Entertainment TV. *American Economic Review* 109 (7): 2497-2530.

EACEA (2023). *Key Data on Teaching Languages at School in Europe – 2023 Edition, Eurydice Report*. European Education and Culture Executive Agency. Luxembourg: Publications Office of the European Union.

EF Education First (2011). *EF EPI - EF EPI English Proficiency Index*. https://www.ef.com/assetscdn/WIBIwq6RdJvcD9bc8RMd/cefcom-epi-site/reports/2011/ef-epi-2011-english.pdf (accessed 1 February 2025).

EF Education First (2024). *EF EPI - EF English Proficiency Index: A Ranking of 116 Countries and Regions by English Skills*.

https://www.ef.com/assetscdn/WIBIwq6RdJvcD9bc8RMd/cefcom-epi-site/reports/2024/ef-epi-2024-english.pdf (accessed 16 February 2025).

Ellis, Rod (2008). *The Study of Second Language Acquisition. Second Edition*. Oxford: Oxford University Press.

Ethnologue (2023). *Ethnologue 200: List of the Top 200 Languages in the World*. Cited according to https://en.wikipedia.org/wiki/List_of_languages_by_total_number_of_speakers (accessed 1 November 2023).

ETS (2024). *TOEFL iBT: Test and Score Data Summary 2023*. https://www.ets.org/pdfs/toefl/toefl-ibt-test-score-data-summary-2023.pdf (accessed 18 February 2025). Princeton, NJ: Educational Testing Service.

European Union (2024). *Europeans and Their Languages. Special Eurobarometer 540*. https://europa.eu/eurobarometer/surveys/detail/2979 (accessed 21 February 2025).

Eurostat (2024). *Adult Education Survey*. Reference Metadata. https://ec.europa.eu/eurostat/cache/metadata/en/trng_aes_12m_esms.htm (accessed 14 February 2025).

Eurostat (2025). *Education and Training Database*. https://ec.europa.eu/eurostat/web/education-and-training/database (accessed 16 February 2025).

Fidrmuc, Jan, Jarko Fidrmuc (2016). Foreign Languages and Trade: Evidence from a Natural Experiment. *Empirical Economics* 50 (1): 31-49.

Gazzola, Michele, François Grin, Bengt-Arne Wickström (2016). A Concise Bibliography of Language Economics. In *The Economics of Language Policy*, edited by Michele Gazzola, Bengt-Arne Wickström. Cambridge, MA: MIT Press: 53-92.

Gentzkow, Matthew, Jesse M. Shapiro (2008). Preschool Television Viewing and Adolescent Test Scores: Historical Evidence from the Coleman Study. *Quarterly Journal of Economics* 123 (1): 279-323.

Ginsburgh, Victor A., Juan Prieto-Rodriguez (2011). Returns to Foreign Languages of Native Workers in the European Union. *ILR Review* 64 (3): 599-618.

Ginsburgh, Victor, Shlomo Weber (2020). The Economics of Language. *Journal of Economic Literature* 58 (2): 348-404.

Gottlieb, Henrik (1994). Subtitling: Diagonal Translation. *Perspectives: Studies in Translation Theory and Practice* 2 (1): 101-121.

Gruenwedel, Erik (2020). *Netflix Primarily Dubs Content in Germany, France and Japan*. https://www.mediaplaynews.com/ampere-netflix-primarily-dubs-content-in-germany-france-and-japan/ (accessed 17 June 2025).

Hahm, Sabrina, Michele Gazzola (2022). The Value of Foreign Language Skills in the German Labor Market. *Labour Economics* 76: 102150.

Hanushek, Eric A. (1986). The Economics of Schooling: Production and Efficiency in Public Schools. *Journal of Economic Literature* 24 (3): 1141-1177.

Hanushek, Eric A., Lavinia Kinne, Philipp Lergetporer, Ludger Woessmann (2022). Patience, Risk-Taking, and Human Capital Investment across Countries. *Economic Journal* 132 (646): 2290-2307.

Hanushek, Eric A., Ludger Woessmann (2011). The Economics of International Differences in Educational Achievement. In *Handbook of the Economics of Education, Vol. 3*, edited by Eric A. Hanushek, Stephen Machin, Ludger Woessmann. Amsterdam: North Holland: 89-200.

Ivarsson, Jan (2009). The History of Subtitles in Europe. In *Dubbing and Subtitling in a World Context*, edited by Gilbert C.F. Fong, Kenneth K.L. Au. Hong Kong: Chinese University Press: 3-12.

Katzner, Kenneth (2002). *The Languages of the World*. 3rd ed. London: Routledge.

Kearney, Melissa S., Phillip B. Levine (2019). Early Childhood Education by Television: Lessons from Sesame Street. *American Economic Journal: Applied Economics* 11 (1): 318-350.

Koolstra, Cees M., Jonannes W. J. Beentjes (1999). Children's Vocabulary Acquisition in a Foreign Language through Watching Subtitled Television Programs at Home. *Educational Technology Research and Development* 47 (1): 51-60.

Modot, Alain, Thierry Baujard, Silvia Angrisani, et al. (2007). *Study on Dubbing and Subtitling Needs and Practices in the European Audiovisual Industry*. Study initiated by the European Commission. Paris: Media Consulting Group.

Piopiunik, Marc, Guido Schwerdt, Lisa Simon, Ludger Woessmann (2020). Skills, Signals, and Employability: An Experimental Investigation. *European Economic Review* 123: 103374.

Reinders, Hayo, Phil Benson (2017). Research Agenda: Language Learning Beyond the Classroom. *Language Teaching* 50 (4): 561-578.

RTL AdAlliance (2024). *TV Key Facts*. Luxembourg.

Rupérez Micola, Augusto, Ainoa Aparicio Fenoll, Albert Banal-Estañol, Arturo Bris (2019). TV or Not TV? The Impact of Subtitling on English Skills. *Journal of Economic Behavior & Organization* 158: 487-499.

Safar, Hayssam, Alain Modot, Silvia Angrisani, Yves Gambier, Carlo Eugeni, Héloïse Fontanel, Najwa Hamaoui, Xavier Verstrepen (2011). *Study on the Use of Subtitling: The Potential of Subtitling to Encourage Foreign Language Learning and Improve the Mastery of Foreign Languages*. Study requested by the European Commission, Directorate-General Education and Culture. Paris: Media Consulting Group.

Stanat, Petra, Stefan Schipolowski, Rebecca Schneider, Sebastian Weirich, Sofie Henschel, Karoline A. Sachse (2023). *IQB-Bildungstrend 2022: Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im dritten Ländervergleich*. Münster: Waxmann.

Strazzeri, Maurizio, Chantal Oggenfuss, Stefan C. Wolter (2022). Much Ado About Nothing? School Curriculum Reforms and Students' Educational Trajectories. CESifo Working Paper 9912. Munich: CESifo.

Van Lommel, Sven, Annouschka Laenen, Géry d'Ydewalle (2006). Foreign-Grammar Acquisition While Watching Subtitled Television Programmes. *British Journal of Educational Psychology* 76 (2): 243-258.

Woessmann, Ludger (2016). The Importance of School Systems: Evidence from International Differences in Student Achievement. *Journal of Economic Perspectives* 30 (3): 3-32.

# Appendix A: Learning Assessment Data

## A.1 The English Proficiency Index (EPI)

The English Proficiency Index (EPI) has been developed by Education First (EF), an international education company that focuses on language, academics, cultural exchange, and educational travel. The goal of EPI is to "measure the average English competency of the working population" (EF Education First (2011)) based on adaptive online tests of English reading and listening skills. The tests can be completed on computers as well as mobile devices.

EPI results are available on an annual basis from 2011 to 2024.[20] EPI has steadily increased its country coverage from 44 in 2011 to 116 in 2024, and the number of test takers from roughly 1 million to over 2 million since 2019.[21] From the 36 non-English-speaking European countries participating in PISA, our EPI analysis excludes the small countries of Iceland, Kosovo, Montenegro, and North Macedonia that are not ranked in EPI and the two Belgian regions that differ in translation mode, leaving us with 30 countries (Table A1).

