

# The role of morphological complexity in predicting the learnability of an additional language: The case of La (additional language) Dutch

Second Language Research

2019, Vol. 35(1) 47–70

© The Author(s) 2017



Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/0267658317691322

[journals.sagepub.com/home/slr](https://journals.sagepub.com/home/slr)

**Frans van der Slik**  
**Roeland van Hout**

Radboud University, Netherlands

**Job Schepens**

Freie Universität Berlin, Germany

## Abstract

Applied linguistics may benefit from a morphological complexity measure to get a better grip on language learning problems and to better understand what kind of typological differences between languages are more important than others in facilitating or impeding adult learning of an additional language. Using speaking proficiency scores of 9,000 adult learners of Dutch as an additional language, we reproduced the findings of the Schepens et al. (2013a) study, using a reduced morphological complexity measure. We wanted to define a reduced measure to reveal which morphological features constitute the really important learning problems. Adult language learners whose first language (L1) has a less complex morphological feature configuration than Dutch turned out to have more learning difficulties in acquiring Dutch the less complex their L1 is in relation to Dutch. The reduced measure contains eight features only. In addition, we found cognitive aging effects that corroborate the construct validity of the morphological measure we used. Generally, adult language learners' speaking skills in Dutch improve when residing longer in the host country. However, this conclusion is only warranted when their L1 morphological complexity is at least comparable to Dutch morphological complexity. If the morphological complexity of their L1 is lower as compared to Dutch, the effect of length of residence may even reverse and have a negative impact on speaking skills in Dutch. It was observed that the negative effect of age of arrival is mitigated when adult language learners have a command of a second language (L2) with higher morphological complexity. We give morphological information for five additional target languages: Afrikaans, Chinese, English, German, and Spanish.

## Corresponding author:

Frans van der Slik, Centre for Language Studies, Department of Linguistics, Radboud University, Erasmusplein 1, 6525 HT Nijmegen, Netherlands.

Email: [f.v.d.slik@let.ru.nl](mailto:f.v.d.slik@let.ru.nl)

## Keywords

age of arrival, cognitive aging, cross-classified multilevel models, length of residence, morphological complexity, second language acquisition

## I Introduction

In acquiring an additional language (La), adult second language learners experience a variety of problems among which acquiring the vocabulary is the most obvious one. Learning the morphological rules of an La is another well-known problem (Housen, De Clercq, Kuiken and Vedder, 2019). It is often a painstaking process marked by transfer issues, i.e. the influences of the first language (L1) on learning an La, that can either be beneficial or hindering, depending on the presence or absence of L1 and La morphological features (De Clercq and Housen, 2019; Ehret and Szmrecsanyi, 2019; Kellerman and Sharwood Smith, 1986; Lado, 1957; O'Grady, 2006; Odlin, 1989; Paquot, 2019; Weinreich, 1963). Schepens et al. (2013b) found that the lower the morphological complexity of the L1 is, in comparison to Dutch, the lower the speaking proficiency scores of adult learners of La Dutch are. This is nicely illustrated by the outcomes of the present study as well. It will turn out that adult L1 German learners of La Dutch had better results than adult L1 Afrikaans learners of La Dutch on an official exam of Dutch as a second language. While the three languages are West Germanic Languages, Afrikaans is a daughter language historically closer to Dutch. Its morphology and grammar became more analytic than Dutch in the course of time (van Rensburg, 1983; van Sluijs, 2013), but their lexicons are still quite similar. German with its extended case system is morphologically more complex than Dutch. L1 German learners have to cope with less morphological complexity; L1 Afrikaans learners have to learn new, additional morphology, like the use of agreement between subject and verb. The observed differences in learning success between the two language groups, both groups performing highly, but the L1 Afrikaans learners performing worse than the L1 German learners, clearly demonstrate the conceivable role morphological complexity may have in adult second language acquisition processes. Linguistic complexity can be defined in many ways (Ehret and Szmrecsanyi, 2019; Szmrecsanyi and Kortmann, 2012). Schepens et al. (2013a) made use of a relative measure of morphological complexity (see Arends, 2001; De Clercq and Housen, 2019; Kusters, 2008; Miestamo, 2008) relating the properties of  $L_x$  to the properties of  $L_y$ . The present study will also make use of such a relative morphological complexity measure.

How can we measure the morphological complexity of languages? Lupyan and Dale (2010) selected 28 morphological features from the WALS (world atlas of language structures) database (Dryer and Haspelmath, 2011). One of these features, for instance, is the encoding of past tense through verb inflection versus its expression in an analytic way by explicit lexical means on the sentence level. Other features relate to the complexity of conjugations and case systems and to the use of lexical means to encode, for instance, evidentiality, negation, aspect, and possession versus synthetic, morphological inflections. Lupyan and Dale (2010) applied these 28 features to more than 2,000 languages. Their Linguistic Niche Hypothesis makes a distinction between languages spoken in esoteric

niches (smaller populations, smaller areas, fewer linguistic neighbours), and exoteric niches with large numbers of speakers. Population size came out as a proper predictor for complexity-related patterns of 23 out of these 28 features, which is in accordance with the Linguistic Niche Hypothesis. Feature values that express lexical coding were found more predominantly among languages spoken in exoteric niches, while values expressing morphological coding were found more often among languages spoken in esoteric niches. By adding up the number of features for which each language relies on either lexical or morphological coding, Lupyan and Dale (2010) constructed a more general morphological complexity measure. It turned out that languages spoken by large groups are more likely to have lexical strategies instead of morphological procedures. Lupyan and Dale (2010) explain the relationship between morphological complexity and population size by the assumption that larger languages are often acquired by adult second language learners who fail to learn the complex features of their additional language, and, therefore, will not pass them on to subsequent learners. L1 speakers probably take their share as well, as suggested by Wray and Grace (2007), by expressing themselves in more analytical ways (see the concept of foreigner talk), to facilitate communication with second language (L2) speakers. As a result, the languages involved will ultimately become more analytical, solving communicative demands of their users by providing explicit lexical means instead of morphological tools.

The link between the impact of adult language learning and decreasing complexity has been explicated in many other studies. McWhorter (2002), for example, explains the simplification of English to massive second language acquisition by Scandinavians from the eighth century onwards. Other examples are Creole languages. McWhorter (2001: 163) gives the following list of grammatical features typically levelled by creolization in the 19 Creoles he investigated ‘[...] ergativity, grammaticalized evidential marking, inalienably marking, switch-reference marking, inverse marking, obviative marking, “dummy” verbs, syntactic asymmetries between matrix and subordinative clauses, grammaticalized subjunctive marking, verb-second, clitic movement, any pragmatic neutral word order but SVO, noun class or grammatical gender marking (analytical or affixial), or lexically contrastive or morphosyntactic tone [...]’. Although McWhorter refers to syntactical and morphological features as communicatively redundant baggage coming from the past, his list is also bound to answer the question why morphosyntactic simplification occurs in the first place. According to Trudgill (2001: 372) ‘[linguistic] complexity increases through time, and survives as the result of the amazing language-learning abilities of the human child, so complexity disappears as a result of the lousy language-learning abilities of the human adult.’ This conclusion may seem rather strong, given the abundant evidence that adult language learners can reach near-native levels of target language proficiency (e.g. Birdsong, 2005; Bongaerts et al., 1995, 2000), and that adult learners show steeper learning curves than children, at least in the initial stages (Krashen et al., 1982; Marinova-Todd et al., 2000; Saville-Troike, 2012; Snow and Hoefnagel-Höhle, 1977). However, the general trend to be observed is that adult language acquisition is marked in the end by lower overall levels of proficiency in the target language.

How can we investigate the role of morphological complexity in adult language learning in more detail? In a previous study (Schepens et al., 2013a), we used the morphological features selected by Lupyan and Dale (2010) to predict speaking proficiency scores

of adult learners of Dutch (for a description of these features, see Lupyan and Dale, 2010: Text S6; Schepens, 2015: 81–82; see also <http://wals.info/feature>). We defined complexity of the L1s of the learners relative to Dutch by comparing their L1s and Dutch complexity values of the features selected by Lupyan and Dale (2010). The correlations we found between feature-specific morphological complexity differences between Dutch and the 73 L1s of the adult learners confirmed that morphological complexity poses serious learning problems and that we could compute an overall morphological complexity measure for differences between the L1s involved and Dutch. To obtain an overall measure of complexity between the L1s and Dutch, we computed the overall correlation between learnability of Dutch (based on the scores on the speaking proficiency test) and the overall set of differences in morphological complexity using weighted correlation coefficients of all features. We first address the preliminary question whether our earlier findings for La Dutch were robust enough to emerge again in analysing a more recently collected set of data (2010–14 instead of 1995–2010):

1. Does our new data set reveal the same correlation pattern between an overall L1 morphological complexity measure and proficiency in Dutch as found in Schepens et al. (2013a)?

However, using weighted correlation coefficients to compute overall morphological complexity is a rather opaque procedure because of intercorrelations between the features involved (the features may for instance represent comparable morphological procedures) and because weights have the disadvantage of being sample specific (they depend for instance on the distribution of the morphological features present in the L1s of the sample of learners). It is our aim in this contribution to define morphological complexity in a more transparent way by focusing on a limited set of features, in relation to La Dutch. A reduced measure may enable us to gain deeper insight in which morphological features facilitate or impede La acquisition.

Specifically, we found in Schepens et al. (2013a) that out of eight features for which there were significant effects of morphological differences across the L1s of the learners, seven were due to an increase in morphological complexity from a L1 perspective and only one was due to a decrease. When La learners have less complex values on these seven features they need to learn new morphological procedures (e.g. with respect to verbal person marking, past tense, coding of negations, or possessives). So, our next step is to dissect the morphological complexity measure in detail by focusing on individual morphological features that really seem to matter by being directly involved in shaping overall morphological thresholds adult Dutch La learners have to cope with. Our second research question therefore is:

2. Can we reduce the overall morphological complexity measure in an insightful way without losing power in predicting Dutch La learnability?

Next, Schepens et al. (2016) found that not only the L1s, but also the L2s morphological complexity had an impact on learning La Dutch, although considerably less than the L1 effect. We want to investigate the role of the L2 in the new data as well, because, again,

a recurrent pattern would substantiate the relevance of the morphological complexity effect and underline the essential role of morphology in adult language acquisition. Moreover, taking this effect into account reflects the multilingual situation of the respondents who in the majority of the cases learned Dutch as a third or later language. This effect is comprised in our third research question:

3. Does L2 morphological complexity have an influence on acquiring La Dutch?

Our final research question pertains to the influence of age-bound factors. We want to provide supportive evidence in favour of our morphological complexity measure by investigating the effects of age-bound factors. Research on language acquisition consistently points out that child learners are dramatically better equipped to master morphology than adult learners, although this distinction seems to evolve gradually over age (see Bialystok and Hakuta, 1994; Birdsong, 2005; Chiswick and Miller, 2008; van der Slik, 2010), and less sharply or discontinuously than suggested in the past (Johnson and Newport, 1989). Language learning capacities seem to decline gradually over a longer period of time, putting forward the question whether cognitive aging directly relates to life-long gradual processes of decay leading to a steady, slow, but persistent decline in language learning capacities (Hakuta et al., 2003), particularly in the domain of morphology (see Birdsong and Flege, 2001). Schepens et al. (2016) found age-bound effects in adult language learners. Learning success appeared to gradually decrease with increasing age of arrival (AoA). The other age-bound variable was length of residence (LoR) in the Netherlands. The longer the LoR was the better the proficiency results were in the exams, implying that experience and practice may in fact compensate for cognitive aging. Both effects seem plausible from the viewpoint of cognitive aging processes and time spent to learn. Consequently, we expect these age-bound effects to recur in the analysis of our new data set. However, age-bound effects may perhaps deliver additional, even crucial evidence that supports the validity of the morphological complexity measure we use. When acquiring morphology is an age-bound process, morphology is not only hard for adults in general, becoming harder for them growing older, but its effect may be moderated by the degree of complexity of the La, in comparison to the complexity of their L1. Salthouse (2004, 2009, 2010a, 2010b) presents ample evidence that age-related cognitive decline starts already around the age of 20, but also that cognitive aging does not occur in all domains equally strong and with the same speed. The capacity of learning new vocabulary, for example, remains remarkably intact. Various studies (see Salthouse, 2004) have shown that the number of words people are familiar with keep increasing until people are in their sixties. On the other hand, a measure of paired associates memory performance, which aims at linking a target word to the best associate only recently acquired, shows a monotonic decrease with age (Salthouse, 2010a). In addition, measures of speed, reasoning and memory have fairly impressive linear declines from the age of 20 onwards (e.g. Harada et al., 2013). The actual effects of cognitive decline appear to be hardly noticeable in everyday life: ‘the greater experience and knowledge associated with increased age probably reduces the need for the type of novel problem solving that declines with age’ (Salthouse, 2004: 142).

Acquiring an additional language however requires novel types of problem solving. Acquiring an additional language that is morphologically more complex than one's first language seems to present an extra challenge that is more susceptible to cognitive aging than acquiring an additional language that is less complex. The harder the learning task, i.e. the more complex the La is in comparison to the L1, the more negative the outcome of cognitive aging is expected to be. This expectation is also sparked by the outcomes of the Birdsong and Flege (2001) study, which reported an interaction between age of arrival and morphological complexity. Keeping length of residence constant, they found a stronger negative effect of age of arrival on grammaticality judgments in morphologically more complex L2 tasks than in morphologically less complex tasks (see also Birdsong, 2005). We therefore formulated the following research question including AoA as the predictor and morphological complexity as the moderator variable:

4. Does morphological complexity help to explain differences in age of arrival (AoA) effects in acquiring La Dutch?

In investigating AoA effects we need to take LoR into account, as done by Birdsong and Flege (2001), because LoR is another and perhaps competing age-bound factor that affects La acquisition.

## II The present study

We made use of the most recent subset of a large-scale database, gathered in the period 2010–14, containing information on 22,638 immigrants to re-examine the impact of the morphological complexity of their L1 and L2 on their speaking proficiency scores on Dutch as an additional language. This is a partial replication of a study that used a less recent subset of the same database, containing information on around 50,000 immigrants (Schepens et al., 2013a, 2016), collected in the period 1994–2010. The large-scale database that we used provides a particular strong testing ground for our research questions given the enormous amount of L1s and countries of origin and given the social-demographic and contextual characteristics we have included in the statistical analyses.

The impact of morphological complexity on the acquisition of Dutch as an additional language was examined in immigrants from 33 mother tongue backgrounds, spoken in 67 countries, 20 of which were Indo-European (IE) and 13 non-Indo-European (non-IE). In the latter group, there were three Afro-Asiatic (Arabian, Berber, Amharic), two Austronesian (Indonesian, Tagalog), and two Uralic languages (Finnish, Hungarian). There were two Altaic (Mongolic, Turkish), one Kartvelian (Georgian), one Japanese, one Sino-Tibetan (Chinese), and one Tai-Kadai (Thai) language. Examinees originated from 29 Western countries (including Canada, the USA, and former East European countries), eight African countries, 11 countries from South and Central America, five from East Asia, seven from the Middle East, and seven South Asian countries. Stable estimates of country and language level effects require a sufficient number of observations in the country-level combinations. The minimum amount of observations is open to discussion, however (see Bell et al., 2010). We opted for the requirement that countries of origin and mother tongues had to contain a minimum of 15 examinees in order to be included in this study, like we did in previous studies.

### III Method

#### I Sample

Currently, the State Examination of Dutch as a Second Language (STEX) is administered almost every week of the year. The examination consists of two separate exams. Program II (STEX II) is meant for immigrants who intend to enrol in higher-level education in the Netherlands, or who have a higher-level occupation. Program I (STEX I) aims at immigrants who intend to follow a lower level of (vocational) education, or who have a lower or middle-level occupation. The requirements for Dutch language proficiency are the same for both levels, but the abstraction (academic) level of Program II is higher (for detailed information, see College voor Toetsen en Examens, 2016; Schepens et al., 2013a).

The Program II exam results were available for the years 2010 up to 2014. The examination covers four language skills: speaking, writing, reading, and listening, which are tested separately. An examinee passes the entire exam when she or he has acquired 500 points or more on each of the four sub-exams. The current study focuses on the productive skill of Dutch speaking proficiency. Writing skills in La Dutch could also have been examined, but we decided to keep our focus on speaking. Writing skills not only develop at a later age, including them would also require additional knowledge on the examinees' command of the Roman script and on the different literacy methods used worldwide. The results of a previous study also indicated that the outcomes for La Dutch writing proficiency are largely comparable to those of La Dutch speaking proficiency (van der Slik et al., 2015).

In total, 22,638 examinees took the speaking sub-exam in the period 2010–14. In case of re-exams, we only used the first available test score. Data on test scores, gender and age were available for all examinees. Only adult second language learners who arrived in the Netherlands between 18 and 65 years of age were included in the study. This reduced the number of examinees by only 3% to 21,700.

As part of the electronic registering procedure, examinees are invited to fill in a brief questionnaire about various background characteristics, such as date of arrival in the Netherlands, country of birth, mother tongue, education, etc. This reduced the number of examinees with valid scores substantially to 12,540. The morphological complexity scores of examinees' mother tongues were not available for all L1s. This reduced the number of valid scores on all the variables of interest to 11,814. Finally, restricting the data to L1s and countries of birth containing at least 15 learners reduced the number of valid scores to 8,754.

#### 2 Variables

All variables involved are summarized in Table 1. These variables are discussed below as criterion variable (speaking proficiency), and learner and context characteristics. Table 1 gives information on the means scores (plus standard deviations) or proportions of these variables.

The Dutch speaking proficiency test was constructed by the Centraal Instituut Toetsontwikkeling (Central Institute for Test Development) and the Bureau Interculturele Evaluatie (Bureau for Intercultural Evaluation), two large test battery constructors in the

**Table 1.** