



The dynamics of recent refugees' language acquisition: how do their pathways compare to those of other new immigrants?

Yuliya Kosyakova, Cornelia Kristen & Christoph Spörlein

To cite this article: Yuliya Kosyakova, Cornelia Kristen & Christoph Spörlein (2022) The dynamics of recent refugees' language acquisition: how do their pathways compare to those of other new immigrants?, Journal of Ethnic and Migration Studies, 48:5, 989-1012, DOI: [10.1080/1369183X.2021.1988845](https://doi.org/10.1080/1369183X.2021.1988845)

To link to this article: <https://doi.org/10.1080/1369183X.2021.1988845>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 12 Oct 2021.



[Submit your article to this journal](#)



Article views: 10596



[View related articles](#)



[View Crossmark data](#)



Citing articles: 32 [View citing articles](#)

The dynamics of recent refugees' language acquisition: how do their pathways compare to those of other new immigrants?

Yuliya Kosyakova ^{a,b}, Cornelia Kristen ^b and Christoph Spörlein ^c

^aInstitute for Employment Research (IAB), Nuremberg, Germany; ^bUniversity of Bamberg, Bamberg, Germany; ^cUniversity of Dusseldorf, Dusseldorf, Germany

ABSTRACT

Do the processes underlying destination-language acquisition differ between recently arrived refugees and other new immigrants? Based on a well-established model of language learning according to which language fluency is a function of efficiency, incentives, and exposure, this study addresses general processes of language learning as well as conditions specific to refugees. Longitudinal data on refugees (IAB-BAMF-SOEP Sample of Refugees in Germany) and other immigrants (IAB-SOEP Migration Sample) in Germany indicate that exposure to the destination language is the main driver of proficiency, followed by efficiency, whereas incentives matter less. Moreover, refugees profit substantially from structured learning in language courses, while other immigrants benefit more from informal exposure in their everyday environments. Overall, the findings suggest that language acquisition is a general process that is similar across different types of immigrants.

ARTICLE HISTORY

Received 27 May 2021
Accepted 29 September 2021

KEYWORDS


Destination-language proficiency; language acquisition; refugees; new immigrants; Germany

Introduction

Humanitarian migration surged in the past decade, driven by war, persecution and oppression (UNHCR 2020). The major driving force of this increase is ongoing violent conflicts. Despite most refugees¹ settling in neighbouring states, the number of individuals coming to wealthy countries has risen, with many settling in European destinations. Between 2015 and 2016, approximately 2.6 million refugees arrived in the European Union (EU), and Germany by far had the most refugees, having received approximately 1.2 million during those years (Eurostat 2020). In comparison, the stock of refugees in classic immigration countries such as the United States amounted to 815,600 by the end of 2016 (UNHCR 2017, 60–63).

The incorporation of refugees into receiving countries is of great concern, and many scholars have examined how this population has fared in recent years. Usually, such

CONTACT Yuliya Kosyakova  Yuliya.Kosyakova@iab.de

 Supplemental data for this article can be accessed <https://doi.org/10.1080/1369183X.2021.1988845>.

This article has been republished with minor changes. These changes do not impact the academic content of the article.

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

studies focus solely on refugees rather than comparing their situation to those of other immigrants who left their home countries for different reasons.² This reluctance seems to be partly due to data limitations, as new data collection efforts aimed at providing information on recent refugee populations typically do not include other new immigrants. This limitation is also somewhat grounded in the notion that refugees are distinct from other immigrants (Feller 2005). This view is reflected in a separate body of work on refugee studies on the one hand and on the sociology of international migration on the other hand (FitzGerald and Arar 2018). Considering the tendency to view refugee integration as special and different from immigrant integration, researchers have called for an overcoming of this scholarly chasm (FitzGerald and Arar 2018; Kogan and Kalter 2020). They have argued that refugee migration is a particular type of migration and that patterns of integration among refugees are subject to regularities similar to those associated with the integration of other kinds of immigrants (Kogan and Kalter 2020, 5; van Tubergen 2010, 530).

In this contribution, we adopt this view and examine destination-language acquisition among refugees and other new immigrants who came to Germany – one of Europe's premier destinations for both groups, during a similar period. Destination-language skills,³ which are key to sustainably incorporating immigrants into the receiving society, are of great importance to the labour market and shape opportunities for employment, earnings and occupational status (Chiswick and Miller 1995; Dustmann and Fabbri 2003; Kanas et al. 2012). In addition, such skills are essential for establishing contact with members of the majority population (Martinovic, van Tubergen, and Maas 2009) and play an important role in the educational careers of offspring (Alba, Sloan, and Sperling 2011; Schnepf 2007).

Against this background, we compare the learning pathways of refugees and other new immigrants throughout the early period after arrival and study a range of conditions that are known to improve proficiency. We expect that refugees and other immigrants respond to most conditions in similar ways. Differential patterns may nevertheless emerge because refugees differ from other immigrants in a range of conditions (Kristen and Seuring 2021). For example, given that refugee migration is often abrupt and therefore diverges from that of other immigrants who can plan and prepare for the experience, refugees are more likely to arrive with lower destination-language skills (Chiswick and Miller 2001; van Tubergen and Kalmijn 2005).

In our theoretical account, we build upon a general and widely applied model of language learning according to which destination-language proficiency is a function of the efficiency with which immigrants learn a new language, the incentives (or motivation) for investing in its acquisition and the degree of exposure to this language (Chiswick and Miller 1995; Chiswick and Miller 2001). We use this model to elaborate on various conditions that shape the development of proficiency after taking up residency in the destination country. Our account combines general processes of language learning with conditions that are specific to refugees and that may set their learning pathways apart from those of other new immigrants.

For the empirical study, we employ two sources of panel data on recently arrived adult immigrants: the *IAB-BAMF-SOEP Sample of Refugees in Germany* (Brücker, Rother, and Schupp 2017; Kühne, Jacobsen, and Kroh 2019) and the *IAB-SOEP Migration Sample* (Brücker et al. 2014). Respondents were interviewed soon after their arrival in

Germany and then were followed for up to either four (refugees) or seven years (other new immigrants). The data are particularly suitable for the empirical inquiry of language acquisition due to the provision of yearly indicators of German language proficiency as well as a range of pre- and postmigration characteristics.

The current study contributes to the existing body of research in various ways. First, by comparing refugees to other new immigrants, we overcome the divide characterising the literature that has mostly focused on either refugees or other immigrants. More importantly, this comparison enables us to examine similarities and differences in how language acquisition takes place among individuals who left their home countries for different reasons. Second, in contrast to many previous studies that focused on immigrants who have been living in the destination country for some time, we focus exclusively on recently arrived immigrants. The early years after arrival are crucial for second-language acquisition and lay the foundations for becoming proficient (Stevens 1999). The learning curve tends to be steeper during this initial period than it is in later years (Hartshorne, Tenenbaum, and Pinker 2018), rendering this phase a very dynamic stage that is crucial for individuals' future prospects. Third, the longitudinal nature of the two datasets allows us to depict the development of language skills over time, taking into account initial skill levels and a range of time-invariant and time-dependent conditions fostering language learning after arrival. Fourth, this setup makes it possible to disentangle premigration from postmigration characteristics and to investigate how both shape the processes of language acquisition in the period after arrival. Doing so is important to our study, which aims to address group differences. For example, while other immigrants can plan and prepare for their migration, refugees frequently have to leave their countries on short notice (Chiswick and Miller 2001, 394). Accordingly, premigration investments in language learning are more likely to be made by other immigrants than they are by refugees. Additionally, after migration, the two groups may have encountered different learning opportunities, for instance, in the form of German language courses offered specifically to refugees (Brücker et al. 2019; Kosyakova and Brenzel 2020). Finally, we seek to assess the relative importance of the three core constructs (i.e. efficiency, incentives and exposure) in an account of refugees' and other recent immigrants' destination-language acquisition. The findings of this work may enable policy makers to direct their efforts toward policies that are particularly productive for improving language proficiency.

The dynamics of language acquisition

The human capital model of language learning

In the literature on immigrants, destination-language proficiency is often viewed as a form of human capital with three characteristic properties (Chiswick and Miller 1995, 248; Chiswick and Miller 2001, 391–392): linguistic skills are embodied in the person, they are productive in that they can be used to create opportunities and increase future returns, and their acquisition requires investments that incur costs in the form of time and other resources. While economists mostly focus on language fluency as a means of improving employment opportunities and wages (e.g. Chiswick and Miller

1995; Dustmann and Fabbri 2003), sociologists also emphasise the significance of language skills for making friends cross-ethnically, navigating receiving-country institutions or succeeding in the educational system (e.g. Alba, Sloan, and Sperling 2011; Espenshade and Fu 1997; Martinovic, van Tubergen, and Maas 2009).

The model of language learning proposed by Chiswick and Miller (1995; 2001) was originally formulated from an economic perspective. Due to its general character, however, it has been adopted by researchers from different disciplines and used widely to address language learning processes. According to this model, language proficiency is a function of efficiency, incentives and exposure.

The term *efficiency* describes the extent to which a certain amount of exposure translates into fluency (Chiswick and Miller 1995, 250). Typically, younger individuals, those with greater cognitive skills and those with more education learn more efficiently (Chiswick and Miller 1995; Chiswick and Miller 2001; Dustmann and Fabbri 2003; Espenshade and Fu 1997; Hartshorne, Tenenbaum, and Pinker 2018).

Incentives refer to individuals' learning motivation, which can be driven by economic returns to language investments such as improved employment opportunities or higher wages (Chiswick and Miller 2001) and by noneconomic returns such as benefits originating from cross-ethnic contacts or from feelings of attachment to the receiving context (Espenshade and Fu 1997). Meanwhile, learning a new language involves direct costs in the form of expenses for language instruction or learning materials as well as indirect costs in terms of earnings forgone due to the time spent acquiring the new language (van Tubergen 2010, 516–517). To render learning investments attractive, future benefits need to outweigh such costs.

Finally, *exposure* refers to the degree to which the new language is present in contexts that immigrants encounter. A substantive body of empirical evidence suggests that exposure is the major source of destination-language proficiency (Kristen 2019, 523), indicating that conditions signalling exposure merit special attention. Most studies focus on linguistic opportunities encountered in receiving societies after migration. They refer to structured forms of learning, such as that in language courses (e.g. van Tubergen 2010), as well as to exposure that is not necessarily aimed at greater fluency but that occurs as a byproduct of activities that are pursued for other reasons (Kristen, Mühlau, and Schacht 2016, 182). Above all, these activities include interactions with speakers of the dominant language in the neighbourhood, the workplace, the family or the circle of friends (Chiswick and Miller 2001; Stevens 1999). In addition to postmigration exposure, an account of new immigrants' language acquisition requires considering exposure to the destination language prior to migration. Empirical studies examining the conditions of premigration exposure typically refer to premigration stays in the destination country, attending language classes in the home country to prepare for migration or acquiring such skills in school (Espinosa and Massey 1997; Kristen, Mühlau, and Schacht 2016).

Application to recent refugees and other new immigrants

Many researchers have elaborated on the processes associated with the model's three basic dimensions and have examined them for various groups of immigrants in different destinations around the world (e.g. Espenshade and Fu 1997; Dustmann and

Fabbri 2003; Espinosa and Massey 1997; Hwang and Xi 2008; van Tubergen and Kalmijn 2005; Kristen, Mühlau, and Schacht 2016). Following their reasoning and the related body of empirical evidence, which primarily points to general processes underlying fluency, we expect recent refugees to respond to most conditions in similar ways as other new immigrants. Differential patterns may nevertheless emerge because refugees are likely to have experienced certain situations or particular conditions that are less common among other new arrivals. Accordingly, in the following, we discuss a range of these conditions, which may be responsible for group differences in the level of proficiency upon arrival and for distinct pathways of language acquisition thereafter. Together, they may contribute to compositional differences across the two groups.

Efficiency. Some of these conditions originate from the forced nature of refugee migration. Refugees tend to be pushed out of their home countries, and they often experience further dangerous and life-threatening situations on their journey to a new destination (Hatton 2020). These defining experiences can have consequences for their future learning prospects. In this context, it has been shown that refugees suffer more often from posttraumatic stress and from related mental health problems than do other populations (Lindert et al. 2009; Dietrich et al. 2019). Poor mental health can in turn hamper the efficiency of language learning (Chiswick and Miller 2001; van Tubergen and Kalmijn 2005). However, studies addressing refugees' language acquisition thus far have not found evidence supporting this line of reasoning (Kristen and Seuring 2021; van Tubergen 2010; Hunkler and Khourshed 2020).

Compared to other new immigrants who came to Europe in recent years, contemporary refugees are also less educated on average (Spörlein and Kristen 2019; Spörlein et al. 2020). This observation is related to the economic development in their origin countries and, relatedly, to the educational expansion, which has not progressed as far in most refugees' home countries as it has in the countries from which other recent immigrants primarily originate (Spörlein and Kristen 2019; Spörlein et al. 2020). The resulting difference in the distribution of educational qualifications across groups suggests that, compared to other new immigrants, refugees are at a disadvantage when learning a new language (Chiswick and Miller 1995; Chiswick and Miller 2001; Dustmann and Fabbri 2003; Espenshade and Fu 1997). Considering that schools also train individuals in how to acquire new skills, refugees who have received very little formal education may be less prepared to acquire the destination language. While individuals who cannot read and write are virtually nonexistent in other recent immigrant populations, the least educated segments among recent refugees include considerable numbers of illiterate individuals (Kristen and Seuring 2021; UNESCO 2020; Brücker, Rother, and Schupp 2017, 47).⁴ For this reason, learning patterns may also differ between refugees and other new immigrants.

Incentives. In light of adverse conditions in their home countries, many refugees do not view returning to their countries as a viable option; in contrast, other new immigrants often have the opportunity to return to their home countries if they want to (Kogan and Kalter 2020; Cortes 2004). Therefore, we expect refugees to be more inclined to stay permanently than other recent immigrants. The longer time horizon in which refugees can realise the returns to their human capital investments should render language learning more attractive for them (Cortes 2004).

A different process relates to refugees' legal status. After their arrival, they frequently experience insecurities about their legal prospects of remaining in the destination country, in some cases, over a prolonged period (Kosyakova and Brenzel 2020). Scholars have argued that such insecurities could hamper individuals' motivation to invest in human capital (Hvidtfeldt et al. 2018; Kosyakova and Brenzel 2020; van Tubergen 2010). This should set refugees apart from other new immigrants who usually enter the receiving country already under the provision of a defined legal status. In the German case, many contemporary immigrants originate from EU countries. Its citizens are free to settle and work anywhere in the EU's territory, so they should not experience legal insecurities of the kind encountered by refugees.

Finally, we turn to the long-standing assumption according to which individuals who immigrate for work-related reasons are favourably selected on motivational features compared to those who migrate for other reasons, such as refugees or family migrants (Chiswick 1999).⁵ Due to their assumed higher level of motivation, economic migrants should be more inclined to attain destination-language proficiency. However, this reasoning and the accompanying focus on migration motives as potential drivers of language acquisition are challenged by the mounting evidence on migrant selectivity. Recent comparisons of refugees and labour migrants in Europe show that there is more individual-level variation in selectivity within these broad categories than between them and that different categories of immigrants and different groups of origin are composed of varying proportions of differently selected individuals (Spörlein and Kristen 2019; Spörlein et al. 2020). These findings attest to the inadequacy of group-level characterisations of motivational and other features and suggest that broadly defined migration motives might be unsuitable to account for differences in language fluency. This does not mean, however, that labour market orientations do not matter. Instead, we argue that all individuals who strive to participate and succeed in the destination country's economic sphere are more likely to invest in language learning than those who exhibit different orientations, such as focusing on family-related matters. This reasoning, however, is not specific to certain immigrant populations but should apply generally. Furthermore, refugees who choose to migrate to the Western Hemisphere rather than to a neighbouring country may exhibit higher selectivity for economic success (Kogan and Kalter 2020, 9; Spörlein et al. 2020). For this reason, their orientations should be similar to those of other new immigrants.

Exposure. Recent refugees can also be expected to differ from other new immigrants in their exposure to the destination language prior to migration. Because they are forced out of their home countries, refugees usually do not prepare for their departure and arrive with very low levels of linguistic skills (Kristen and Seuring 2021; Brücker, Kosyakova, and Vallizadeh 2020; Hunkler and Khoureshed 2020). In contrast, other new immigrants to Germany, especially those from neighbouring states such as Poland, have sometimes been exposed to the language of the destination country before migration, whether in school, during prior visits or in the form of investments in language learning aimed at preparing for their stay (Kristen, Mühlau, and Schacht 2016; Kristen and Seuring 2021). The level of proficiency upon arrival has implications for newcomers' early linguistic pathways. For individuals who learn a new language from scratch, the learning curve tends to be steeper, whereas for individuals who have already acquired a certain level of

proficiency, linguistic gains are usually smaller (Hartshorne, Tenenbaum, and Pinker 2018). Over time, however, as individuals with very low levels of initial skills catch up, these curves should gradually align.

In terms of postmigration exposure, two refugee-specific institutional features may create differential conditions for language acquisition. The first pertains to the provision of collective accommodation for refugees. In the early months after arrival, refugees often stay in collective accommodations and live with other refugees in close proximity (Brücker, Kosyakova, and Vallizadeh 2020). These kinds of shared accommodations are organised by official authorities and, since they target refugees, are not available to other new immigrants. Living in such circumstances may restrict refugees' exposure to speakers of the destination language, whereas living outside a reception centre should increase opportunities for establishing contact with speakers of the dominant language (van Tubergen 2010).

A second institutional feature that is likely to increase the linguistic exposure of refugees relative to other new immigrants concerns the German government's substantive efforts to provide language training to recent refugee populations (Brücker et al. 2019; Kosyakova and Brenzel 2020). Structured exposure in formal settings is known to raise proficiency (Hayfron 2001; Hoehne and Michalowski 2016; van Tubergen 2010; Ortman's 2020), and many refugees in Germany have made use of these opportunities (Brücker et al. 2019; Kristen and Seuring 2021; Brücker, Kosyakova, and Vallizadeh 2020). Higher rates of participation in language courses among refugees should reflect greater linguistic gains compared to the gains achieved by other new immigrants.

Assigning empirical indicators to the three core constructs

To empirically assess the three core constructs in an overall account of language proficiency, it is necessary to assign specific indicators to each of them. However, linking an indicator directly to incentives, efficiency or exposure is not always possible. For example, age at immigration is a classic measure of learning efficiency, with individuals who arrive at younger ages outperforming those who arrive later in life (Chiswick and Miller 1995; Chiswick and Miller 2001; Espenshade and Fu 1997; Stevens 1999; van Tubergen and Kalmijn 2005). This negative relationship between age at arrival and language proficiency is usually traced back to the cognitive advantages of younger individuals (Kristen 2019, 524). Notwithstanding this prominent interpretation, the finding of a negative relationship could also result from age-specific differences in incentives, with younger individuals facing a longer time horizon in which to reap the benefits of their learning investments.

In view of such ambiguities, we decided to base the assignment of each measure to one of the three constructs on substantive arguments linking the respective indicator to a specific construct and/or on the available empirical evidence. This proceeding is often-times justifiable, as is the case for age at immigration. Greater learning efficiency has been shown to be largely a function of higher neural plasticity, a universal finding among humans (Center on the Developing Child 2007). In contrast, incentive arguments related to the age at immigration are likely to be less general in scope. They require drawing on additional assumptions, which may apply differently to specific groups. For example, the time horizon in which individuals can realise the returns to their

achieved level of destination-language proficiency should be more relevant for the investment decisions of labour migrants, who focus on labour market returns, compared to family migrants, who migrate for reasons that may not necessarily require language skills. Moreover, in our empirical models, we include measures that capture, at least to some extent, alternative explanations. In the example on incentives, we consider the intention to stay in Germany. At the same time, we are aware that this proceeding may not always result in a fully adequate assignment of a given measure to a specific construct. The assignment, nevertheless, can be considered most plausible.

Data and methods

Data

The empirical analyses are based on two longitudinal data sources, which include samples of new immigrants in Germany. The *IAB-BAMF-SOEP Sample of Refugees in Germany* (Brücker, Rother, and Schupp 2017; Kühne, Jacobsen, and Kroh 2019) was launched in 2016, in the aftermath of the surge of refugee migration to Europe in 2015. It is based on three subsamples that were drawn from the *Central Register of Foreigners*. The empirical analyses are restricted to refugees aged 18–55 at arrival, with a duration of stay no longer than six years and valid information on language proficiency at the time of the first interview. Our analyses cover 7,091 individuals (85 percent of the original sample).

The *IAB-SOEP Migration Sample* (Brücker et al. 2014) was launched in 2013. It consists of two subsamples drawn from register data from the *Federal Employment Agency* (BA), the so-called *Integrated Employment Biographies*. The target population consists of individuals who immigrated to Germany between 1995 and 2013 and of second-generation individuals born after 1976. Given our interest in language acquisition among newly arrived immigrants, we exclusively focus on the part of the sample that covers first-generation immigrants. Again, we restrict our analyses to immigrants aged 18–55 at arrival, with a duration of stay no longer than six years and valid information on language proficiency at the time of the first interview. The resulting sample covers 1,829 respondents (71 percent of all first generation immigrants of the original sample). We provide detailed information on these two data sources in the Online Appendix A.

Measures

Language proficiency in German is based on an additive index comprising information on respondents' self-rated competences in speaking, reading, and writing. Each scale ranges from 0 ('very good') to 4 ('not at all'). We reversed these scales before calculating the index so a greater value indicates a higher level of proficiency. The index can take values between 0 and 12. The measure shows a high degree of internal consistency (for refugees: Cronbach's alpha = 0.94; for other immigrants: Cronbach's alpha = 0.93), with individual variables loading on a single factor (for refugees: eigenvalue = 2.68; for other immigrants: eigenvalue = 2.63).

To address the model of language learning, we rely on a range of indicators, which reflect the discussion in the theory section. Table A.2 in the Online Appendix presents the definitions of all variables and indicates whether the measures differ between the two data sources. Table A.3 in the Online Appendix also illustrates the panel waves in which the time-varying and time-invariant measures were collected. Finally, Table 1 provides information about the distributions of all variables in the two datasets.

In terms of efficiency, we consider the respondent's *age at immigration*; their *cognitive skills* measured by the Symbol-Digit Test, a speed-constrained measure of information-processing capacities (Lang et al. 2007), or math grades in the last school report; and the level of education acquired in the country of origin (henceforth, CO), i.e. *education in CO*. For refugees, we additionally examine whether individuals have acquired *CO literacy*. Possessing writing and reading skills in their first language allows second language learners to fall back on previously acquired learning strategies and writing skills (Scheible 2018, 2). Similarly, teaching materials and language course curricula often assume that language learners can read and write in their origin languages (Sharifian et al. 2021, 75). Furthermore, we consider *mental health*, as poor mental health is known to affect cognitive functioning (e.g. Trivedi 2006).

Regarding the incentives for learning the destination language, we include the respondents' *economic orientation*. Fleeing their country of origin due to armed conflict or persecution is typically the main migration motive among refugees (Hatton 2020). Moving to a neighbouring country in many cases allows escape from the threat of immediate harm. Hence, for refugees to undergo the additional burden of an arduous and dangerous journey to Germany suggests that other motives, such as economic incentives, also play a role (Brücker et al. 2016). We further consider the *intention to stay* in Germany and the type of *residency status*, which provides the legal basis for individuals to assess their prospects for permanent residence. For refugees in pending asylum proceedings, the length of the pending procedure is taken into account. The degree of *connection to CO* enters the analyses as an indicator of the individuals' emotional attachment to the country of origin rather than their destination. Moreover, we examine the respondents' *premigration status*, referring to their self-assessed position in the economic distribution prior to migration, their former occupational status, and their former labour market participation. Individuals who, prior to their migration, held a higher status position in the origin country may continue to see themselves as having a high status when experiencing status loss or unemployment (Ichou 2014). This subjective status perception can reflect a higher motivation to succeed (Destin et al. 2012) and, hence, a greater inclination to become proficient in the destination language.

Premigration exposure is measured by the level of self-reported *premigration proficiency* and by variables indicating whether the respondents had attended a language course prior to migration (i.e. *premigration language course*) as well as whether they had previously stayed in Germany (i.e. *premigration stay in Germany*). Postmigration exposure refers to any activity or condition that signals exposure to the destination language after arriving in Germany. Indicators include whether individuals took a *language course* and, if so, which language certificate they obtained and whether they have been or are currently enrolled in education (i.e. *education in Germany*). We also include three indicators of *language use*: with family members, with friends, and in

Table 1. Descriptive statistics for refugees and other immigrants

	Refugees			Other immigrants			Range
	N	Mean	SD	N	Mean	SD	
<i>Language proficiency</i>	16090	4.80	2.91	3908	6.92	3.02	0–12
<i>Duration of stay (in months)</i>	16096	22.24	10.60	5621	43.29	17.99	0–150
Efficiency							
<i>Age at immigration</i>	16096	31.48	9.37	5621	31.31	8.25	18–55
<i>Cognitive skills</i>							
Digit-Symbol Test (standardised)	7674	−0.01	1.02	3931	−0.02	1.06	−11–0.4
Math grades				5214	4.63	1.04	1–6
<i>Education in CO</i>	15058			5453			
Less than primary		0.20			0.01		0/1
Primary		0.19			0.04		0/1
Lower secondary		0.22			0.17		0/1
Upper secondary		0.24			0.32		0/1
Postsecondary nontertiary		0.02			0.12		0/1
Tertiary		0.13			0.34		0/1
<i>CO literacy</i>	16091	0.81	0.39				0/1
<i>Mental health index</i>	9804	48.03	11.77	2057	54.15	8.37	4.6–79.4
Incentives							
<i>Economic orientation</i>	15928			5399			
Economic		0.43			0.45		0/1
Family-/network-related		0.19			0.46		0/1
Other					0.09		0/1
<i>Intention to stay (permanently)</i>	15673	0.96		5416	0.77		0/1
<i>Residency status</i>							
Status of asylum request	14912						
Pending		0.24					0/1
Approved		0.63					0/1
Rejected		0.09					0/1
Resettlement		0.03					0/1
Length of pending asylum procedure	1459	17.19	10.56				1–140
Permanent residency				5418	0.77		0/1
<i>Connection to CO</i>	12666	3.41	1.33	2661	3.76	1.12	1–5
<i>Premigration status</i>							
Position in economic distribution	15281						
Below average		0.25					0/1
Average		0.47					0/1
Above average		0.27					0/1
Occupational status	9894	39.35	16.46	3627	41.93	17.53	10–89
Labour market participation	15711	0.65		5165	0.70		0/1
Premigration exposure							
<i>Premigration proficiency</i>	16054	0.30	1.31	5603	2.75	3.47	0–12
<i>Premigration language course</i>				5608	0.26		0/1
<i>Premigration stay in Germany</i>	16096	0.01		5621	0.13		0/1
Postmigration exposure							
<i>Language course</i>	16048			5613	0.48		0/1
No course attended		0.27					0/1
Course without certificate		0.50					0/1
Course with A1/A2		0.13					0/1
Course with B1/B2		0.10					0/1
Course with C1/C2		0.01					0/1
<i>Education in Germany</i>	16050	0.06		5620	0.15		0/1
<i>Language use</i>							
With family members	8255			3393			
Mainly in CO or another language		0.80			0.71		0/1
Equally		0.17			0.17		0/1
Mainly in German		0.03			0.12		0/1
With friends	8722			3467			
Mainly in CO or another language		0.45			0.43		0/1
Equally		0.31			0.32		0/1
Mainly in German		0.24			0.25		0/1

(Continued)

Table 1. Continued.

	Refugees			Other immigrants			Range
	N	Mean	SD	N	Mean	SD	
Media consumption	8401			2128			
Mostly in CO language		0.38			0.33		0/1
In both languages		0.31			0.38		0/1
Mostly in German		0.31			0.29		0/1
Improvement via family members/friends	16025	0.36					0/1
Improvement via media consumption	16025	0.67					0/1
Contact with Germans	16048	0.57		3476	0.83		0/1
Shared accommodation	16039	0.33		5225			0/1
Single	15942	0.33		5552	0.39		0/1
Controls							
Female	16096	0.41		5621	0.55		0/1
Child < age 6	15869	0.46		5619	0.39		0/1
Country of origin	16095			5621			
Syria		0.51					0/1
Afghanistan		0.12					0/1
Iraq		0.14					0/1
Eritrea		0.05					0/1
Other MENA		0.04					0/1
West Balkan		0.03					0/1
Former USSR		0.03					0/1
Other Africa		0.04					0/1
Other		0.02			0.00		0/1
Stateless		0.01					
Western EU					0.03		0/1
Southern EU					0.12		0/1
Eastern EU					0.09		0/1
Bulgaria					0.08		0/1
Romania					0.17		0/1
Poland					0.15		0/1
Middle East					0.05		0/1
North America					0.01		0/1
South America/Central America/Caribbean					0.03		0/1
Balkans/Russia/Eurasia					0.18		0/1
Sub-Saharan Africa					0.02		0/1
Asia					0.05		0/1
Sample	16096			5621			
M1					0.32		0/1
M2					0.68		0/1
M3		0.33					0/1
M4		0.35					0/1
M5		0.32					0/1

Notes: For the number of cases (N) and the range, we considered person-year observations; for the calculation of the mean and the SD, we considered the first observation for each person.

Sources: Refugees: IAB-BAMF-SOEP Survey of Refugees 2016–2019. Other immigrants: IAB-SOEP Migration Sample 2013–2019.

media consumption. Since these measures were collected only in later waves for refugees, in the initial wave, we additionally consider whether they were taught German by family members or friends and whether they consume media in German. We further consider whether respondents are in *contact with Germans*. For refugees, we include the accommodation type, which, especially in the early days after immigration, frequently meant staying in shared accommodations with other refugees.

As controls, we consider the respondent's *gender*, whether s/he has *children* under the age of 6 living in the household and the *country of origin*. We also record the *sample* to which each person belongs.

Analytical strategy

To model the language acquisition process, we rely on multilevel growth curve models in which measurement occasions are hierarchically nested in individuals (Hox, Moerbeek, and van the Schoot 2017). By using random intercepts, we model that individuals start from different competence levels, while the use of random slopes allows us to account for individual differences in the speed of language acquisition. In addition, we model the covariance between random intercepts and random slopes, considering that the initial proficiency level (at t_0 of the panel) can become relevant for the pace of language growth. A negative covariance could indicate that individuals who have already acquired greater destination-language skills at t_0 have less room for improvement and therefore show slower progress. The coefficients of the time-variant variables indicate within-person differences in language acquisition and, thus, address developments over time. For example, the measures of how postmigration exposure changes with the duration of stay, such as by attending a language course, account for patterns of language development. The time-invariant measures address the between-person differences in language skills. The coefficients presented for the constant attributes, such as most measures of learning efficiency, indicate why some individuals with a given initial value (e.g. a certain age at immigration) on average become more fluent than others, irrespective of their linguistic development.

To compare the relative strength of the influence of a group of variables, namely, the three sets of variables associated with the core constructs, we estimate sheaf coefficients. These are standardised multiple-partial regression coefficients that summarise the coefficients of a set of variables when other variables are controlled for (Heise 1972).⁶ Sheaf coefficients assume that a group of variables influences the dependent variable through a latent variable. In our analyses, the latent variables of interest are the three constructs of efficiency, incentives and exposure. Comparing the coefficients of these latent variables allows us to tentatively assess their relative importance for language learning. As discussed above, the grouping of variables comes with ambiguities, because a one-to-one assignment of the various measures to the constructs is not always feasible. In addition, the variables included may not capture the constructs fully, and they might do so unevenly. For these reasons, we present the sheaf coefficients as a descriptive tool, which summarises the patterns for the measures included and which points to differences in these patterns across the two groups. As the indicators included for refugees and other immigrants are not fully equivalent, we refrain from formally testing for group differences.

To address item nonresponse, we apply multiple imputation using chained equations (van Buuren 2012). We estimate 25 imputed datasets with complete information. Following Rubin's (1987) approach, we then combine the results of the analyses performed on each dataset. Table 1 illustrates that missing information was present to varying degrees across measures. The replication codes for data preparation and analyses are available at <https://osf.io/94myu/>.

Results

The development of destination-language proficiency over time

Before turning to the substantive findings, we note that the individuals in the refugee sample differ from the sample of other immigrants in their average length of stay in

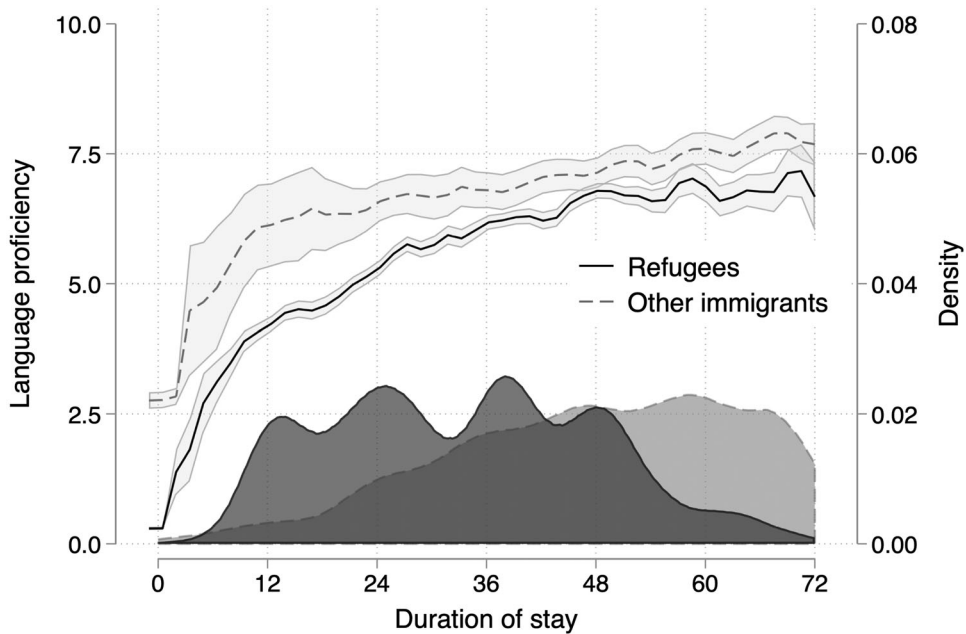


Figure 1. Development of language proficiency and density distribution of the duration of stay.

Sources: Refugees: IAB-BAMF-SOEP Survey of Refugees 2016–2019; Other immigrants: IAB-SOEP Migration Sample 2013–2019.

Germany. Figure 1 at the bottom highlights this pattern by presenting the density distributions for the duration of stay. The refugee sample is primarily composed of individuals early on in their settlement phase, whereas other immigrants have been living in Germany for a longer period of time.

The upper part of Figure 1 depicts refugees' and other immigrants' levels of German language proficiency upon arrival and the development of these skills thereafter. The curves illustrate that most refugees enter with little to no destination-language skills. Other immigrants do somewhat better at the beginning, with the initial levels amounting to approximately 2.5 points, which reflects 'poor' to 'fair' language skills on the scale from 0 to 12. Over time, both groups substantially gain in proficiency and eventually achieve values of approximately 7.5 points. The gap between refugees and other immigrants that is present upon entry diminishes over time and closes after approximately 48 months. The curves depicted in Figure 1 are in line with the well-established finding of greater linguistic improvements at the beginning (i.e. throughout the first year), reflecting a steeper learning curve, and slowed progress thereafter, reflecting a flattening curve (Hartshorne, Tenenbaum, and Pinker 2018). The baseline models of the multivariate account with their negative quadratic terms reveal the same pattern of decreasing linguistic returns to a longer duration of stay (see M1.0 for refugees and M2.0 for other immigrants in Table 2).

The dynamics of language learning among refugees and other immigrants

In the next step, we investigate the conditions associated with the core constructs of efficiency, incentives and exposure and focus on how the recent refugees are similar to

Table 2. Multilevel growth curve models of language proficiency.

	Refugees				Other immigrants			
	M1.0		M1.1		M2.0		M2.1	
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
<i>Duration of stay</i>	0.14*	(0.00)	0.05*	(0.00)	0.03*	(0.00)	0.03*	(0.00)
<i>Duration of stay, squared</i>	-0.00*	(0.00)	-0.00*	(0.00)	-0.00*	(0.00)	-0.00*	(0.00)
Efficiency								
<i>Age at immigration</i>			-0.06*	(0.00)			-0.06*	(0.01)
<i>Cognitive skills</i>								
<i>Digit-Symbol Test (standardised)</i>			0.10*	(0.03)			0.23*	(0.05)
<i>Math grades</i>							0.15*	(0.04)
Premigration education								
<i>Education in CO (Ref. primary)</i>								
<i>Less than primary</i>			-0.40*	(0.07)			-0.74	(0.44)
<i>Lower secondary</i>			0.45*	(0.06)			0.27	(0.23)
<i>Upper secondary</i>			0.77*	(0.06)			0.61*	(0.22)
<i>Postsecondary nontertiary</i>			0.91*	(0.15)			0.87*	(0.24)
<i>Tertiary</i>			1.14*	(0.09)			1.21*	(0.24)
<i>CO literacy</i>			0.57*	(0.06)				
<i>Mental health</i>			0.00	(0.00)			-0.01	(0.01)
Incentives								
<i>Economic orientation (Ref. not mentioned/other)</i>								
<i>Economic</i>			0.07	(0.04)			-0.21	(0.14)
<i>Family-/network-related</i>			0.03	(0.05)			-0.21	(0.14)
<i>Intention to stay</i>			0.20*	(0.09)			-0.07	(0.10)
Residency status								
<i>Status of asylum request (Ref. pending)</i>								
<i>Approved</i>			0.23*	(0.10)				
<i>Rejected</i>			0.10	(0.12)				
<i>Resettlement</i>			0.30*	(0.15)				
<i>Length of pending asylum procedure</i>			0.01	(0.01)				
<i>Permanent residency</i>							0.30*	(0.11)
<i>Connection to CO</i>			0.00	(0.01)			0.02	(0.06)
Premigration status								
<i>Position in economic distribution (Ref. below average)</i>								
<i>Average</i>			0.16*	(0.05)				
<i>Above average</i>			0.23*	(0.06)				
<i>Occupational status</i>			0.01*	(0.00)			0.00	(0.00)
<i>Labour market participation</i>			-0.30*	(0.09)			-0.08	(0.18)
Premigration exposure								
<i>Premigration German proficiency</i>			0.15*	(0.02)			0.27*	(0.01)
<i>Premigration language course</i>							0.29*	(0.09)
<i>Premigration stay in Germany</i>			0.40*	(0.19)			0.54*	(0.12)
Postmigration exposure								
<i>Language course (Ref. no course attended)</i>								
<i>Course without certificate</i>			0.99*	(0.05)				
<i>Course with A1/A2</i>			1.69*	(0.06)				
<i>Course with B1/B2</i>			2.78*	(0.07)				
<i>Course with C1/C2</i>			3.69*	(0.13)				
<i>Education in Germany</i>			0.49*	(0.06)			0.78*	(0.10)
Language use								
<i>With family members (Ref. mainly in CO or another language)</i>								
<i>Equally</i>			0.45*	(0.06)			0.32*	(0.11)
<i>Mainly in German</i>			0.54*	(0.11)			0.54*	(0.16)
<i>With friends (Ref. mainly in CO or another language)</i>								

(Continued)

Table 2. Continued.

	Refugees				Other immigrants			
	M1.0		M1.1		M2.0		M2.1	
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
Equally			0.56*	(0.05)			0.76*	(0.10)
Mainly in German			0.94*	(0.06)			1.20*	(0.12)
Media consumption (Ref. mostly in CO language)								
In both languages			0.28*	(0.05)			0.52*	(0.14)
Mostly in German			0.31*	(0.05)			1.00*	(0.22)
Improvement via family members/friends		0.16*		(0.04)				
Improvement via media consumption			0.63*	(0.04)				
Contact with Germans			0.40*	(0.04)			0.46*	(0.11)
Shared accommodation			-0.18*	(0.04)				
Single			0.13*	(0.05)			0.10	(0.08)
_cons	3.17*	(0.10)	2.16*	(0.21)	6.75*	(0.36)	3.97*	(0.59)
var(duration)	0.00		0.00		0.00		0.00	
var(_cons)	5.20		1.73		6.65		1.02	
corr(duration,_cons)	-0.34		-0.65		-1.00		-1.00	
var(Residual)	2.75		2.71		3.28		3.07	
N observations	16096		16096		5621		5621	
N individuals	7091		7091		1829		1829	
N imputations	25		25		25		24	

Notes: * $p < 0.05$. Additionally, we controlled for gender, children under the age of 6, country of origin, and sample. Sources: Refugees: IAB-BAMF-SOEP Survey of Refugees 2016–2019. Other immigrants: IAB-SOEP Migration Sample 2013–2019.

or differ from other new immigrants. For this purpose, Table 2 presents the results of the multilevel growth curve models of language proficiency separately for refugees (M1.1) and other immigrants (M2.1).

Regarding efficiency, our findings point to remarkable similarities across the two types of immigrants that are largely in line with prior research. Younger individuals, those with higher scores on the cognitive skill measures and the respondents with more education are more successful learners of the destination language. We also find evidence in favour of our reasoning on the less educated segments of the refugee population. We argued that according to Western standards, refugees can seem poorly educated (see Table 1). At the same time, they have attained levels of formal education that set them apart from a sometimes substantive share of their origin population, who has not reached such levels (Spörlein et al. 2020). In accordance with this observation, our analyses demonstrate that those who have completed primary education, the lowest level of formal education, already have an advantage for language acquisition over less educated individuals; this finding is also true for those who are literate in the origin language. Mental health, in contrast, does not constitute a major hurdle for becoming proficient.

Turning to incentives, only some of the findings follow the expectations outlined in the theoretical section. For example, refugees' economic orientation is not relevant to their language learning, but those who intend to stay in Germany permanently increase their proficiency. For the comparison group of other immigrants, only one of the indicators is in line with the theoretical reasoning: that is, recent immigrants with permanent residency achieve higher proficiency values than do those with a temporary residency status. For refugees, one of the most important results regarding incentives concerns

their legal status: individuals whose asylum requests are approved or who are allowed to stay permanently as resettled refugees achieve higher scores on the proficiency measure than those who are still awaiting decisions. In addition, there seems to be a linguistic penalty for refugee immigrants who have been active in the labour market prior to their migration. Within the group of individuals with labour market experience, however, those who previously held a higher occupational status and a better position in the economic distribution of their home country attain greater fluency than those who lack such advantages.

While our results on incentives are mixed, this is not the case for the third set of measures of exposure. In both groups, the findings are completely in line with the reasoning outlined earlier that being exposed to the destination language is essential for becoming proficient. Premigration German skills and premigration stays are positively related to language learning among refugees and other immigrants. Other immigrants also benefit from language course attendance prior to their migration.

Regarding postmigration exposure, we observe that structured learning opportunities, in the form of either attending language courses or receiving training in the educational system, can be a major driver of fluency. Among refugees, the completion of a language course without a certificate is associated with an increase of approximately 1 point on the proficiency scale, while completing it with a certificate results in even greater gains. To put this in perspective, it takes refugees slightly more than 7 months on average ($1/0.14$, see M1.0 in Table 2) to increase their proficiency level by 1 point. Given that courses for lower qualification levels such as A1 typically take approximately 160 course hours and are often completed within 1–2 months, these gains are considerable. Notably, each additional course level reached pays off linguistically. Obtaining the highest certificate (i.e. C1/C2 with approximately 200 course hours usually completed within 3 months) translates into a proficiency improvement of approximately 3.7 points. Achieving a similar linguistic gain takes close to 2 years on average ($3.7/0.14 \sim 26$ months). Structured exposure is also beneficial for other immigrants, albeit to a lower extent: course attendance is associated with a 0.5-point increase on the proficiency scale. However, the course measure used for other immigrants is less detailed than the one available for refugees (see Table A.1 in the Online Appendix), and it was included only once in the panel (see Table A.2 in the Online Appendix), preventing a direct comparison of the findings.

In addition to linguistic improvements following exposure to structured learning environments, informal exposure that takes place in everyday contexts matters as well. Both types of immigrants benefit from speaking with family members and friends in German, from being in contact with Germans and from consuming media in the destination language. As for refugee-specific circumstances, the results indicate that residing in shared accommodations with other refugees is less helpful, as it seems to constitute a hurdle to becoming proficient.

In an additional step, we reran all analyses for the separate language domains (i.e. speaking, reading and writing). We present the findings in Table B.1 and Figure B.1 in the Online Appendix. Most of the coefficients (93 percent) show the same direction and significance, and some even show the same results. This high degree of consistency is accompanied by various differences in the effect sizes across the domains. For example, measures for efficiency—such as cognitive skills, educational attainment and literacy—in

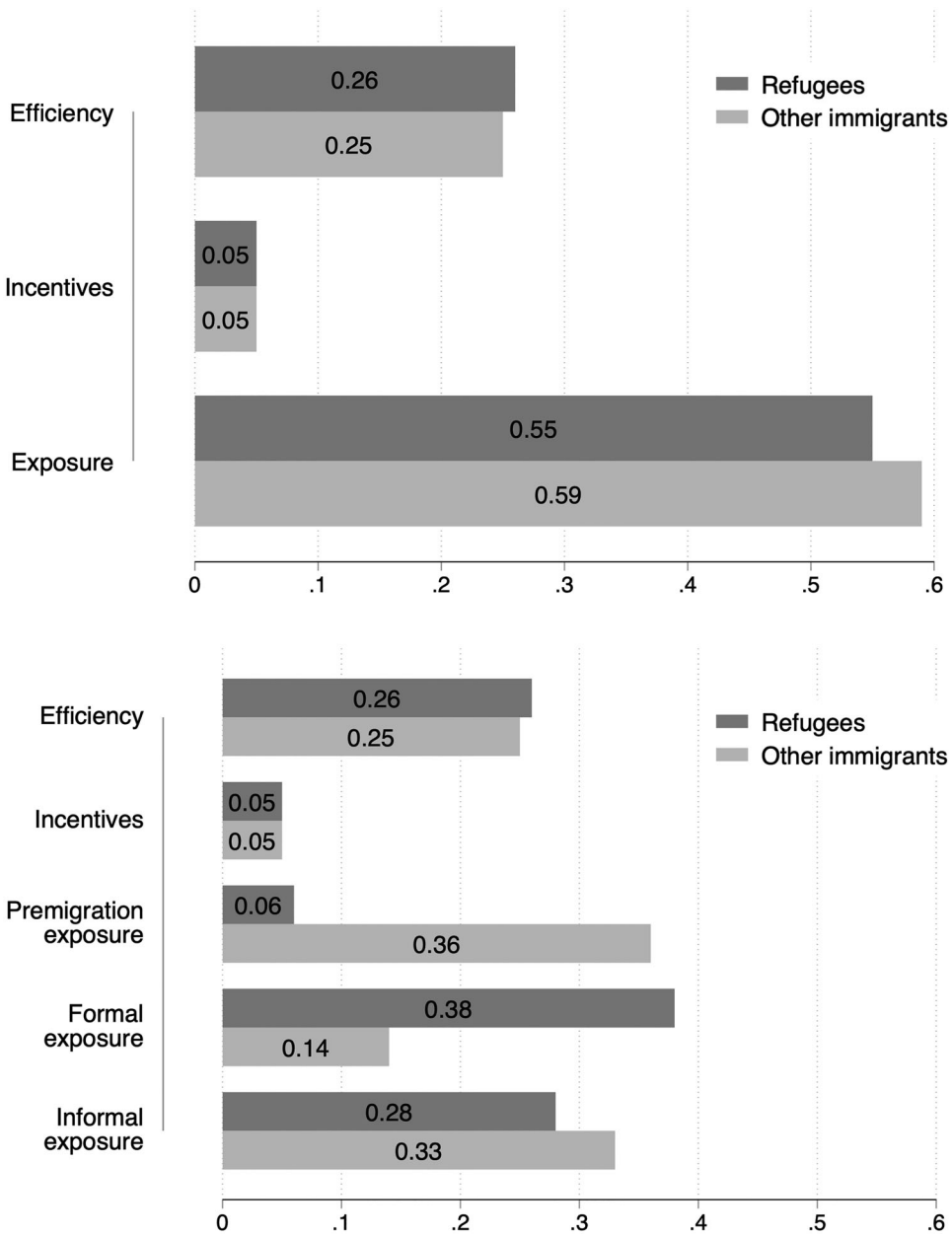


Figure 2. Relative influence of efficiency, incentives and exposure on language proficiency, standardised sheaf coefficients.

Sources: Refugees: IAB-BAMF-SOEP Survey of Refugees 2016–2019. Other immigrants: IAB-SOEP Migration Sample 2013–2019.

the origin language turn out to be more important for the relatively more demanding domains of reading and writing than they are for speaking. Similarly, language courses foster reading and writing skills to a greater extent than speaking skills, while nevertheless being a crucial contributor to proficiency in all domains. Using the destination

language with friends, in contrast, is a more advantageous tool for acquiring speaking skills than it is for acquiring reading or writing skills. Overall, these additional analyses allow for more fine-grained descriptions, without substantively challenging the main results reported above for the additive index of destination-language proficiency.

What matters most: efficiency, incentives or exposure?

To assess the relative importance of the three theoretical constructs, Figure 2 presents the standardised sheaf coefficients. The panel in the upper part of Figure 2 illustrates that, for refugees and other immigrants alike, exposure is the driving force of language acquisition. Exposure is also more important than efficiency and incentives taken together. Efficiency is the second most important construct, while incentives seem to be less consequential for language learning. The latter result is at odds with the effort to present incentive-related arguments in the literature and in our account.

In view of the central role that exposure plays in the acquisition of the destination language, the illustration at the bottom of Figure 2 further differentiates between the key components of exposure: (a) premigration exposure, (b) formal exposure in terms of attending language courses and enrolment in education in Germany, and (c) informal exposure, covering a range of opportunities to use the destination language in everyday life. Zooming in on exposure reveals notable differences between the two types of immigrants under study. While premigration exposure is of minor relevance to language learning for refugees, it is an important source of proficiency for other immigrants. This finding may reflect that most refugees were not exposed to German prior to their arrival, whereas other immigrants more frequently were (see Table 1). However, the key finding seems to be that for refugees, structured learning substantially contributes to fluency and is more important than informal exposure. Informal exposure, in contrast, seems to be slightly more important for other immigrants than it is for refugees.

Conclusions

Rather than viewing refugees as distinct from other types of immigrants, we proposed that refugee migration is a special case of immigration and that the processes of incorporating refugees into the receiving society are subject to regularities similar to those of the integration of other kinds of immigrants (Kogan and Kalter 2020, 5; van Tubergen 2010, 530). To exemplify this reasoning, we studied the destination-language acquisition of recently arrived refugees and other new immigrants in Germany. This study was theoretically based on a widely used model of language learning according to which destination-language proficiency is a function of the efficiency with which individuals learn a new language, the incentives for investing in its acquisition and the degree of exposure to this new language (Chiswick and Miller 1995, 2001). From this model of language learning, we derived arguments regarding the three core constructs. They refer to general processes of language learning as well as to a range of conditions that are specific to refugees and distinguish them from other new immigrants. To empirically investigate these considerations, we employed two longitudinal datasets covering recent arrivals in Germany: the *IAB-BAMF-SOEP Sample of Refugees in Germany* and the *IAB-SOEP Migration Sample*.

According to the descriptive results, most refugees arrive with little knowledge of German, whereas other immigrants often have some command of the destination language upon entry. In both groups, language growth is most pronounced in the early period of their stay and then levels off. Refugees display particularly strong gains in proficiency in the first year. After approximately three years, they close the linguistic gap relative to other new immigrants. The multivariate account supports this descriptive picture. The findings are also in line with much of the literature on immigrants' language acquisition and confirm the notion that language learning follows a general logic that applies similarly to different types of immigrants. At the same time, the results point to the need to consider group-specific conditions and circumstances.

Prior research has emphasised the role of incentives for investing in human capital. In this respect, several arguments as to why refugees may differ from other immigrants have been brought forward: group differences in the time horizon that different types of immigrants envisage for their stay (e.g. Chiswick and Miller 2001; Cortes 2004; Kogan and Kalter 2020), the legal hurdles they typically encounter (e.g. Hvidtfeldt et al. 2018; Kosyakova and Brenzel 2020; van Tubergen 2010) and motivational differences related to group-specific selective migration (e.g. Chiswick 1999). However, our findings provide little support for the importance of incentives in accounting for differential developments of destination-language proficiency. Quite to the contrary, the overall contribution of incentives to language learning is relatively small, and most estimates of the growth curve models are irrelevant or run counter to expectations. Hence, there seems to be an incongruity between the prominence of incentive-related arguments on the one hand and the lack of consistent empirical evidence supporting them on the other hand (Kristen 2019, 525). In our account, probably the most important finding regarding incentives is related to a refugee-specific condition: individuals whose asylum requests are approved achieve greater linguistic fluency than those whose asylum requests are still pending.

While incentives seem to be less relevant to language learning, our results for efficiency and exposure are strikingly consistent for refugees and other immigrants. Efficiency and, to a larger extent, exposure to destination language are the key drivers of proficiency. A closer look at exposure points to an important distinction in exposure to formal versus informal contexts. Whereas refugees have benefited substantially from the large-scale provision of targeted language training that offers intense and—compared to everyday interactions—systematic, conscious and focused exposure to the destination language (Hoehne and Michalowski 2016, 134), informal exposure to German in everyday settings has been found to be relatively more important for other immigrants. This observation does not suggest that language acquisition processes differ between refugees and other immigrants but rather that the structural conditions under which they acquire the new language differ. The day-to-day lives of refugees, especially in the early period after arrival, are influenced to a greater extent by administrative regulations, such as the duty to reside in a specific place for a certain time or rules connected to their (provisional) legal status. Refugees are also more frequent recipients of targeted offers such as those related to language training than are other immigrant types. Other immigrants, in contrast, are subject to less restrictive administrative conditions and therefore have more opportunities to self-select into informal, less structured exposure contexts.

The comparison of the relative importance of incentives, efficiency and exposure provides important insights, for example, regarding the question in which realms to best support language learning. At the same time, the value of this descriptive tool relies on whether the analyses captured the most relevant measures and whether they have been accurately assigned to the constructs. As pointed out before, there is not always a one-to-one correspondence between them, so the eventual ordering resulting from our calculations must be considered with caution. Nevertheless, since the assignments were based on substantive arguments and the available evidence, the indicators, in most cases, could not have been just as well assigned to another construct.

Issues of harmonisation are challenging for any empirical study that includes different data sources. Comparisons are difficult if relevant variables are included in only one of the two datasets (e.g. regarding respondents' premigration status or certain measures of language use) or are measured differently, such as in the case of enrolment in language courses. The two panels also covered somewhat different periods (2016–2019 for refugees and 2013–2019 for other immigrants). Another shortcoming is that the dependent variable was based on the respondents' self-assessed proficiency. These self-report measures have been implemented in surveys for quite some time, and most research on immigrants' language acquisition is based on them. However, the measures do not fully match actual skills as measured by standardised tests and are biased to some extent (Edele et al. 2015). Finally, we considered immigrants and refugees as aggregate groups; correspondingly, we were unable to consider variation within them (e.g. among refugees between approved refugees, asylum seekers, or internally displaced persons; among immigrants between economic immigrants, family immigrants, or repatriates). To address this issue, larger sample sizes are required than those currently available.

The policy implications of our results for other prominent refugee destinations seem straightforward: increasing exposure to the destination language should be high on the agenda of policy makers seeking to support new immigrants in achieving fluency in the destination language. This is important because exposure is the key to acquiring a new language. Focusing on either efficiency or incentives is less advisable because increasing learning efficiency hardly seems feasible considering characteristics such as age or cognitive skills and because incentives are empirically inconsequential. Our findings suggest that the large-scale provision of language courses for refugees has been a successful tool for fostering their early incorporation. The long-term returns will likely outweigh the associated costs.

Notes

1. In using the term 'refugee', we refer to individuals who seek asylum outside their country of origin or any other form of protection, irrespective of their legal status.
2. In using the term 'other immigrants', we refer to immigrants who leave their countries of origin for economic, family-related, educational or other reasons that are distinct from the drivers of humanitarian migration.
3. Throughout this study, we use the terms 'language proficiency' and 'language skills' interchangeably to describe a broad range of competences without referring to a particular domain (Kristen, Mühlau, and Schacht 2016, 204).

4. For Germany, this situation can be illustrated by comparing the literacy rates in the countries of origin of the two largest refugee groups with those of the two largest groups of other new immigrants. According to the World Bank, the most recently available numbers for people aged 15 and older point to a literacy rate of 43 percent in Afghanistan (2018), 81 percent in Syria (2004), 99 percent in Poland (2008) and 99 percent in Romania (2018; The World Bank 2020).
5. In addition to motivational characteristics, Chiswick (1999) refers to a range of attributes according to which the selection of economic migrants is expected to differ from that of refugees. These include disparities in cognitive ability. Therefore, his contribution can be used to argue for group differences in efficiency.
6. For the empirical implementation in Stata, we use SHEAFCOEF by Buis (2009).

Acknowledgements

We would like to thank Irena Kogan for providing helpful advice on an earlier version of this paper and all those who commented on the presentation of this work at the research colloquium ‘Empirical Sociological Research’ at the University of Konstanz (July 2021). Replication codes for data preparation and analyses are available at <https://osf.io/94myu/>.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Yuliya Kosyakova  <http://orcid.org/0000-0002-9621-1755>

Cornelia Kristen  <http://orcid.org/0000-0003-3426-9174>

Christoph Spörlein  <http://orcid.org/0000-0002-6148-9490>

References

- Alba, R., J. Sloan, and J. Sperling. 2011. “The Integration Imperative: The Children of Low-Status Immigrants in the Schools of Wealthy Societies.” *Annual Review of Sociology* 37 (1): 395–415. doi:10.1146/annurev-soc-081309-150219.
- Brücker, H., J. Croisier, Y. Kosyakova, H. Kröger, G. Pietrantuono, N. Rother, and J. Schupp. 2019. “Second Wave of the IAB-BAMF-SOEP Survey: Language Skills and Employment Rate of Refugees Improving with Time.” *IAB Brief Report* 2019 (3): 1–19. http://doku.iab.de/kurzber/2019/kb0319_englisch.pdf.
- Brücker, H., Y. Kosyakova, and E. Vallizadeh. 2020. “Has There Been a ‘Refugee Crisis’? New Insights on the Recent Refugee Arrivals in Germany and Their Integration Prospects.” *Soziale Welt* 71 (1–2): 24–53. doi:10.5771/0038-6073-2020-1-2-24.
- Brücker, H., M. Kroh, S. Bartsch, J. Goebel, S. Kühne, E. Liebau, P. Trübswetter, I. Tucci, and J. Schupp. 2014. The New IAB-SOEP Migration Sample: An Introduction Into the Methodology and the Contents.” *SOEP Survey Papers* 216: Series C. Berlin: DIW/SOEP. https://www.diw.de/documents/publikationen/73/diw_01.c.570700.de/diw_ssp0216.pdf.
- Brücker, H., N. Rother, and J. Schupp. 2017. “IAB-BAMF-SOEP-Befragung von Geflüchteten 2016: Studiendesign, Feldergebnisse Sowie Analysen Zu Schulischer Wie Beruflicher Qualifikation, Sprachkenntnissen Sowie Kognitiven Potenzialen”. *IAB- Forschungsbericht*, 13/ 2017. Nürnberg: IAB. <http://doku.iab.de/forschungsbericht/2017/fb1317.pdf>.
- Brücker, H., N. Rother, J. Schupp, C. Babka von Gostomski, A. Böhm, T. Fendel, M. Friedrich, et al. 2016. “Forced Migration, Arrival in Germany, and First Steps Toward Integration.” *DIW Economic Bulletin* 6 (48): 541–556. <http://hdl.handle.net/10419/148611>.

- Buis, M. L. 2009. *SHEAFCOEF: Stata Module to Compute Sheaf Coefficients*. Statistical Software Components S456995, Boston College Department of Economics, revised 04 Sep 2010.
- Center on the Developing Child. 2007. *The Science of Early Childhood Development (InBrief)*. <https://developingchild.harvard.edu/resources/inbrief-science-of-ecd/>.
- Chiswick, B. R. 1999. "Are Immigrants Favorably Self-Selected?" *American Economic Review* 89 (2): 181–185. doi:10.1257/aer.89.2.181.
- Chiswick, B. R., and P. W. Miller. 1995. "The Endogeneity Between Language and Earnings: International Analyses." *Journal of Labor Economics* 13 (2): 246–288. doi:10.1086/298374.
- Chiswick, B. R., and P. W. Miller. 2001. "A Model of Destination-Language Acquisition: Application to Male Immigrants in Canada." *Demography* 38 (3): 391–409. doi:10.2307/3088354.
- Cortes, K. E. 2004. "Are Refugees Different from Economic Immigrants? Some Empirical Evidence on the Heterogeneity of Immigrant Groups in the United States." *Review of Economics and Statistics* 86 (2): 465–480. doi:10.1162/003465304323031058.
- Destin, M., S. Richman, F. Varner, and J. Mandara. 2012. "'Feeling' Hierarchy: The Pathway from Subjective Social Status to Achievement." *Journal of Adolescence* 35 (6): 1571–1579. doi:10.1016/j.adolescence.2012.06.006.
- Dietrich, H., A. Radwan A, S. Tagay, J. Hebebrand, and V. Reissner. 2019. "Screening for Posttraumatic Stress Disorder in Young Adult Refugees from Syria and Iraq." *Comprehensive Psychiatry* 90 (April): 73–81. doi:10.1016/j.comppsy.2018.11.001.
- Dustmann, C., and F. Fabbri. 2003. "Language Proficiency and Labour Market Performance of Immigrants in the UK." *The Economic Journal* 113 (489): 695–717. doi:10.1111/1468-0297.t01-1-00151.
- Edele, A., J. Seuring, C. Kristen, and P. Stanat. 2015. "Why Bother with Testing? The Validity of Immigrants' Self-Assessed Language Proficiency." *Social Science Research* 52, Elsevier Inc. 99–123. doi:10.1016/j.ssresearch.2014.12.017.
- Espenshade, T. J., and H. Fu. 1997. "An Analysis of English-Language Proficiency among U.S. Immigrants." *American Sociological Review* 62 (2): 288–305. doi:10.2307/2657305.
- Espinosa, K. E., and D. S. Massey. 1997. "Determinants of English Proficiency among Mexican Migrants to the United States." *International Migration Review* 31 (1): 28–50. doi:10.2307/2547256.
- Eurostat. 2020. *Asylum Seekers and First-Time Asylum Seekers by Citizenship, Age and Sex. Annual Aggregated Data (Rounded)*. Luxembourg: Eurostat. http://ec.europa.eu/eurostat/data/database?node_code=migr.
- Feller, E. 2005. "Refugees Are Not Migrants." *Refugee Survey Quarterly* 24 (4): 27–35. doi:10.1093/rsq/hdi077.
- FitzGerald, D. S., and R. Arar. 2018. "The Sociology of Refugee Migration." *Annual Review of Sociology* 44 (1): 387–406. doi:10.1146/annurev-soc-073117-041204.
- Hartshorne, J. K., J. B. Tenenbaum, and S. Pinker. 2018. "A Critical Period for Second Language Acquisition: Evidence from 2/3 Million English Speakers." *Cognition* 177: 263–277. doi:10.1016/j.cognition.2018.04.007.
- Hatton, T. J. 2020. "Asylum Migration to the Developed World: Persecution, Incentives, and Policy." *Journal of Economic Perspectives* 34 (1): 75–93. doi:10.1257/jep.34.1.75.
- Hayfron, J. E. 2001. "Language Training, Language Proficiency and Earnings of Immigrants in Norway." *Applied Economics* 33 (15): 1971–1979. doi:10.1080/00036840010018630.
- Heise, D. R. 1972. "Employing Nominal Variables, Induced Variables, and Block Variables in Path Analyses." *Sociological Methods & Research* 1 (2): 147–173. doi:10.1177/004912417200100201.
- Hoehne, J., and I. Michalowski. 2016. "Long-Term Effects of Language Course Timing on Language Acquisition and Social Contacts: Turkish and Moroccan Immigrants in Western Europe." *International Migration Review* 50 (1): 133–162. doi:10.1111/imre.12130.
- Hox, J. J., M. Moerbeek, and R. van de Schoot. 2017. *Multilevel Analysis: Techniques and Application*. New York: Routledge.
- Hunkler, C., and M. Khourshed. 2020. "The Role of Trauma for Integration. The Case of Syrian Refugees." *Soziale Welt* 71 (1–2): 90–122. doi:10.5771/0038-6073-2020-1-2-90.

- Hvidtfeldt, C., M. L. Schultz-Nielsen, E. Tekin, and M. Fosgerau. 2018. "An Estimate of the Effect of Waiting Time in the Danish Asylum System on Post-Resettlement Employment among Refugees: Separating the Pure Delay Effect from the Effects of the Conditions Under Which Refugees Are Waiting." *PLoS ONE* 13 (11): 4–8. doi:10.1371/journal.pone.0206737.
- Hwang, S.-S. Shong, and J. Xi. 2008. "Structural and Individual Covariates of English Language Proficiency." *Social Forces* 86 (3): 1079–1104. doi:10.1353/sof.0.0017.
- Ichou, M. 2014. "Who They Were There: Immigrants' Educational Selectivity and Their Children's Educational Attainment." *European Sociological Review* 30 (6): 750–765. doi:10.1093/esr/jcu071.
- Kanas, A., B. R. Chiswick, T. van der Lippe, and F. van Tubergen. 2012. "Social Contacts and the Economic Performance of Immigrants: A Panel Study of Immigrants in Germany." *International Migration Review* 46 (3): 680–709. doi:10.1111/j.1747-7379.2012.00901.x.
- Kogan, I., and F. Kalter. 2020. "An Empirical–Analytical Approach to the Study of Recent Refugee Migrants in Germany." *Soziale Welt* 71 (1–2): 3–23. doi:10.5771/0038-6073-2020-1-2-3.
- Kosyakova, Y., and H. Brenzel. 2020. "The Role of Length of Asylum Procedure and Legal Status in the Labour Market Integration of Refugees in Germany." *Soziale Welt* 71 (1–2): 123–159. doi:10.5771/0038-6073-2020-1-2-123.
- Kristen, C. 2019. "Language Assimilation and the Education of Immigrant Students." In *Research Handbook on the Sociology of Education*, edited by Rolf Becker, 519–534. Cheltenham (UK) and Northampton (MA, USA): Edward Elgar Publishing. doi:10.4337/9781788110426.00039.
- Kristen, C., P. Mühlau, and D. Schacht. 2016. "Language Acquisition of Recently Arrived Immigrants in England, Germany, Ireland, and the Netherlands." *Ethnicities* 16 (2): 180–212. doi:10.1177/1468796815616157.
- Kristen, C., and J. Seuring. 2021. "Destination-Language Acquisition of Recently Arrived Immigrants: Do Refugees Differ from Other Immigrants?" *Journal of Educational Research Online* 13 (1): 128–156. doi:10.31244/jero.2021.01.05.
- Kühne, S., J. Jacobsen, and M. Kroh. 2019. "Sampling in Times of High Immigration: The Survey Process of the IAB-BAMF-SOEP Survey of Refugees." *Survey Methods: Insights from the Field*, 1–9. doi:10.13094/SMIF-2019-00005.
- Lang, D., D. Weiss, A. Stocker, and B. von Rosenblatt. 2007. "Assessing Cognitive Capacities in Computer-Assisted Survey Research: Two Ultra-Short Tests of Intellectual Ability in the German Socio-Economic Panel (SOEP)." *Schmollers Jahrbuch: Journal of Applied Social Science Studies* 127 (1): 183–192.
- Lindert, J., O. S. von Ehrenstein, S. Priebe, A. Mielck, and E. Brähler. 2009. "Depression and Anxiety in Labor Migrants and Refugees – A Systematic Review and Meta-Analysis." *Social Science & Medicine* 69 (2): 246–257. doi:10.1016/j.socscimed.2009.04.032.
- Martinovic, B., F. van Tubergen, and I. Maas. 2009. "Dynamics of Interethnic Contact: A Panel Study of Immigrants in the Netherlands." *European Sociological Review* 25 (3): 303–318. doi:10.1093/esr/jcn049.
- Ortmanns, V. 2020. *Issues in Measuring Education in Cross-National and Migration Surveys*. Mannheim: University of Mannheim.
- Rubin, D. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.
- Scheible, J. A. 2018. Literacy Training and German Language Acquisition among Refugees: Knowledge of German and the Need for Support among Integration Course Attendees Learning a Second Alphabet and Those with No Literacy Skills." *BAMF-Brief Analysis*, 1/2018. Nürnberg: BAMF-FZ. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-67560-1>.
- Schnepf, S. V. 2007. "Immigrants' Educational Disadvantage: An Examination Across Ten Countries and Three Surveys." *Journal of Population Economics* 20 (3): 527–545. doi:10.1007/s00148-006-0102-y.
- Sharifian, F., M. Sadeghpour, S. M. Barton, J. Barry, G. Barton, and I. Yilmaz. 2021. "English Language Learning Barriers of Afghan Refugee Women in Australia." *International Journal of Applied Linguistics* 31 (1): 65–78. doi:10.1111/ijal.12320.

- Spörlein, C., and C. Kristen. 2019. "Why We Should Care About Regional Origins: Educational Selectivity Among Refugees and Labor Migrants in Western Europe." *Frontiers in Sociology* 4: 1–18. doi:10.3389/fsoc.2019.00039.
- Spörlein, C., C. Kristen, R. Schmidt, and J. Welker. 2020. "Selectivity Profiles of Recently Arrived Refugees and Labour Migrants in Germany." *Soziale Welt* 71 (1–2): 54–89. doi:10.5771/0038-6073-2020-1-2-54.
- Stevens, G. 1999. "Age at Immigration and Second Language Proficiency among Foreign-Born Adults." *Language in Society* 28 (04): 555–578. doi:10.1017/S0047404599004030.
- Trivedi, J. K. 2006. "Cognitive Deficits in Psychiatric Disorders: Current Status." *Indian Journal of Psychiatry* 48 (1): 10–20. doi:10.4103/0019-5545.31613.
- UNESCO. 2020. *Syrian Refugee Youth Literacy Assessment Study*. Beirut: UNESCO Regional Bureau for Education in the Arab States. https://en.unesco.org/sites/default/files/literacy_assessment_of_syrian_youth_refugees-revised_2020_final_2509072_0.pdf.
- UNHCR. 2017. *Global Trends. Forced Displacement in 2016*. Genf: UNHCR The UN Refugee AgencyCR. <https://www.unhcr.org/5943e8a34.pdf>.
- UNHCR. 2020. *Trends at a Glance: Global Trends Forced Displacement in 2019*. Genf: UNHCR The UN Refugee Agency. <https://www.unhcr.org/5ee200e37.pdf>.
- van Buuren, S. 2012. *Flexible Imputation of Missing Data*. Boca Raton: Chapman & Hall/CRC.
- van Tubergen, F. 2010. "Determinants of Second Language Proficiency among Refugees in the Netherlands." *Social Forces* 89 (2): 515–534. doi:10.1353/sof.2010.0092.
- van Tubergen, F., and M. Kalmijn. 2005. "Destination Language Proficiency in Cross National Perspective: A Study of Immigrant Groups in Nine Western Countries." *American Journal of Sociology* 110 (5): 1412–1457. doi:10.1086/428931.
- The World Bank. 2020. "Literacy Rate, Adult Total (% of People Ages 15 and Above)." <https://data.worldbank.org/indicator/SE.ADT.LITR.ZS>.