A concern about EPI is the self-selection of its test takers. EPI is a freely available online test usually taken by people interested in learning about their English proficiency. As an online test, it naturally excludes people without internet access. In addressing possible concerns about selection, Education First reports that its sample is roughly gender balanced, that the ranking includes only countries with a minimum of 400 test takers, and that EPI scores are highly correlated with TOEFL scores. According to Education First, there is also no incentive for cheating or cramming in the low-stakes test environment. Nevertheless, test takers are arguably

[20] Country scores were first calculated as a percentage of correct answers, then moved to a 100-point scale in 2018. Starting in 2020, EPI is measured on an 800-point scale and includes scores from previous years since 2019 to stabilize scores over time. We standardize all scores in our empirical analysis.

[21] The first published editions in 2011 and 2012 are based on test data from 2007-2009 and 2009-2011, respectively, whereas all following EPI editions are based on test data from the immediately prior year.

not representative of a country's adult population but rather positively selected in terms of educational attainment and English knowledge. Test takers are almost all working adults or college students with a median age of 26 in 2024 (EF Education First (2024)). This selection could introduce bias into our analysis if it were systematically different between subtitling and dubbing countries—e.g., if the test were more wide-spread in one group of countries. However, our robustness analysis indicates no impact on our results of selective test taking.

### A.2 Test of English as a Foreign Language (TOEFL)

An alternative measure of English skills are the average country scores from the Test of English as a Foreign Language (TOEFL). Administered by the Educational Testing Service (ETS), TOEFL is widely accepted by universities worldwide as proof of English proficiency (ETS (2024)). We use all annual releases available from 1991 to 2023.

The format of TOEFL has changed over time. From 1991-1997, the test was paper-based (PBT) with reading comprehension, listening comprehension, structure, and written expression domains. In 1998, a computer-based version (CBT) replaced the PBT in our sample of European countries. In 2005, this was replaced by an internet-based version (iBT) that also added a speaking section. The tests are extensive, with TOEFL iBT taking four hours.

TOEFL reports country averages if at least 30 test takers participated. In total, the TOEFL test was taken by about 600,000-700,000 people every year between 1991 and 2006 (no data on test takers has been released since). The number of test takers across countries varies greatly, e.g., on the TOEFL CBT 2001-2002 from 20 test takers in Finland to over 100,000 in China. TOEFL iBT is available for 34 non-English-speaking European countries, while TOEFL PBT and CBT are available for 31 countries with Kosovo, Montenegro, and Serbia missing.

Because TOEFL is most commonly taken by students applying for English-language university programs, sample selection concerns also apply to TOEFL. However, the available data on the number of test takers as a percentage of a country's population allows us to test implications of overall selectivity, again suggesting that selection does not affect our results.

### A.3 Adult Education Survey (AES)

The Adult Education Survey (AES) by Eurostat covers representative country samples and provides self-rated language skills (Eurostat (2024)). After being piloted in 2007, AES has been conducted in 2011, 2016, and 2022. It targets random samples of 2000-3000 respondents of a country's resident population in slightly varying age ranges (e.g., 25-64 in 2007, 16-75 in 2011). Participation increased from 26 countries in the pilot study to 33 in 2016. Our analysis uses the 28 non-English-speaking AES countries that also have PISA data. Because regional identifiers are available in the AES micro data, we can include the Belgian regions of Flanders and Wallonia (with different translation modes) separately in the analysis.

In AES, respondents list up to seven languages that they can use and identify the two foreign languages they know best (Eurostat (2024)). They rate their proficiency in these languages on a four-point scale as very basic (only understand a few words), basic, good, or proficient (mastered the language almost completely), and we use the average self-rated English proficiency.

In contrast to EPI and TOEFL, AES offers a representative sample of a country's resident adult population, allowing us to assess the relevance of selectivity of the other two. On the other hand, AES provides only self-rated foreign-language proficiencies and does not elicit them with a standardized test. This could introduce bias if people from dubbing or subtitling countries systematically either over- or underrate their English proficiency.

## A.4 Programme for International Student Assessment (PISA) and Programme of the International Assessment of Adult Competencies (PIAAC)

Math skills are measured using data from the Programme for International Student Assessment (PISA) conducted by the Organization for Economic Co-operation and Development (OECD). PISA collects internationally comparable data on the performance of representative samples of 15-year-old students with a target sample size per country of at least 5000 students. PISA has been administered every three years since 2000, and our analysis employs all eight waves from 2000 to 2022. Through 2022, 94 countries participated in PISA, and we focus on the sample of 36 European countries that do not have English among their official languages. In 2015, PISA switched from paper-based to computer-based assessment in most participating countries. Tests covering math, reading, and science last about two hours.

The placebo analysis uses the PISA scores in native-language reading and science that were administered in conjunction with the PISA math tests.

We further employ numeracy scores from an adult skill test, the Programme of the International Assessment of Adult Competencies (PIAAC), also conducted by the OECD. PIAAC tests representative samples of at least 5000 participants covering the population aged 16-65. Countries participated in different waves between 2011 and 2018.

# Appendix B: Additional Analyses

## B.1 Subtitling Effects over Time

The TOEFL tests allow for an analysis of subtitling effects over time. TOEFL provides information on international achievement since 1991. However, the changing format of the TOEFL test—from paper-based (PBT) to computer-based (CBT) to internet-based (iBT)—restricts direct comparability to three separate subperiods. We split each of these subperiods in two: 1991-1993 and 1995-1997 for PBT, 1998-2001 and 2002-2005 for CBT, and 2006-2014 and 2015-2023 for iBT.

Results in Table A5 show that changes over time are small, inconsistent, and not statistically significant. Within each of the three testing subperiods, changes are very limited and do not show a clear pattern (slightly down in PBT and iBT, more strongly up in CBT). Overall, the subtitling effect is consistent across all subperiods, with no overall pattern over time.[22]

## B.2 Alternative Counterfactuals: Adult Numeracy and Student Reading and Science Skills

Potential measurement problems related to the age of the tested individuals in the different subjects do not affect our results. While the median age of participants in the EPI test is 26, PISA assesses 15-year-olds. Although available for a smaller country sample, PIAAC provides an assessment of numeracy skills for a sample of young adults aged 16-24. When using PIAAC numeracy instead of PISA math to form the counterfactual, the point estimate is very similar and statistically significant in the smaller country sample (Table A6, column 1).

If the driving force was general language competency and if this was related to subtitling, we could again have biased estimates of subtitling—but this is not the case. To test this, we use

[22] The TOEFL iBT estimate is somewhat larger, at 0.979 (standard error 0.381), when estimated on the same smaller sample of countries available in the TOEFL PBT and CBT tests.

reading skills in the native language instead of math skills to provide the counterfactual for English skills in our analysis. Subtitling estimates are very similar when using PISA reading scores instead of PISA math scores for the counterfactual (Table A6, column 2). The same is true when using science skills to produce another alternative counterfactual (column 3).

## Appendix C: Additional Figures and Tables

**Figure A1: TV translation mode in European countries**

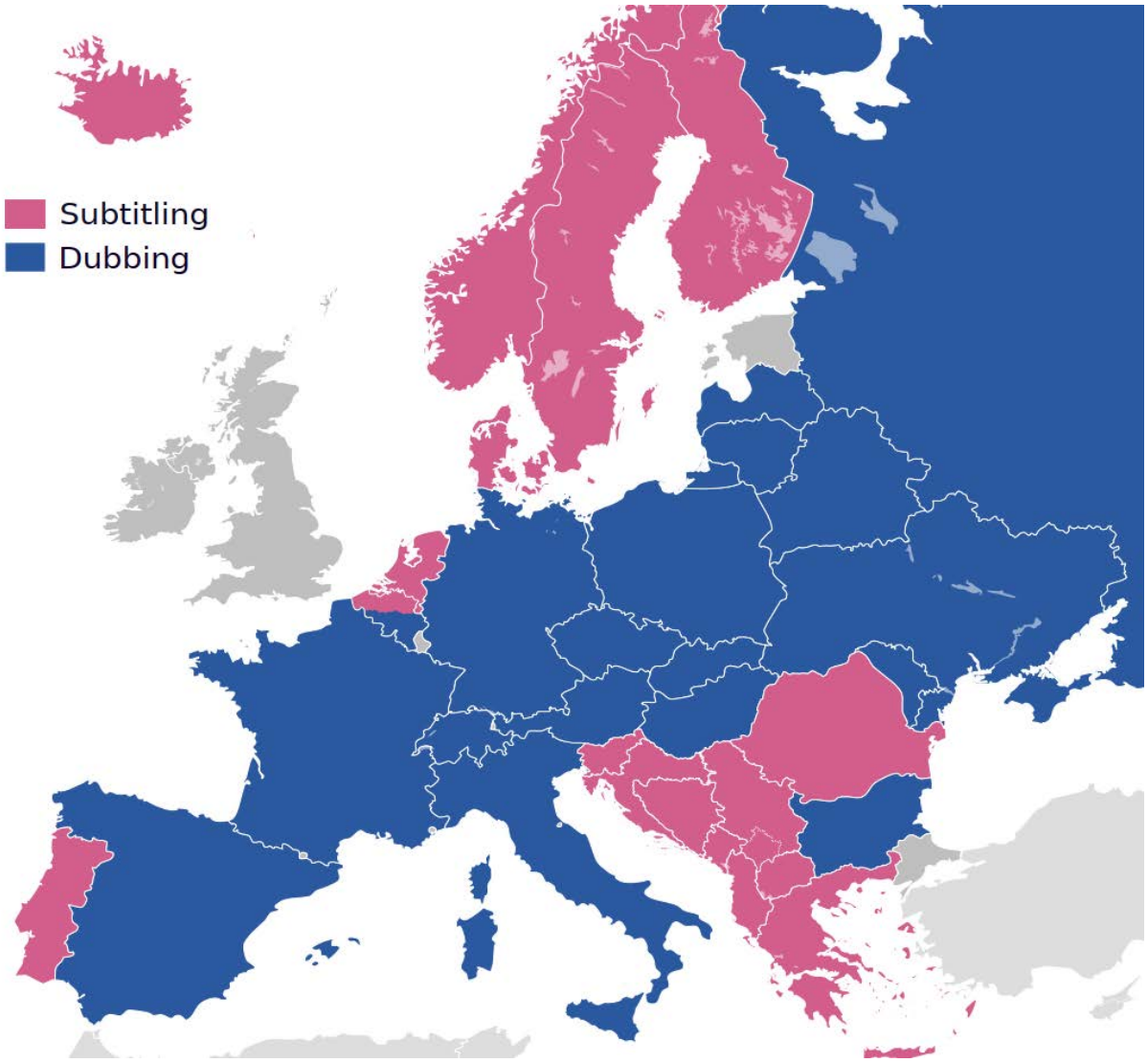
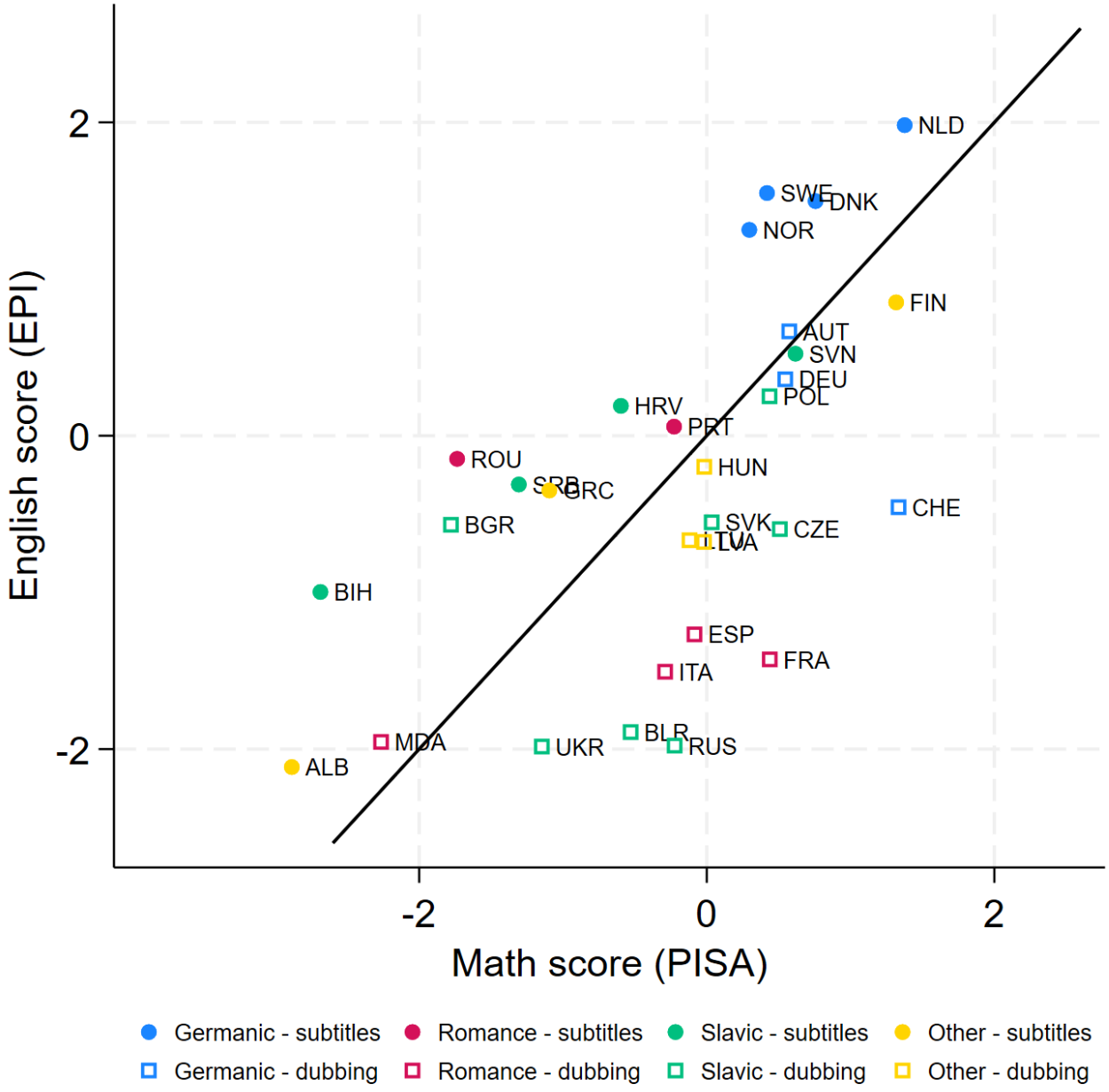


**Figure A2: English and math achievement by translation mode and language family**




Notes: Test scores in English and math, country average across all waves, standardized in joint country sample. See Appendix Table A1 for country abbreviations. Data sources: EPI, PISA.

**Table A1: Descriptive statistics**

| | Subtitles | EPI | TOEFL | AES | PISA |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Albania (ALB) | 1 | -2.115 | -2.618 | | -2.886 |
| Austria (AUT) | 0 | 0.666 | 1.520 | 0.335 | 0.574 |
| Belarus (BLR) | 0 | -1.891 | -1.324 | | -0.529 |
| Belgium: Flanders (BE-VLG) | 1 | | | 0.431 | 1.392 |
| Belgium: Wallonia (BE-WAL) | 0 | | | -0.509 | 0.183 |
| Bosnia and Herzegovina (BIH) | 1 | -0.997 | -1.511 | -1.748 | -2.686 |
| Bulgaria (BGR) | 0 | -0.568 | -0.362 | -0.127 | -1.779 |
| Croatia (HRV) | 1 | 0.190 | -0.107 | 0.060 | -0.597 |
| Czech Republic (CZE) | 0 | -0.596 | -0.431 | -0.727 | 0.508 |
| Denmark (DNK) | 1 | 1.497 | 1.559 | 1.675 | 0.757 |
| Finland (FIN) | 1 | 0.851 | 0.919 | 0.806 | 1.317 |
| France (FRA) | 0 | -1.428 | -0.996 | -1.402 | 0.438 |
| Germany (DEU) | 0 | 0.360 | 1.106 | -0.227 | 0.546 |
| Greece (GRC) | 1 | -0.349 | -0.671 | 0.022 | -1.095 |
| Hungary (HUN) | 0 | -0.198 | -0.331 | -0.105 | -0.017 |
| Iceland (ISL) | 1 | | 0.672 | | 0.435 |
| Italy (ITA) | 0 | -1.506 | -1.244 | -1.480 | -0.290 |
| Kosovo (KSV) | 1 | | -3.496 | | -3.953 |
| Latvia (LVA) | 0 | -0.678 | -1.048 | -0.407 | -0.019 |
| Lithuania (LTU) | 0 | -0.667 | -1.020 | -0.681 | -0.120 |
| Moldova (MDA) | 0 | -1.954 | -1.380 | | -2.265 |
| Montenegro (MNE) | 1 | | -2.367 | | -2.379 |
| Netherlands (NLD) | 1 | 1.982 | 1.831 | 0.883 | 1.376 |
| North Macedonia (MKD) | 1 | | -1.579 | 0.470 | -3.218 |
| Norway (NOR) | 1 | 1.314 | 0.414 | 1.896 | 0.296 |
| Poland (POL) | 0 | 0.252 | -0.878 | -0.619 | 0.437 |
| Portugal (PRT) | 1 | 0.058 | 0.436 | -0.428 | -0.227 |
| Romania (ROU) | 1 | -0.148 | 0.103 | -0.705 | -1.736 |
| Russian Federation (RUS) | 0 | -1.978 | -1.380 | | -0.223 |
| Serbia (SRB) | 1 | -0.311 | -1.248 | 0.240 | -1.307 |
| Slovak Republic (SVK) | 0 | -0.553 | -0.535 | 0.160 | 0.034 |
| Slovenia (SVN) | 1 | 0.523 | 0.699 | 1.299 | 0.617 |
| Spain (ESP) | 0 | -1.268 | -0.777 | -0.633 | -0.086 |
| Sweden (SWE) | 1 | 1.549 | 0.381 | 1.821 | 0.419 |
| Switzerland (CHE) | 0 | -0.456 | 0.940 | 0.333 | 1.334 |
| Ukraine (UKR) | 0 | -1.984 | -1.684 | | -1.147 |
| Number of countries | 36 | 30 | 34 | 28 | 36 |

Notes: Subtitles: translation mode with 1 = subtitling, 0 = dubbing. EPI, TOEFL: English test scores. AES: self-rated English proficiency. PISA: math test score. Country averages across all waves, standardized in joint country sample. Data sources: EPI, TOEFL, AES, PISA.

**Table A2: Correlations of test scores**

| | EPI | TOEFL | AES |
|---|---|---|---|
| | (1) | (2) | (3) |
| TOEFL | 0.856 | | |
| | *(0.000)* | | |
| | [30] | | |
| AES | 0.846 | 0.619 | |
| | *(0.000)* | *(0.001)* | |
| | [25] | [26] | |
| PISA | 0.590 | 0.790 | 0.354 |
| | *(0.001)* | *(0.000)* | *(0.065)* |
| | [30] | [34] | [28] |

Notes: Pair-wise correlations. *p*-values in parentheses. Number of observations in brackets. EPI, TOEFL: English test scores. AES: Self-rated English proficiency. PISA: math test score. Country averages across all waves, standardized in joint country sample. Data sources: EPI, TOEFL, AES, PISA.

**Table A3: The effect of subtitling on English skills: Constant country sample**

| | EPI | TOEFL | AES |
|---|---|---|---|
| | (1) | (2) | (3) |
| Subtitles * English | 1.383*** | 0.905** | 1.273** |
| | *(0.418)* | *(0.428)* | *(0.456)* |
| Subject fixed effects | Yes | Yes | Yes |
| Country fixed effects | Yes | Yes | Yes |
| Observations | 50 | 50 | 50 |
| Countries | 25 | 25 | 25 |

Notes: Least squares regressions. Dependent variable: test scores in English and math. Unit of observation: subject-specific country score, standardized in joint country sample. Subtitles: translation mode with 1 = subtitling, 0 = dubbing. Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: EPI, TOEFL, AES, PISA.

**Table A4: Analysis of linguistic similarity, language size, and school instruction**

| | Correlation | | Coefficient on interaction with English in Table 5 regression excluding subtitles |
|---|---|---|---|
| | with EPI English score | with subtitles | |
| | (1) | (2) | (3) |
| Linguistic similarity | 0.482*** | -0.008 | 0.336 |
| | *(0.007)* | *(0.967)* | *(1.976)* |
| Population size | -0.328* | -0.390** | -0.016*** |
| | *(0.077)* | *(0.033)* | *(0.005)* |
| Language size | -0.329* | -0.332* | -0.004** |
| | *(0.076)* | *(0.074)* | *(0.002)* |
| Starting age | 0.183 | -0.054 | -0.021 |
| | *(0.372)* | *(0.795)* | *(0.150)* |
| Share instruction time | -0.083 | -0.400* | -0.035 |
| | *(0.708)* | *(0.059)* | *(0.115)* |

Notes: Columns 1 and 2: correlation coefficient; *p*-value in parentheses. Column 3: Each entry refers to one least squares regression. Dependent variable: test scores in English and math. Unit of observation: subject-specific country score, standardized in joint country sample. See notes to Table 5 for details. Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: EPI, PISA.

**Table A5: The effect of subtitling on English skills over time (TOEFL)**

**Panel A: TOEFL PBT (1991-1997)**

| | Total | 1991-1993 | 1995-1997 |
|---|---|---|---|
| | (1) | (2) | (3) |
| Subtitles * English | 1.085** | 1.138** | 1.050** |
| | *(0.425)* | *(0.443)* | *(0.423)* |
| Subject fixed effects | Yes | Yes | Yes |
| Country fixed effects | Yes | Yes | Yes |
| Observations | 62 | 62 | 62 |
| Countries | 31 | 31 | 31 |

**Panel B: TOEFL CBT (1998-2005)**

| | Total | 1998-2001 | 2002-2005 |
|---|---|---|---|
| | (1) | (2) | (3) |
| Subtitles * English | 1.001** | 0.894* | 1.079** |
| | *(0.428)* | *(0.440)* | *(0.436)* |
| Subject fixed effects | Yes | Yes | Yes |
| Country fixed effects | Yes | Yes | Yes |
| Observations | 62 | 62 | 62 |
| Countries | 31 | 31 | 31 |

**Panel C: TOEFL iBT (2006-2023)**

| | Total | 2006-2014 | 2015-2023 |
|---|---|---|---|
| | (1) | (2) | (3) |
| Subtitles * English | 0.880** | 0.887** | 0.866** |
| | *(0.364)* | *(0.344)* | *(0.421)* |
| Subject fixed effects | Yes | Yes | Yes |
| Country fixed effects | Yes | Yes | Yes |
| Observations | 68 | 68 | 68 |
| Countries | 34 | 34 | 34 |

Notes: Least squares regressions. Dependent variable: test scores in English and math. Unit of observation: subject-specific country score, standardized in joint country sample. Subtitles: translation mode with 1 = subtitling, 0 = dubbing. Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: TOEFL, PISA.

**Table A6: Alternative counterfactuals: Adult numeracy and student reading and science skills**

| | PIAAC numeracy (age 16-24) | PISA reading | PISA science |
|---|---|---|---|
| | (1) | (2) | (3) |
| Subtitles * English | 1.335** | 1.346*** | 1.522*** |
| | *(0.524)* | *(0.440)* | *(0.429)* |
| Subject fixed effects | Yes | Yes | Yes |
| Country fixed effects | Yes | Yes | Yes |
| Observations | 36 | 60 | 60 |
| Countries | 18 | 30 | 30 |

Notes: Least squares regressions. Dependent variable: test scores in English and numeracy/reading/science. Unit of observation: subject-specific country score, standardized in joint country sample. Subtitles: translation mode with 1 = subtitling, 0 = dubbing. Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: EPI, PIAAC, PISA.