



ARTICLE



<https://doi.org/10.1057/s41599-024-03034-0>

OPEN

Does timing matter? Language course participation and language outcomes amongst new immigrants

Jonas Wood ¹✉ & Dries Lens²

Host country language acquisition is widely considered an important factor to prevent the social and economic exclusion of immigrants. High-income countries invest considerably in language courses for new immigrants, yet hitherto our understanding of when participation in such courses seems to stimulate language acquisition remains limited. As a result, this article addresses whether the timing of language course enrolment is related to course outcomes in terms of credentialed language proficiency. This study benefits from longitudinal population data including all new migrants entering Flanders (Belgium) in 2009–2021, to assess whether associations between language course timing and outcomes persist, for different legal categories of immigrants, when controlling for compositional factors. Findings indicate that for most immigrant categories, negative associations between course postponement and host country language attainment are explained by migration background and human capital. However, amongst family and humanitarian migrants, a negative association between course timing and outcomes persists. This persistent negative linkage suggests a stronger vulnerability to less effective (informal) learning trajectories and downward spiralling opportunities and motivation in the case of language course postponement, but also highlights the need for future research using causal estimation strategies. As such, research on language course timing and outcomes contributes to the literature highlighting that timing – e.g. regarding asylum procedures, housing, but also language programmes – matters in immigrants’ trajectories of integration in host countries.

¹Sociology Department—Center for Population, Family and Health, University of Antwerp, Sint Jacobstraat 2, B-2000 Antwerpen, Belgium. ²Sociology Department—Center for Social Policy, University of Antwerp, Sint Jacobstraat 2, B-2000 Antwerpen, Belgium. ✉email: Jonas.Wood@uantwerpen.be

Introduction

The integration of newcomers is a multifaceted process which is typically evaluated based on positions in various societal domains, such as the labour market, education, housing, social contacts, or cultural participation. In the context of growing and increasingly diverse inflows of migrants in many high-income countries, policy-makers have put forward language acquisition as an essential requirement for integration in most – if not all – of these domains (OECD 2021; Hoehne and Michalowski 2016; Vlaamse Regering 2019; OECD 2015). It is widely accepted that supporting language acquisition is essential to prevent social and economic exclusion of new immigrants. In some Western European countries such as Belgium, large investments in language programmes for newcomers entering today contrast with the widely perceived policy failure of not providing comprehensive language courses for guest workers¹ and their family members in the 1950s to 1980s (Hoehne and Michalowski 2016; Timmerman 2017).

However, in contrast to policy-makers' focus on formal language training, available literature on host country language acquisition amongst immigrants pays relatively little attention to formal language programmes. Previous studies at most include programme participation as one of many factors determining one's post-migration exposure to the host country language, in combination with various indicators of efficiency of language acquisition and incentives to invest in language learning (van Tubergen and Wierenga 2011; Gonzalez 2000; Hayfron 2001; Beenstock 1996). This knowledge gap is remarkable as formal language training is likely to provide more structured and validated language acquisition compared to language acquisition in informal contexts, and formal language credentials play an important role on resumes of immigrant job applicants.

This study aims to enhance our understanding of the conditions under which formal language training programmes seem to facilitate language acquisition, focusing on the timing of participation in language courses. Drawing upon mechanisms studied in Linguistics as well as Sociology and Social psychology, available literature routinely assumes that formal language acquisition shortly upon arrival lays foundations for later language proficiency (Kosyakova et al. 2021; Kristen et al. 2016; Stevens 1999; Hoehne and Michalowski 2016). However this assumption is rarely tested (Hoehne and Michalowski 2016). As a result, this article exploits longitudinal individual-level population data to address the question of whether later enrolment into formal language programmes goes hand in hand with lower credentialed language acquisition.

As such, this study contributes to the available literature on immigrants' integration and host country language acquisition in two ways. First, the overwhelming majority of studies of language acquisition amongst migrant groups that do take into account formal language training simply take into account whether an individual has taken up or completed any language training (van Tubergen and Wierenga 2011; Gonzalez 2000; Hayfron 2001; Beenstock 1996). Such a static account of the potential role played by participation in language courses contrasts with the life course principle of "timing", implying that the consequences of events vary according to their timing in a person's life (Elder et al. 2003). This limited understanding of whether timing of language courses matter (Hoehne and Michalowski, 2016) is remarkable given the available literature focussing on detrimental effects of other types of delay in integration processes, such as lengthy asylum procedures (Kosyakova and Brenzel 2020; Jutvik and Robinson 2020; Hvidtfeldt et al. 2018; Hainmueller et al. 2016). The only study assessing the importance of timing of course enrolment by

Hoehne and Michalowski (2016) focusses on pre-1975 Turkish and Moroccan migrants and their descendants. Since the authors argue that their findings cannot be generalised to more recent migrants, this study is the first to address the potential importance of the timing of language programme enrolment host country language proficiency.

Second, the usage of exceptionally rich Flemish (Belgian) longitudinal microdata drawn from population-wide registers covering immigrants arriving between 2009 and 2021, allows us to overcome many limitations witnessed by the large body of cross-sectional and/or survey studies (Wood and Neels 2020). Unlike previous research (Hoehne and Michalowski 2016), the longitudinal data allows us to distinguish pre-enrolment language acquisition from progress in language skills since enrolment, and also provides a more refined account of language acquisition over time compared to panel data with gaps in between waves of data collection often amounting up to several years (Kristen et al. 2016; Adamuti-Trache 2013). The scale of population data allows for a more extensive test of the association between language course timing and outcomes (Hoehne and Michalowski 2016), by controlling for a wide range of compositional differences, and by testing population heterogeneity in the association of interest by legal migration category. Furthermore, the official registration of language course outcomes in terms of language certificates in Register data contrasts with arguably unreliable self-assessments of objective language skills (Dustmann and Soest 2001; Charette and Meng 1994).

Language training for newcomers in flanders

During the 2009–2021 observation period, the annual number of immigrants arriving in Belgium is relatively stable and close to 160,000, with a short peak of 174,591 in 2019 and a low of 144,169 during the onset of the covid-19 pandemic in 2020. The Flemish rate of international immigration relative to population size is close to the EU-average, however a large number of international migrants arrives in Flanders via the Brussels region (Statistics Belgium 2023; Statistics Flanders 2023). Available documentation of 18–54 year old immigrants' legal migration category for 2010–2016 indicates that about half of immigrations to Flanders is legally categorized as educational (15 percent amongst 18–29 year olds, 2–3 percent for 30–54 year olds) or labour migration (35 and 50 percent respectively). In both age groups, the share of family migration approximates 35 percent, whereas the share of humanitarian migration is close to 10 percent. Immigrations categorized as regularisation² and other or unknown contribute less than 10 percent of all immigrations (UNIA and FOD WASO 2017).

From an international perspective, in 2020 Flanders exhibits a relatively generous system of language training for immigrants, as all legal categories of migrants are eligible to publicly funded language courses, which is the case in ten out of 37 studied OECD countries (Denmark, Germany, Iceland, Italy, Latvia, Luxembourg, Portugal, Sweden, Switzerland). The Flemish system provides moderate flexibility with evening lessons and differentiated learning tracks, yet no childcare support (OECD 2021). An extremely small share of new migrants enters Flanders with Dutch language skills (OECD 2023), and evidence indicates that the share of non-European migrants reporting limited host country skills as the main obstacle to employment entry (16 percent) is amongst the highest in the EU (OECD 2023).

Since the 2000s, Flanders has installed the civic integration programme - including civic integration courses, guidance for labour market integration, and formal language training

(including literacy training if required) – which is mandatory for non-EU family and humanitarian migrants and optional for immigrants (and family members) from EU and European Free Trade Agreement (EFTA) countries, non-EU migrants who come for work or study, and long-term residents of EU and EFTA countries (see OECD 2023 for an overview). In addition, immigrants who are not obliged to enrol into the civic integration programme, can also participate in language courses outside the civic integration programme (Wood and Neels 2020).

Language training for new immigrants is monitored and organised by the Agency for Integration, in close collaboration with centres for adult education and centres for basic education. In this paper we define “formal language training” as all language courses that are provided by these institutions, which is largely comparable to language training for immigrants in other OECD countries (OECD 2021). Despite the existence of other private providers of language courses, it is noteworthy that only credentials from the aforementioned institutions are recognised in evaluating whether language requirements are met when applying for a Belgian nationality, social housing, or social assistance benefits (Blommaert 2011; OECD 2023). Furthermore, language certificates are widely known and thus play an important role for immigrant job applicants’ resumes. Previous research indicates that 80 percent of all vacancies in Flanders require (very) good Dutch language proficiency (OECD 2023). The formal language programmes studied in this article do not include additional language support in the rare context in which immigrants enrol internships (co-) organised by the public employment services (Kasztan Flechner et al. 2022; Wood and Neels 2020), nor language support in secondary schools which mostly serves minors.

Flemish language policy includes intake tests of Dutch proficiency to assign new migrants to the most appropriate language training, depending on literacy, language knowledge, and cognitive abilities. Language classes follow the language levels of the Common European Framework of Reference (CEFR). The latter firstly focus on basic interactions with support (i.e. A1-courses), then gradually increasing independence in specific situations (i.e. A2-courses), and later also providing skills to deal with most situations (i.e. B1- or higher-level courses) (Council of Europe, s.d.; Devlieger et al. 2014). During this study’s observation period, language courses are subsidized and free of charge³. The minimum level of Dutch that is aimed for and thus is used as an indicator of “a successful civic integration programme” (entailing a civic integration certificate) was A1 until 2014 and A2 thereafter.

Available evidence indicates large variation in the timing of participation in language courses between new migrants, which is routinely related to supply- and demand-side factors (Devlieger et al. 2014; OECD 2023). With respect to the supply, the timing of enrolment in language courses is related to the organisation of language training. As soon as the intake level is known, reservations are made for language courses geared towards the next level. However, available government documentation puts forward numerous supply-side factors which might affect the timing of enrolment, such as fixed starting dates for courses (e.g. penalising those who want to start during summer), or waiting lists particularly affecting migrants who are not targeted as priority groups (e.g. labour migrants). With respect to demand-side factors, available documentation highlights housing and mobility problems, internal migration and legal procedures (e.g. asylum procedure), information deficits, and lack of childcare as common reasons to postpone language course attendance (Devlieger et al. 2014; OECD 2023).

Theory

Exposure, efficiency and incentives. Most empirical contributions about migrants’ language acquisition rely on a three-pillar theoretical framework originally proposed by Chiswick and Miller (2001, 1995). The framework assumes that language acquisition is determined by (I) exposure, (II) efficiency, and (III) incentives.

First, exposure refers to passive or active learning opportunities. This pillar includes various indicators for pre-migration exposure such as formal host country language training in the origin country, pre-migration trips to the host country, or linguistic distance between the origin and destination country language (Kristen et al. 2016; van Tubergen 2010; Stevens, 1999; Kosyakova et al. 2021; Adamuti-Trache 2013; Espenshade and Fu 1997; Hwang and Xi 2008). In addition, post-migration exposure can refer to duration of stay in the host country, language course participation, partner’s language proficiency, the intensity of social interaction with native speakers, and the degree of segregation at the community level (Kristen et al. 2016; van Tubergen 2010; Stevens 1999; Kosyakova et al. 2021; Adamuti-Trache 2013; Espenshade and Fu 1997; Hwang and Xi 2008).

Second, efficiency is the extent to which exposure translates into fluency (Chiswick and Miller 2001). Commonly used measures include education, literacy, age at migration as younger individuals generally learn languages more easily, gender as women adopt a wider range of language learning strategies, and linguistic distance between origin and destination country language (Kosyakova et al. 2021; Chiswick and Miller 1995; Chiswick and Miller 2001; Kristen et al. 2016; van Tubergen and Wierenga 2011).

Third, incentives concern the expected returns to versus the costs of language investments. The latter includes direct costs in terms of tuition fees or study materials and indirect opportunity costs of the time spent on language learning (van Tubergen 2010). Returns include easier access to host country human capital and increased employment/wage potential (Wood and Neels 2020) but also non-economic factors such as inter-ethnic contact and increased feeling of belonging to the host society (Espenshade and Fu 1997). The perceived returns depend on individual features such as the expected length of stay or attitudes towards integration, and macro-level factors like the perceived orientation of locals toward diversity, political regimes and tolerance towards other languages (Chiswick and Miller 2001; Kristen et al. 2016; Kosyakova et al. 2021; Adamuti-Trache 2013; Chiswick and Miller 2002; Mesch 2003; van Tubergen and Kalmijn 2005).

The effect of timing of enrolment. It should be noted that hitherto research using this three-pillar framework has been unable to distinguish effects of the aforementioned factors from one another, not only as a result of indirect measurements (e.g. education as indicator for exposure and efficiency), but also due to their undeniably endogenous relationship (e.g. higher efficiency might boost incentives) (e.g. van Tubergen and Wierenga 2011). Consequently, theorising about migrants’ acquisition of language credentials should not only feature exposure, efficiency and incentives as the main concepts, but also take into account their interplay over time in a more dynamic way. Furthermore, exposure, efficiency and incentives are also likely to vary jointly depending on the timing of formal training participation. As a result, we suggest three potential mechanisms through which the timing of formal training participation is assumed to influence the outcome of formal training trajectories in terms of language credentials.

First, the timing of formal training participation determines the sequencing of exposure in formal training versus informal

exposure in daily life. Following linguistic literature, we assume that relying solely on formal training is inefficient due to limited opportunities for practice, whereas exclusive reliance on informal practice unlikely transcends basic levels of proficiency. Early formal language training provides structured and validated knowledge, which can be practiced in subsequent daily situations (Hoehne and Michalowski 2016; Mehlem 2003). This implies that early timing of participation in formal language programmes is likely to yield more favourable outcomes. In addition, the fact that language proficiency will halt at a level that precludes most meaningful social contacts when exclusively relying on informal practice, in turn might limit social opportunities to practice the host country language (Schuller, 2011). Early enrolment into formal language training might prevent this negative feedback effect between proficiency and exposure.

Second, the timing of formal language training is also likely to affect learning outcomes through efficiency. Available literature argues that adults who learn a new language typically limit themselves to the linguistic strategies they first acquired (i.e. a language learning plateau⁴). This implies that later enrolment into language programmes might entail lower-quality linguistic strategies, or even stabilised errors and systematic usage of erroneous forms accumulated through unstructured and unguided informal practice (Hoehne and Michalowski 2016). Available literature distinguishes phonological language learning plateaus (i.e. pronunciation), from morphological plateaus (i.e. forms of and relation between words), syntactic plateaus (i.e. how linguistic elements such as words form constituents), and semantic or pragmatic plateaus (i.e. meaning of words and constituents in a given context) (Wei 2008). Language skill plateaus have been argued to create a limit to the development of linguistic competences, which is cumbersome to remediate afterwards in formal training, yet can be prevented by early enrolment in language courses (Wei 2008; Hoehne and Michalowski 2016).

Third, the timing of formal language training is also likely to affect learning outcomes through incentives and motivation. Social psychology literature and Sociological theories of reactive ethnicity highlight the recursive relation between integration outcomes and motivation (Portes and Zhou 1993; Hoehne and Michalowski 2016). In the absence of formal training, migrants are at risk of experiencing negative vicious circles including lacking proficiency, decreasing motivation, and confirmed perceptions of limited integration (Pozzo 2022; Temple 2010; Norton 1997). In contrast, early enrolment in formal language training programmes potentially provides an early signal that new migrants are accepted to participate in language training, but also in wider society, in turn boosting migrants' self-confidence and feeling of belonging (Hoehne and Michalowski 2016). This reasoning also aligns with documented positive effects of brief social belonging interventions at the start of minority college students' academic career (Esser 2006; Schuller 2011; Hoehne and Michalowski 2016; van Tubergen and Kalmijn 2008).

As a result of the aforementioned mechanisms, we hypothesise that *migrants who start participating in language training later will be less likely to exhibit progression in credentialed language skills*.

(Self-)selection. In addition to potential causal effects of the timing of language training, mechanisms of (self-)selection into early enrolment might also entail associations between language course timing and outcomes. We distinguish six broad categories of potential (self-)selection mechanisms. The first group is related to individuals' migration background. Documented mechanisms such as waiting lists penalising non-priority groups who are not

obliged to participate, like labour migrants or EU-citizens, yield variation in timing by migration background, and as waiting lists have gradually been eliminated over time, migrants arriving later are more likely to enrol sooner. To the extent that migration background also affects the outcome of language course participation (e.g. lower incentive for labour migrants who plan to return), these mechanisms might yield a spurious relationship between language course timing and outcomes. Second, an association between language course timing and outcomes might also be misinterpreted as a causal effect due to demographic compositional differences between early and later participants. For instance, older immigrants' postponement and lower language course outcomes might both be driven by hesitations related to efficiency and use of the Dutch language for them. Similarly, childcare duties and cumbersome access to formal childcare (Biegel et al. 2021; Maes et al. 2023) have been documented as a factor postponing enrolment (Devlieger et al. 2014), and might also affect the later performance in language courses. Third, in line with documented problems in terms of information deficits amongst migrants (Devlieger et al. 2014), it is likely that migrants with stronger learning capabilities (e.g. higher education) will navigate the system more swiftly, and also exhibit higher language course outcomes. Fourth, the degree of labour market integration potentially affects the incentive to start language training as well as incentives to perform well (e.g. a job in an Dutch-speaking workplace).

The fifth and sixth categories of potential selection mechanisms concern concepts which are typically more cumbersome to measure. The fifth category includes spuriousness due to the impact of social integration on the necessity to enrol swiftly and perform well in Dutch language courses, for instance in the case longer residing family members can provide Dutch language lessons. The sixth category includes ideational mechanisms such as the degree to which immigrants value host country language acquisition, which also potentially drives variation in language course timing and outcomes.

Data and methods

This study uses unique longitudinal population data covering all immigrants who entered Belgium in 2009–2021 and subsequently resided in Flanders. The source is the Crossroads Bank for Civic Integration, which is the digital monitoring and tracking instrument used by the Flemish Agency for Integration (“Agentschap Inburgering”) for intake meetings, language class reservation, and coaching of new immigrants (Devlieger et al. 2014). The data contain 117,818 new immigrants who exhibit language course participation at some point. We exclude educational migrants who are likely to acquire unobserved language skills in the educational system. We also exclude a very small group of migrants holding a B1 or even higher level certificate at intake. The resulting analytical sample consists of 114,022 new immigrants, whose language level is typically tested at intake to determine the most appropriate starting course, as well as at the end of a module. This analytical sample generates 6,899,353 monthly observations.

The dependent variable is language course participants' progression to the next language certificate level. Our data allows us to distinguish four levels: (I) no certificates, (II) A1-level, (III) A2-level, and (IV) B1 or higher levels. As the number of levels one can progress depends on the intake level, all models are stratified by intake level (i.e., models A, B and C). Our sample includes 88,845 migrants without language credentials at intake, 14,793 with an A1 credential, and 10,384 with an A2 credential at intake. Ordered logit models are used to estimate progress in migrants' language certificate level over time, as a function of the timing of

Table 1 Nested model specifications: included independent variables of interest and control variables.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Timing and time since enrolment						
Months since start participation	☑	☑	☑	☑	☑	☑
Timing of participation	☑	☑	☑	☑	☑	☑
Migration background						
Legal type of migration		☑				☑
Country of origin		☑				☑
Year of arrival		☑				☑
Demographic						
Gender			☑			☑
Children			☑			☑
Age at arrival			☑			☑
Human capital						
Educational attainment				☑		☑
Learning capacity				☑		☑
Language knowledge				☑		☑
Labour market attachment						
Labour market status at arrival					☑	☑
First contact with Employment Office					☑	☑

language course participation. Exponentiated ordered logit parameters are interpreted as the odds-ratio of having progressed k levels versus having progressed less than k levels in which k stands for the cut-off between different levels of Dutch language credentials.

For every intake level, the same sequence of six nested model specifications is used, as illustrated in Table 1. All models include *time since enrolment* in language courses using a cubic specification, and *timing of participation*, a variable scaled to indicate 6-month shifts in timing, which aligns realistically with the variation in timing observed in the data. In contrast to Model 1, Models 2–6 add additional covariates to test whether the association between language course timing and outcomes can be statistically explained in terms of compositional differences.

Controls for migration background include *legal migration type* (labour, family, humanitarian, regularisation, other or unknown), *region of origin* (EU15, EU10, EU3, Other European, North African, Other African, North American, Other American, Asian, Australian, Unknown), and *year of arrival* (linear). Demographic controls are *gender* (female versus male), *children* (parent versus childless), and *age at arrival* (18–29, 30–39, 40–49, 50+). The adopted controls for human capital include *educational attainment at intake* (low (ISCED 0–2), medium (ISCED 3–4), high (ISCED 5–6), unknown), *learning capacity* as measured through intake tests (capacity sufficient for enrolment in basic education, long adult education, standard adult education, short-track adult education, unknown), as well as an extensive set of separate dummies indicating self-reported *language knowledge* at arrival (i.e., Dutch, English, French, German, Italian, Greek, other Romance or Basque languages, other Germanic languages, Baltic Slavic Paleo-Siberian and Ural languages, other Indo-European languages, Arabic, Berber languages, other Afro-Asian languages, Turkish, other Asian languages). Controls for labour market attachment are *labour market status at intake at the agency for integration* (not employed nor studying, employed or studying, unknown), and a time-varying indicator for *first registration as jobseeker at the employment office* (no contact, max 1 year ago, max 2 years ago, max 4 years ago, longer ago). Table A1 in appendix provides an overview of the control variable distributions, as well as mean durations since enrolment in language courses, and mean progression of Dutch language level.

It is noteworthy that all control variables are either time-constant (e.g. region of origin), or measured at arrival or intake by

the agency for integration, which implies that they cannot be affected by language course outcomes. The indicator for first registration at the employment office is the only exception.

Finally, additional tests of heterogeneity (model 7) will be performed to assess to what degree the association between timing of language course participation and subsequent outcomes vary depending on *legal migration type*, after controlling for composition in terms of migration background, demographic characteristics, human capital, and labour market attachment.

Results

Timing of language course participation. Before addressing the association between the timing of language course participation and course outcomes in terms of Dutch language credentials, Fig. 1 illustrates considerable variation in the timing of course enrolment amongst new migrants in Flanders, regardless of whether the group without, or with a starting level of Dutch proficiency is considered. This variation will be exploited in multivariate models of enrolment timing and course outcomes.

Timing of language course participation and subsequent outcomes. The main results are presented in Tables 2–4 which exhibit estimated associations between timing of course enrolment and all covariates on the one hand, and language course outcomes on the other, in terms of odds-ratios.

Regarding migrants with sub-A1 Dutch language proficiency at intake (Table 2), which is the overwhelming majority, results (model A1) indicate that a six month delay in language training enrolment associates with an $(1-(0.885)^{*100})$ 11.5 percentage decrease in the odds of subsequent progression in language level. A more extreme, yet routinely occurring, difference in timing of course enrolment of 24 months associates with an 38.7 percentage decrease in the odds of progression. Whereas migrants with A1-level proficiency at intake exhibit a similar association (Table 3, model B1), the association between language course timing and outcomes is considerably weaker amongst the smaller group of migrants with A2-level proficiency at intake (Table 4, model C1). These findings imply that there is a negative association between language course postponement and outcomes, in particular for the largest group that starts without any Dutch language skills. In line with the theoretical framework (see Section 3.2) this association might be driven by the fact that early formal training

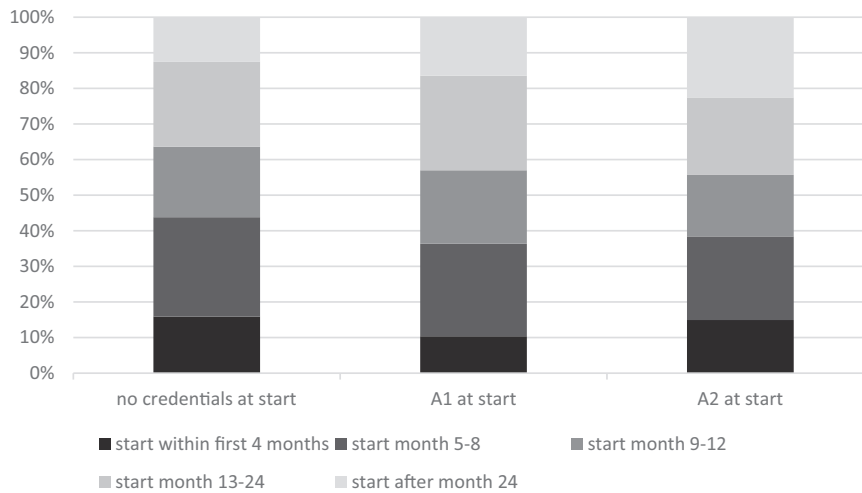


Fig. 1 Timing of language course enrolment in months since arrival, by starting level of Dutch proficiency, Flanders 2009–2021 immigrant cohorts.

Table 2 Progression in Dutch language credentials^a amongst immigrants without Dutch language credentials at the start of language course participation, Flanders 2009–2021.

	Model A1		Model A2		Model A3		Model A4		Model A5		Model A6	
	OR	sig	OR	sig	OR	sig	OR	sig	OR	sig	OR	sig
Months since start participation												
Linear	1.110	***	1.112	***	1.112	***	1.133	***	1.098	***	1.127	***
Quadratic	0.999	***	0.999	***	0.998	***	0.998	***	0.998	***	0.998	***
Cubic	1.000	***	1.000	***	1.000	***	1.000	***	1.000	***	1.000	***
Timing of participation (effect of 6 month postponement)												
6 month delay	0.885	***	0.937	***	0.888	***	0.935	***	0.879	***	0.980	***
Type of migration (Labour migration is reference)												
Family			1.408	***							1.223	***
Humanitarian			1.017								1.323	***
Regularisation			0.914								1.053	
Other/Unknown			0.757	***							0.735	***
Country of origin (xxx is reference)												
EU10			1.414	***							1.327	***
EU3			1.207	***							1.329	***
Other European			0.970								1.530	***
North African			0.772	***							1.190	*
Other African			1.019								1.565	***
North American			1.075								0.867	
Other American			1.111	*							1.302	***
Asian			1.224	***							1.790	***
Australian			0.938								0.695	
Unknown			2.037	***							2.272	***
Year of arrival												
linear			1.104	***							1.081	***
Gender (male is reference)												
female					1.381	***					1.618	***
Children (childless is reference)												
parent					0.999						0.982	
Age at arrival (18–29 is reference)												
30–39					0.898	***					0.954	**
40–49					0.620	***					0.856	***
50+					0.341	***					0.673	***
Educational attainment (low is reference)												
medium							1.114	***			1.245	***
high							1.735	***			1.831	***
unknown							0.915				0.997	

Table 2 (continued)

	Model A1		Model A2		Model A3		Model A4		Model A5		Model A6	
	OR	sig	OR	sig	OR	sig	OR	sig	OR	sig	OR	sig
Learning capacity												
Basic education							3.238	***			3.065	***
long adult education							9.221	***			7.890	***
standard adult education							13.636	***			12.670	***
short adult education							20.519	***			18.526	***
unknown							1.480	***			1.484	***
Language knowledge												
Dutch							1.522	***			1.432	***
English							1.402	***			1.412	***
French							0.763	***			0.845	***
German							0.905				1.042	
Italian							0.836	***			1.021	
Greek							0.864				1.000	
Other Romance or Basque							0.865	***			1.053	
Other Germanic							0.560	***			0.779	
Baltic, Slavic, Pal-Siberian,Ural							0.821	***			1.042	
Other Indo-European							1.253	***			1.103	***
Arabic							1.472	***			1.252	***
Berber							0.753	***			1.120	**
Other Afro-Asian							2.038	***			1.513	***
Other African							1.188	***			0.940	
Turkish							0.786	***			0.797	***
Other Asian							1.220	***			1.082	*
Labour market status at arrival (not working, nor studying is reference)												
employed or studying									0.872	***	0.787	***
unknown									1.133	***	1.350	***
First contact with Employment Office (no contact is reference)												
max 1 year ago									2.378	***	2.224	***
max 2 year ago									2.701	***	2.573	***
max 4 year ago									2.847	***	2.586	***
longer ago									3.529	***	3.206	***
Model parameters												
Cut 1	0.825	***	1.746	***	0.874	***	3.737	***	1.068	***	5.464	***
Cut 2	1.674	***	2.634	***	1.739	***	4.759	***	1.975	***	6.589	***
Cut 3	3.358	***	4.734	***	3.451	***	6.721	***	3.770	***	8.735	***
N Persons	88,845		88,845		88,845		88,845		88,845		88,845	
N Person-periods	5,409,894		5,409,894		5,409,894		5,409,894		5,409,894		5,409,894	

Significance levels: $p < 0.050$ (*), $p < 0.010$ (**), $p < 0.001$ (***). Source: Crossroads Bank for Civic Integration (KBI).
 *Exponentiated coefficients (odds ratios [OR]) from ordered logit models.

facilitates immigrants to reap the benefits from informal learning opportunities, prevents language learning plateaus, and potential negative impacts on motivation.

However, associations between language course timing and outcomes also potentially reflect mechanisms of (self-)selection (see section “(Self-)selection”). Results of models 2–6 indicate that the aforementioned associations between the timing of course enrolment and subsequent outcomes change considerably when controlling for compositional differences. With respect to migration background, besides other differentiating factors such as region of origin and year of arrival, regression results (Tables 2–4) consistently indicate that family migrants (model A2-C2), as well as humanitarian migrants with A2-level proficiency at intake (model C2), are more likely to exhibit progression in Dutch language credentials and enrol in language courses earlier (cf. Table A1 in appendix) compared to other migrant types. When controlling for composition in terms of migration background, the aforementioned 11.5 percentage disadvantage of a six-month enrolment has weakened to a 6.3 percentage disadvantage amongst the large group of migrants with sub-A1 proficiency at intake. The group with A1-level intake proficiency exhibits a similar drop in the disadvantage associated with postponement when controlling for migration background, whereas the smaller

group with A2-level intake proficiency exhibits an increase in the postponement-induced disadvantage in language acquisition. These findings imply that – except for the small group with A2-level intake levels – a considerable part of the negative association between language course timing and outcomes is explained by variation in migration background.

Regarding the impact of demographic composition, results consistently indicate that female migrants and younger migrants are considerably more likely to progress to higher language levels (model A3-C3, Tables 2–4). However, as these demographic groups do not systematically enrol in language courses earlier (cf. Table A1 in appendix), controlling for demographic composition does not alter the association between language course timing and outcomes, as illustrated by similar odds-ratios in model 1 and model 3.

With respect to human capital, regression parameters indicate clear positive educational gradients, clear positive effects of learning capacity, and differentiation depending on knowledge of other languages at arrival (model A4-C4, Tables 2–4). In addition, the negative association between timing of enrolment and course outcomes as identified in the first model weakens considerably when including the aforementioned indicators of human capital, particularly for the majority group of migrants without any Dutch

Table 3 Progression in Dutch language credentials^a amongst immigrants with A1 Dutch language credentials at the start of language course participation, Flanders 2009–2021.

	Model B1		Model B2		Model B3		Model B4		Model B5		Model B6	
	OR	sig	OR	sig	OR	sig	OR	sig	OR	sig	OR	sig
Months since start participation												
<i>Linear</i>	1.097	***	1.097	***	1.112	***	1.104	***	1.092	***	1.100	***
<i>Quadratic</i>	0.998	***	0.999	***	0.998	***	0.998	***	0.998	***	0.999	***
<i>Cubic</i>	1.000	***	1.000	***	1.000	***	1.000	***	1.000	***	1.000	***
Timing of participation (<i>effect of 6 month postponement</i>)												
<i>6 month delay</i>	0.873	***	0.907	***	0.888	***	0.906	***	0.889	***	0.946	***
Type of migration (Labour migration is reference)												
<i>Family</i>			1.196	*							0.982	
<i>Humanitarian</i>			1.031								0.989	
<i>Regularisation</i>			0.498								0.605	**
<i>Other/Unknown</i>			0.890	***							0.840	
Country of origin (<i>xxx is reference</i>)												
<i>EU10</i>			1.355	**							1.417	**
<i>EU3</i>			1.473	***							1.569	***
<i>Other European</i>			1.216	*							1.783	***
<i>North African</i>			0.782	*							1.159	
<i>Other African</i>			1.088								1.699	***
<i>North American</i>			1.060								1.376	
<i>Other American</i>			0.837								1.339	*
<i>Asian</i>			1.354	**							1.792	***
<i>Australian</i>			1.004								1.241	
<i>Unknown</i>			2.002	***							2.493	***
Year of arrival												
<i>linear</i>			1.073	***							1.084	***
Gender (<i>male is reference</i>)												
<i>female</i>					1.687	***					1.723	***
Children (<i>childless is reference</i>)												
<i>parent</i>					0.976						0.961	
Age at arrival (<i>18–29 is reference</i>)												
<i>30–39</i>					0.899	**					0.910	*
<i>40–49</i>					0.770	***					0.863	**
<i>50+</i>					0.632	***					0.741	**
Educational attainment (<i>low is reference</i>)												
<i>medium</i>							1.088				1.170	**
<i>high</i>							1.554	***			1.547	***
<i>unknown</i>							1.052				1.133	
Learning capacity												
<i>Basic education</i>							1.313				1.286	
<i>long adult education</i>							2.155	***			2.075	***
<i>standard adult education</i>							2.837	***			2.669	***
<i>short adult education</i>							3.850	***			3.532	***
<i>unknown</i>							0.280	***			0.270	***
Language knowledge												
Dutch							0.866	*			0.893	
English							1.261	***			1.297	***
French							0.747	***			0.862	*
German							0.841				0.955	
Italian							1.063	***			1.188	
Greek							1.322				1.586	*
Other Romance or Basque							1.094				1.229	**
Other Germanic							0.486	***			0.581	*
Baltic, Slavic, Pal-Siberian,Ural							1.024				1.085	
Other Indo-European							1.024				0.940	
Arabic							1.242	***			1.237	**
Berber							0.609	***			0.864	
Other Afro-Asian							1.479	***			1.240	*
Other African							1.037				0.944	
Turkish							0.900				0.865	
Other Asian							1.215	*			1.045	
Labour market status at arrival (<i>not working, nor studying is reference</i>)												
<i>employed or studying</i>									0.616	***	0.743	***
<i>unknown</i>									1.079		1.412	***

Table 3 (continued)

	Model B1		Model B2		Model B3		Model B4		Model B5		Model B6	
	OR	sig	OR	sig	OR	sig	OR	sig	OR	sig	OR	sig
First contact with Employment Office (no contact is reference)												
max 1 year ago									1.885	***	1.824	***
max 2 year ago									1.982	***	2.029	***
max 4 year ago									1.780	***	1.871	***
longer ago									1.734	***	1.967	***
<i>Model parameters</i>												
Cut 1	0.325	***	1.013	***	0.521	***	1.779	***	0.561	***	3.219	***
Cut 2	2.039	***	2.776	***	2.264	***	3.627	***	2.331	***	5.183	***
N Persons	14,793		14,793		14,793		14,793		14,793		14,793	
N Person-periods	959,451		959,451		959,451		959,451		959,451		959,451	

Significance levels: $p < 0.050$ (*), $p < 0.010$ (**), $p < 0.001$ (***). Source: Crossroads Bank for Civic Integration (KBI).
 *Exponentiated coefficients (odds ratios [OR]) from ordered logit models.

language credentials at enrolment. This implies that a considerable part of the negative association between language course timing and outcomes could be mistakenly interpreted as a causal effect, which is actually driven by human capital, which affects both the timing of language course enrolment and subsequent progression in Dutch language level (cf. Table A1 in appendix).

Regarding labour market attachment, results for the three subgroups of migrants (model A5–C5, Tables 2–4) consistently indicate that migrants who were employed or studying at intake are less likely to exhibit progression in language credentials after enrolment, but also that contact with the employment agency positively associates with language course outcomes. Including composition in terms of labour market attachment does not seem to change the association between timing of enrolment and course outcomes, except for the smallest group with A2 intake level. These findings thus imply that labour market attachment is a clear predictor of language course outcomes, yet does not explain the negative association between timing and outcomes of language outcomes.

Finally, as comparisons between odds-ratios from different models have been shown prone to bias (Mood, 2009), and the previous results do not indicate how language course timing is related to progression between the different language levels over time, Figs. 2–3 provide a dynamic overview of the results in terms of Average Marginal Effects (AME) over time. Furthermore AME allow us to quantify the association in terms of the probability of holding a particular language credential, which is arguably easier to interpret than odds-ratios.

With respect to the largest group of participants with sub-A1 intake proficiency (Fig. 2), estimates (model 1) indicate that a six-month enrolment delay associates positively with no progression (Fig. 2), lower probabilities of swift progression to A1 proficiency (Fig. 2), and lower progression to A2 (Fig. 2), or higher language credentials (Fig. 2). These patterns of association between timing of enrolment and subsequent outcomes remain similar when controlling for demographic control variables (model 3) or labour market attachment (model 5). However, when controlling for either migration background (model 2) or human capital (model 4) the magnitude of the aforementioned associations is halved. When controlling for all the control variables combined (model 6) the initial association is weakened, mostly due to the compositional effects in terms of migration background and human capital.

Similar conclusions can be drawn regarding participants with A1 intake level, with higher probabilities of no progression (Fig. 3), but also lower short-term probabilities of progression to A2 proficiency (Fig. 3), and consistently lower progression to B1 or higher language credentials (Fig. 3) associated with enrolment

postponement (model 1). Again, the magnitude of the association halves when controlling for migration background (model 2) or human capital (model 4), and approximately one third of the gap persists when controlling for all covariates simultaneously.

Regarding the small group of participants with A2 intake level, findings similarly indicate that a six-month postponement of enrolment is negatively associated with further progression in proficiency (model 1). This association persists when controlling for demographic controls (model 3). However, in contrast to previous results for participants with lower intake levels, timing-related gaps widen when controlling for migration background (model 2) or human capital (model 4) and decrease moderately when controlling for labour market attachment (model 5). Consequently, the full model (model 6) illustrates a persistent postponement disadvantage regarding language course outcomes.

Association heterogeneity by legal migration type. Finally, results of analyses assessing variation in the association between course timing and outcomes by legal migration type (model 7) are provided in Table 5. Migrants with sub-A1 intake proficiency (model A7) indicate a $(1 - (1.036 * 100))$ 3.6 percentage advantage in the odds of (higher) progressions amongst labour migrants in case of a six month enrolment delay. In contrast, family and humanitarian migrants respectively exhibit a $(1 - (1.036 * 0.914) * 100)$ 5.3 percentage and 2.9 percentage decrease in the odds of (higher) progressions as a result of a six-month enrolment postponement. These associations are not negligible as a 24-month postponement – which occurs routinely in our data – is related to a $((1.036 * 0.914)^4 * 100)$ 19.6 percentage and a 11.1 percentage decrease in the odds of progress in language levels amongst family and humanitarian migrants. These findings suggest that – whilst labour migrants might benefit from language course postponement in terms of later language course outcomes (e.g. due to investments in their job) – the negative linkage between language course postponement and outcomes holds consistently for family and humanitarian migrants.

Differentiation in the association between timing and outcomes of language programmes by migration type is more limited amongst smaller groups with higher intake levels. However, family migrants again exhibit the strongest negative associations, with 8.3 and 19.9 percentage six-month postponement-related penalties respectively for those holding A1 and A2 language certificates at intake.

Robustness checks. Finally, we ran two robustness checks. First, as commonly suggested determinants of delays in

Table 4 Progression in Dutch language credentials^a amongst immigrants with A2 Dutch language credentials at the start of language course participation, Flanders 2009–2021.

	Model C1		Model C2		Model C3		Model C4		Model C5		Model C6	
	OR	sig	OR	sig	OR	sig	OR	sig	OR	sig	OR	sig
Months since start participation												
<i>Linear</i>	1.112	***	1.134	***	1.114	***	1.125	***	1.114	***	1.137	***
<i>Quadratic</i>	0.998	***	0.998	***	0.998	***	0.998	***	0.998	***	0.998	***
<i>Cubic</i>	1.000	***	1.000	***	1.000	***	1.000	***	1.000	***	1.000	***
Timing of participation (<i>effect of 6 month postponement</i>)												
<i>6 month delay</i>	0.914	***	0.802	***	0.916	***	0.907	***	0.939	***	0.821	***
Type of migration (Labour migration is reference)												
<i>Family</i>			1.262	*							1.042	
<i>Humanitarian</i>			1.576	***							1.240	
<i>Regularisation</i>			0.439	**							0.565	
<i>Other/Unknown</i>			1.035								0.989	
Country of origin (<i>xxx is reference</i>)												
<i>EU10</i>			1.589	***							1.331	
<i>EU3</i>			1.218								1.011	
<i>Other European</i>			1.016								1.184	
<i>North African</i>			0.546	***							0.872	
<i>Other African</i>			0.654	**							0.800	
<i>North American</i>			0.629	*							0.601	*
<i>Other American</i>			0.389	***							0.576	***
<i>Asian</i>			1.049								1.390	*
<i>Australian</i>			0.464								0.654	
<i>Unknown</i>			1.492	**							1.853	**
Year of arrival												
<i>linear</i>			0.939	***							0.944	***
Gender (<i>male is reference</i>)												
<i>female</i>					1.612	***					1.571	***
Children (<i>childless is reference</i>)												
<i>parent</i>					0.773						0.898	
Age at arrival (<i>18–29 is reference</i>)												
<i>30–39</i>					0.813	***					0.915	
<i>40–49</i>					0.625	***					0.792	*
<i>50 +</i>					0.433	***					0.579	**
Educational attainment (<i>low is reference</i>)												
<i>medium</i>							1.444	***			1.218	
<i>high</i>							1.757	***			1.570	***
<i>unknown</i>							1.395				1.085	
Learning capacity												
<i>Basic education</i>							28.725	**			35.784	**
<i>long adult education</i>							85.318	***			113.179	***
<i>standard adult education</i>							187.664	***			222.987	***
<i>short adult education</i>							213.310	***			259.733	***
<i>unknown</i>							11.938	*			21.608	**
Language knowledge												
Dutch							1.100				1.316	***
English							1.217	**			1.260	***
French							0.853				0.885	
German							1.167				1.144	
Italian							0.818				0.907	
Greek							1.597				1.843	*
Other Romance or Basque							1.125				1.367	**
Other Germanic							0.577	*			0.784	
Baltic, Slavic, Pal-Siberian,Ural							1.265	**			1.062	
Other Indo-European							1.049				0.751	**
Arabic							0.995				1.001	
Berber							0.755				0.922	
Other Afro-Asian							1.202				1.531	*
Other African							1.056				1.112	
Turkish							0.958				0.997	
Other Asian							1.115				0.714	*
Labour market status at arrival (<i>not working, nor studying is reference</i>)												
<i>employed or studying</i>									0.560	***	0.624	***
<i>unknown</i>									1.237	***	0.953	***

Table 4 (continued)

	Model C1		Model C2		Model C3		Model C4		Model C5		Model C6	
	OR	sig	OR	sig	OR	sig	OR	sig	OR	sig	OR	sig
First contact with Employment Office (no contact is reference)												
max 1 year ago									1.157	**	1.513	***
max 2 year ago									1.347	***	1.937	***
max 4 year ago									1.068		1.688	***
longer ago									0.895		1.587	***
<i>Model parameters</i>												
Cut 1	2.212	***	0.966	***	2.337	***	7.920	***	2.318	***	7.304	***
N Persons	10,384		10,384		10,384		10,384		10,384		10,384	
N Person-periods	530,008		530,008		530,008		530,008		530,008		530,008	

Significance levels: $p < 0.050$ (*), $p < 0.010$ (**), $p < 0.001$ (***) Source: Crossroads Bank for Civic Integration (KBI).
^aExponentiated coefficients (odds ratios [OR]) from ordered logit models.

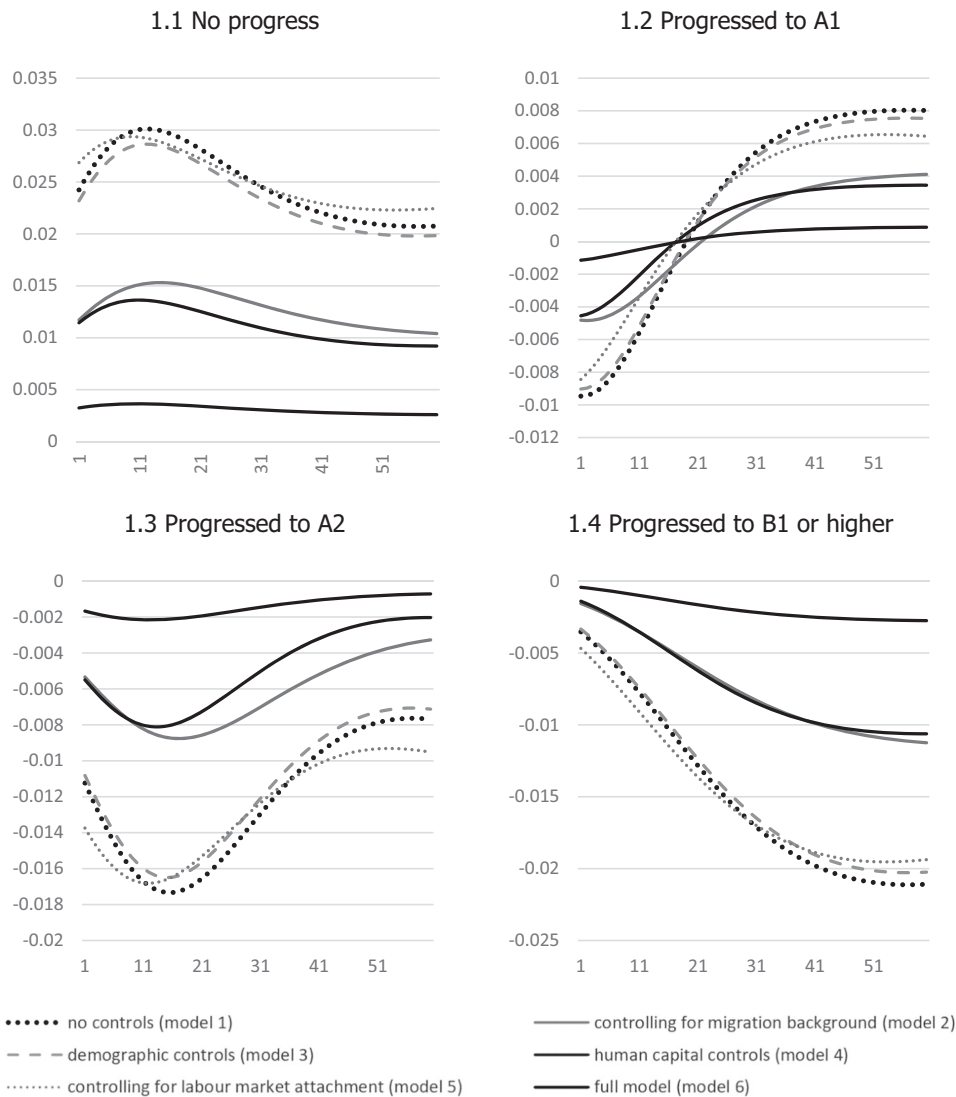


Fig. 2 Effect of 6 month postponement of start language course attendance on probability of progression in credentialed Dutch proficiency in terms of Average Marginal Effects (y-axis) over months since start participation (x-axis), models A1-A6, Flanders 2009-2021 immigrant cohorts without any Dutch language credentials at the start of language course participation.

language course participation are likely to differ between legal migration categories (e.g. the importance of language for labour market integration), it is also possible that the association between control variables on the one hand and

language course outcomes on the other vary by legal migration category (e.g. the effect of labour market status at arrival). As this might induce bias in the estimates of association between the timing and outcomes of language courses, we ran the full

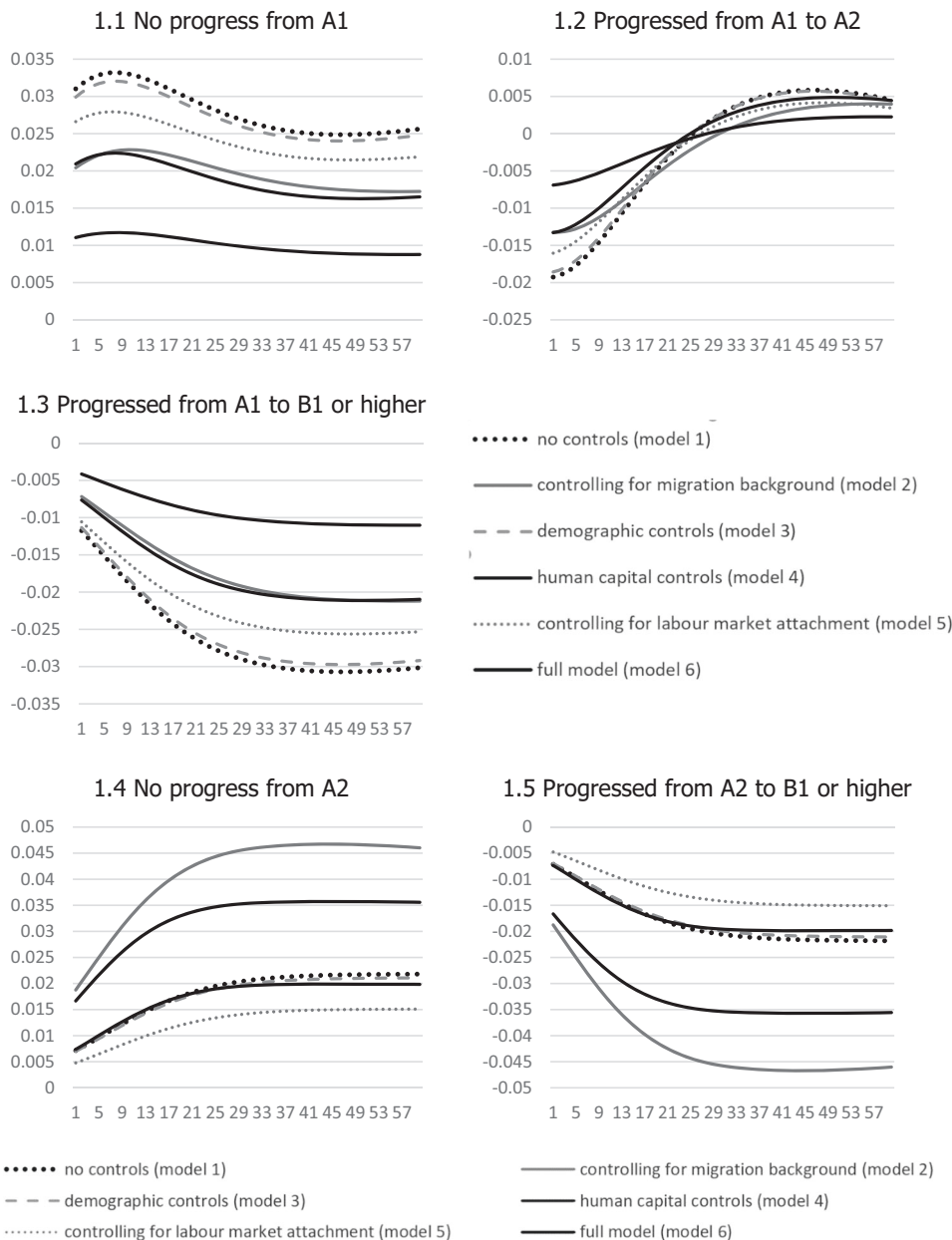


Fig. 3 Effect of 6 month postponement of language course attendance on probability of progression in credentialed Dutch proficiency in terms of Average Marginal Effects (y-axis) over months since start participation (x-axis), models A1-A6, Flanders 2009–2021 immigrant cohorts with Dutch language credentials at the start of language course participation.

model stratified by legal migration category as a robustness check. With respect to migrants without Dutch language credentials at the start of course participation – the overwhelming majority of our sample – estimated associations between timing of enrolment and language course outcomes are similar to results presented in Table 5 for labour migrants (1.036 versus 1.044) family migrants (0.947 versus 0.952) humanitarian migrants (0.970 versus 0.941) regularised migrants (1.023 versus 1.027), and migrants with other or unknown legal status (0.984 versus 0.981). This robustness check could not be reliably performed for groups with A1 or A2 credentials at the start of language course participation due to low cell frequencies.

Second, in contrast to all other control variables which are either time-constant or measured at the start of the language course participation, the included indicator for first contact with

the employment office is time-varying. As a result, it is possible that this variable also captures part of the effect of language course timing (e.g. increased language proficiency stimulates contact with the employment office, which in turn might affect credentialed language proficiency as a result of higher exposure and incentives). Consequently, all models were also estimated excluding this variable. In line with the finding that controlling for labour market attachment does not alter the association between language course timing and outcomes, this robustness check also did not change the main findings of this study.

Discussion and conclusion

In contrast to longer traditions of addressing migrants’ language acquisition in the United States or Canada (e.g. Alba et al. 2002; Carliner 2000; Chiswick and Miller 2001), patterns and underlying dynamics of migrants’ language acquisition have only

Table 5 Progression in Dutch language credentials depending on timing of language course participation, type of migration, and interactions^a, Flanders 2009–2021.

	Model A7 (no credentials at start)		Model B7 (A1 credential at start)		Model C7 (A2 credential at start)	
	OR	sig	OR	sig	OR	sig
Months since start participation						
Linear	1.127	***	1.099	***	1.137	***
Quadratic	0.998	***	0.998	***	0.997	***
Cubic	1.000	***	1.000	***	1.000	***
Timing of participation (effect of 6 month postponement)						
6 month delay	1.036	***	0.981		0.849	***
Type of migration (Labour migration is reference)						
Family	1.523	***	1.203		1.289	
Humanitarian	1.593	***	1.098		1.443	*
Regularisation	1.077		0.640	*	0.580	
Other/Unknown	0.853	**	1.023		1.147	
Type of migration*Timing of participation						
Family	0.914	***	0.935	***	0.943	**
Humanitarian	0.937	***	0.982		0.971	
Regularisation	0.987		0.989		1.002	
Other/Unknown	0.950	**	0.932	*	0.965	
Model parameters						
N Persons	88,845		14,793		10,384	
N Person-periods	5,409,894		959,451		530,008	

Significance levels: $p < 0.050$ (*), $p < 0.010$ (**), $p < 0.001$ (***) Source: Crossroads Bank for Civic Integration (KBI).

^aExponentiated coefficients (odds ratios [OR]) from ordered logit models All models include all control variables regarding migration background, demographic characteristics, human capital and labour market attachment.

recently gained similar attention in European countries (e.g. Kosyakova et al. 2021; Kristen and Sueuring 2021; van Tubergen and Wierenga 2011; Auer 2018). This study exploits Flemish longitudinal individual-level population data to assess whether patterns of credentialed language acquisition amongst new migrants vary depending on the timing of course participation, a question which has only been put to the test by Hoehne and Michalowski (2016), studying pre-1975 Turkish and Moroccan migrants and their descendants in Western Europe.

At first sight, the results of this study corroborate those of Hoehne and Michalowski (2016), as postponement of enrolment is negatively associated with language course outcomes in terms of acquiring Dutch language certificates. However, benefitting from a richer set of control variables, this study indicates that a large part of the differentiation in language course outcomes by timing of participation is due to composition in terms of migration background and human capital. The fact that the association does not disappear completely can however not be interpreted as hard causal evidence as not all potential mechanisms of (self-)selection (see section “(Self-)selection”) could be controlled for in this study. Potential mechanisms of selectivity due to variation in social networks, or ideational factors for instance could not be observed in the data at hand. Consequently, the main finding of this study is that a large part of the association between language course timing is spurious, not that the remaining association evidences a causal effect.

Available literature displays recent calls to compare different patterns of integration (Kogan and Kalter 2020; FitzGerald and Arar 2018), but also host country language acquisition between different subgroups of migrants (Kosyakova et al. 2021; Kristen and Sueuring 2021). For example, quantitative research on particular patterns and determinants of language acquisition has only recently started to fully acknowledge the fact that refugees exhibit vulnerabilities distinctive from other migrants (Kosyakova et al. 2021; Kristen and Sueuring 2021; van Tubergen 2010; Morrice et al. 2019; Bernhard and Bernhard 2022; Pozzo 2022; Pozzo and Nerghes 2020). Consequently this study also assesses whether the

association between enrolment timing and language course outcomes varies depending on the legal migration category considered. Results indicate that whereas labour migrants exhibit a positive association between enrolment postponement and course outcomes, the reverse holds for family and humanitarian migrants. This finding seems to align with previously published statements suggesting that refugees are more dependent upon early formal language programmes providing a systematic, conscious and focused learning trajectory (Schuller 2011; Kosyakova et al. 2021). In the absence of early course participation, refugees and family migrants might rely exclusively on language acquisition in informal settings, which has been found to stall at the level where proficiency is functional due to the so-called plateaus of language skills (Hoehne and Michalowski 2016). In addition, research in social psychology (Cohen and Garcia 2008; Walton and Cohen 2011) and theories of reactive ethnicity (Portes et al. 2005; Portes and Zhou 1993; Çelik 2015) indicate that early interventions signalling social belonging prevent negative vicious circles of poor learning outcomes and reactive identification in terms of ethnic origin.

We conclude that our finding that enrolment postponement associates negatively with language course outcomes amongst humanitarian and family migrants contributes new insights regarding vulnerabilities in language acquisition in the case of late course enrolment. Hence, we contribute to research on path-dependent vulnerabilities across the life courses of refugees connected to initial lengthy asylum procedures and legal residence status (Kosyakova and Brenzel 2020; Jutvik and Robinson 2020; Hvidtfeldt et al. 2018; Hainmueller et al. 2016). Although the identification of causal effects remains a challenge, available literature at least suggests that language proficiency indeed entails a wide range of advantages for immigrants’ integration in terms of economic participation (Dustmann and Fabbri 2003; Neureiter 2019; Kanas et al. 2012; Beiser and Hou 2000; Hsieh 2021; Chiswick et al. 2020; Lang 2022; Wong 2023; Miyar-Busto et al. 2020), but also social integration (Martinovic et al. 2008; Hoehne and Michalowski 2016; Morrice et al. 2019; Özmete et al. 2021), and mental health (Montemitro et al. 2021; Beiser and Hou 2001).

Finally, we present two avenues for future research. First, this study's main findings should be interpreted as facilitating future research adopting research designs more strongly tailored to the identification of causal effects. We indicate that migration background, demographic composition, human capital, and labour market attachment are important confounders which should be taken into account, but also that a persistent negative association between language course timing and outcomes amongst humanitarian and family migrants warrants attention in studies using research designs that control for (un)observed heterogeneity. Despite the possibility of exogenous variation regarding language course timing in Flanders (e.g. seasonal waiting times, or regional variation in administrative waiting times), due to the co-existence of many patterns of (self-)selection (see section "(Self-)selection") we were unable to identify such exogenous delays. Second, despite the fact that the general scarcity of empirical research on the timing of language programmes and similarities between language programmes across countries (OECD 2021) implies that our findings are relevant for international scholars and policy-makers, future research assessing language course timing and outcomes in other countries or regions should be encouraged.

Data availability

The datasets generated during and/or analysed during the current study are not publicly available since the author(s) do not have the legal right to publicly share this information which was used under licence for the current study. Example code excerpts are available from the corresponding author on reasonable request.

Received: 13 February 2023; Accepted: 12 April 2024;

Published online: 07 May 2024

Notes

- 1 A guest worker is defined as a person with temporary permission to work in another country in the Oxford dictionary, and in the Western European context the term mostly refers to post-WWII temporary workers that were recruited from Southern Europe and later non-European countries (e.g. Morocco, Turkey, Algeria) to re-build the economy, more specifically to work in coal mines and heavy industries.
- 2 Regularisation refers to the exceptional practice in which undocumented migrants ineligible for other procedures are granted a permanent residence permit, either due to medical or humanitarian reasons.
- 3 However due to a 2023 policy-change the integration programme is no longer free of charge for some migrant groups (cf. OECD 2023)
- 4 This phenomenon has also been referred to using "fossilisation" (e.g., Hoehne and Michalowski 2016), yet this article adopts the term language learning plateau, which is being increasingly used in the literature.

References

- Adamuti-Trache M (2013) Language acquisition among adult immigrants in Canada: the effect of premigration language capital. *Adult Educ Q* 63:103–126
- Alba R, Logan J, Lutz A et al. (2002) Only English by the third generation? Loss and preservation of the mother tongue among the grandchildren of contemporary immigrants. *Demography* 39:467–484
- Auer D (2018) Language roulette: the effect of random placement on refugees' labour market integration. *J Ethn Migr Stud* 44:341–362
- Beenstock M (1996) The acquisition of language skills by immigrants: the case of Hebrew in Israel. *Int Migr* 34:3–30
- Beiser M, Hou F (2000) Gender differences in language acquisition and employment consequences among Southeast Asian refugees in Canada. *Can Public Policy / Anal de. Politiques* 26:311–330
- Beiser M, Hou F (2001) Language acquisition, unemployment and depressive disorder among Southeast Asian refugees: a 10-year study. *Soc Sci Med* 53:1321–1334
- Bernhard S, Bernhard S (2022) Gender differences in second language proficiency: evidence from recent humanitarian migrants in Germany. *J Refugee Stud* 35:282–309

- Biegel N, Wood J, Neels K (2021) Migrant-native differentials in the uptake of (in) formal childcare in Belgium: The role of mothers' employment opportunities and care availability. *J Fam Res* 33:467–508
- Blommaert J (2011) The long language-ideological debate in Belgium. *J Multicult Discourses* 6(3):241–256
- Carliner G (2000) The language ability of U.S. immigrants: assimilation and cohort effects. *Int Migr Rev* 34:158–182
- Çelik Ç (2015) Having a German passport will not make me German': reactive ethnicity and oppositional identity among disadvantaged male Turkish second-generation youth in Germany. *Ethn Racial Stud* 38:1646–1662
- Charette M, Meng R (1994) Explaining language proficiency: objective versus self-assessed measures of literacy. *Econ Lett* 44:313–321
- Chiswick BR, Miller PW (1995) The endogeneity between language and earnings: international analyses. *J Labor Econ* 13:246–288
- Chiswick BR, Miller PW (2001) A model of destination-language acquisition: application to male immigrants in Canada. *Demography* 38:391–409
- Chiswick BR, Miller PW (2002) immigrant earnings: language skills, linguistic concentrations and the business cycle. *J Popul Econ* 15:31–57
- Chiswick BR, Rebhun U, Beider N (2020) Language acquisition, employment status, and the earnings of Jewish and non-Jewish immigrants in Israel. *Int Migr* 58:205–232
- Cohen GL, Garcia J (2008) Identity, belonging, and achievement: a model, interventions, implications. *Curr Directions Psychol Sci* 17:365–369
- Devlieger M, Lambrechts D, Steverlyncx C et al. (2014) Verbindingen: eindrapport werkjaar 2012–2013 van de inhoudelijke inspectie inburgering ["Connections: 2012_2013 Report of Inspection Integration"]. Vlaamse Overheid - Onderwijsinspectie - Inburgering [Flemish Government - Educational Inspection - Integration], Brussels
- Dustmann C, Fabbri F (2003) Language proficiency and labour market performance of immigrants in the UK. *Econ J* 113:695–717
- Dustmann C, Soest AV (2001) Language fluency and earnings: estimation with misclassified language indicators. *Rev Econ Stat* 83:663–674
- Elder GH, Johnson MK, Crosnoe R (2003) The emergence and development of life course theory. in: Mortimer JT and Shanahan MJ (eds) *Handbook of the life course*. Boston, MA: Springer US, 3–19
- Espenshade TJ, Fu H (1997) An analysis of English-language proficiency among U.S. immigrants. *Am Sociol Rev*. 62:288–305
- Esser H (2006) *Migration, sprache und integration: die sozialen bedingungen und folgen des spracherwerbs von migranten*. ["Migration, language and integration: the social conditions and consequences of migrants' language acquisition."] Berlin: Arbeitsstelle Interkulturelle Konflikte und gesellschaftliche Integration (AKI) ["Office for Intercultural Conflicts and Social Integration"] 119p
- FitzGerald DS, Arar R (2018) The sociology of refugee migration. *Annu Rev Sociol* 44:387–406
- Gonzalez A (2000) The acquisition and labor market value of four English skills: new evidence from NALS. *Contemp Econ Policy* 18:259–269
- Hainmueller J, Hangartner D, Lawrence D (2016) When lives are put on hold: lengthy asylum processes decrease employment among refugees. *Sci Adv* 2:e1600432
- Hayfron JE (2001) Language training, language proficiency and earnings of immigrants in Norway. *Appl Econ* 33:1971–1979
- Hoehne J, Michalowski I (2016) Long-term effects of language course timing on language acquisition and social contacts: Turkish and Moroccan immigrants in Western Europe. *Int Migr Rev* 50:133–162
- Hsieh YJT (2021) Learning language and gaining employment: problems for refugee migrants in Australia. *Equality Diversity Incl*. 40:1013–1031
- Hvidtfeldt C, Schultz-Nielsen ML, Tekin E et al. (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:e0206737
- Hwang S-S, Xi J (2008) Structural and individual covariates of English language proficiency. *Soc Forces* 86:1079–1104
- Jutvik K, Robinson D (2020) Permanent or temporary settlement? A study on the short-term effects of residence status on refugees' labour market participation. *Comp Migr Stud* 8:44
- Kanas A, Chiswick BR, van der Lippe T et al. (2012) Social contacts and the economic performance of immigrants: a panel study of immigrants in Germany. *Int. Migr. Rev.* 46:680–709
- Kasztan Flechner T, Neels K, Wood J, Biegel N (2022) Exploring Women's Uptake of Active Labour Market Programmes: The Role of Household Composition Across Migrant Origin Groups [active labour market policies; Belgium; household composition; migrant origin; mothers; training. *Soc Inclusion* 10(2):15
- Kogan I, Kalter F (2020) An empirical-analytical approach to the study of recent refugee migrants in Germany. *SozW Soz Welt* 71:3–23
- Kosyakova Y, Brenzel H (2020) The role of length of asylum procedure and legal status in the labour market integration of refugees in Germany. *SozW Soz Welt* 71:123–159

- Kosyakova Y, Kristen C, Spörlein C (2021) The dynamics of recent refugees' language acquisition: how do their pathways compare to those of other new immigrants? *J Ethn Migr Stud* 48:1–24
- Kristen C, Mühlau P, Schacht D (2016) Language acquisition of recently arrived immigrants in England, Germany, Ireland, and the Netherlands. *Ethnicities* 16:180–212
- Kristen C, Sueuring J (2021) Destination-language acquisition of recently arrived immigrants: do refugees differ from other immigrants? *J Educ Res Online* 13:128–156
- Lang J (2022) Employment effects of language training for unemployed immigrants. *J Popul Econ* 35:719–754
- Maes J, Neels K, Biegel N, Wood J (2023) Uptake of formal childcare among second generation and native mothers in Belgium: can increasing local childcare availability narrow migrant-native gaps? *Genus* 79(1):7
- Martinovic B, van Tubergen F, Maas I (2008) Dynamics of interethnic contact: a panel study of immigrants in the Netherlands. *Eur Sociol Rev* 25:303–318
- Mehlem U (2003) *Qualitätsanforderungen für die Sprachförderung im Rahmen der Integration von Zuwanderern* ["Quality requirements for language support in the context of the integration of immigrants"]: IMIS, Institut für Migrationsforschung und Interkulturelle Studien ["Institute for Migration Research and Intercultural Studies"]
- Mesch GS (2003) Language proficiency among new immigrants: the role of human capital and societal conditions: the case of immigrants from the FSU in Israel. *Sociol Perspect*. 46:41–58
- Miyar-Busto M, Mato Diaz FJ, Gutiérrez R (2020) Immigrants' educational credentials leading to employment outcomes: the role played by language skills. *Revista Internacional de Organizaciones*: 167–191
- Montemitto C, D'Andrea G, Cesa F et al. (2021) Language proficiency and mental disorders among migrants: a systematic review. *Eur. Psychiatry* 64:18
- Mood C (2009) Logistic regression: why we cannot do what we think we can do, and what we can do about it. *Eur Sociol Rev* 26:67–82
- Morrice L, Tip LK, Collyer M et al. (2019) You can't have a good integration when you don't have a good communication: English-language learning among resettled refugees in England. *J Refugee Stud* 34:681–699
- Neureiter M (2019) Evaluating the effects of immigrant integration policies in Western Europe using a difference-in-differences approach. *J Ethn Migr Stud* 45:2779–2800
- Norton B (1997) Language, identity, and the ownership of English. *TESOL Q* 31:409–429
- OECD (2015) *International migration outlook 2015*. OECD Publishing, Paris
- OECD (2021) *Language training for adult migrants, making integration work*. OECD Publishing, Paris
- OECD (2023) *Skills and labour market integration of immigrants and their children in Flanders*. OECD Publishing, Paris
- Özmete E, Pak M, Duru S (2021) Problems and issues concerning social integration of elderly refugees in Turkey. *J Refugee Stud* 35:93–121
- Portes A, Fernández-Kelly P, Haller W (2005) Segmented assimilation on the ground: the new second generation in early adulthood. *Ethn Racial Stud* 28:1000–1040
- Portes A, Zhou M (1993) The new second generation: segmented assimilation and its variants. *ANNALS Am Acad Polit Soc Sci* 530:74–96
- Pozzo M (2022) Moving on from Dutch to English: young refugees feeling betrayed by the Dutch language integration policy and seeking for more inclusive environments. *J Refugee Stud* 35:348–367
- Pozzo M, Nerghes A (2020) Dutch without the Dutch: discourse, policy, and program impacts on the social integration and language acquisition of young refugees (ages 12–23). *Soc Identities* 26:842–860
- Schuller K (2011) *Der Einfluss des Integrationskurses auf die Integration russisch- und türkischstämmiger Integrationskursteilnehmerinnen*. Qualitative ergänzungsstudie zum Integrationspanel. ["The influence of the integration course on the integration of integration course participants of Russian and Turkish origin. Qualitative supplementary study to the integration panel."] Nürnberg: Bundesamt für Migration und Flüchtlinge ["Federal Office for Migration and Refugees"]: Working Paper number 37
- Statistics Belgium (2023). *Migrations: Total international migration*. Brussels: Statbel: dataset available on <https://statbel.fgov.be/nl/themas/bevolking/loop-van-de-bevolking/migraties#figures>
- Statistics Flanders (2023). *International migration: Statistics Flanders: dataset available on* <https://www.vlaanderen.be/statistiek-vlaanderen/bevolking/internationale-migratie#internationale-immigratie-op-europes-gemiddelde>
- Stevens G (1999) Age at immigration and second language proficiency among foreign-born adults. *Lang Soc* 28:555–578
- Temple B (2010) Feeling special: language in the lives of Polish people. *Sociol Rev* 58:286–304
- Timmerman C (2017) Social sciences and Moroccan migration in Belgium. In: Timmerman C, Fadil N, Goddeeris I, et al. (eds) *Moroccan migration in Belgium: More than 50 years of settlement*. Leuven: Leuven University Press: 9–20
- UNIA, & FOD WASO (2017). *Socio-economische monitoring - arbeidsmarkt en origine* [socio-economic monitoring – labour market and migration background]. Retrieved from Brussel: <https://www.unia.be/nl/publicaties-statistieken/publicaties/socio-economische-monitoring-2017-arbeidsmarkt-en-origine>
- van Tubergen F (2010) Determinants of second language proficiency among refugees in the Netherlands. *Soc Forces* 89:515–534
- van Tubergen F, Kalmijn M (2005) Destination-language proficiency in cross-national perspective: a study of immigrant groups in nine western countries. *Am J Soc* 110:1412–1457
- van Tubergen F, Kalmijn M (2008) Language proficiency and usage among immigrants in the Netherlands: incentives or opportunities? *Eur Sociol Rev* 25:169–182
- van Tubergen F, Wierenga M (2011) The language acquisition of male immigrants in a multilingual destination: Turks and Moroccans in Belgium. *J Ethn Migr Stud* 37:1039–1057
- Vlaamse Regering. (2019) *Beleidsnota 2019–2024 gelijke kansen, integratie en inburgering* ["Policy Note 2019–2024: equal opportunities & integration"]. Brussels: Flemish Government, 141
- Walton GM, Cohen GL (2011) A brief social-belonging intervention improves academic and health outcomes of minority students. *Science* 331:1447–1451
- Wei X (2008) Implication of IL fossilization in second language acquisition. *Engl Lang Teach* 1:127–131
- Wong L (2023) The effect of linguistic proximity on the labour market outcomes of the asylum population. *J Popul Econ* 36:609–652
- Wood J, Neels K (2020). *Wegwijs naar werk. Longitudinale analyse en evaluatie van inburgerings- en activeringstrajecten in Vlaanderen* ["A route to employment. Longitudinal analysis of integration and activation programmes in Flanders"], 2005–2016 (Report VIONA Chair Labour Market). Antwerp: University of Antwerp Press

Author contributions

DL was responsible for data acquisition and cleaning. JW and DL both contributed to the literature review. JW was responsible for the analyses and first version of the article. JW and DL both revised the manuscript during the review procedure.

Competing interests

The authors declare no competing interests.

Ethical approval

Ethical approval was granted by the Ethics Committee for the Social Sciences and Humanities of the University of Antwerp, file number SHW_20_102.

Informed consent

This article does not contain any studies with human participants performed by any of the authors. As a result, no informed consent was required.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1057/s41599-024-03034-0>.

Correspondence and requests for materials should be addressed to Jonas Wood.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024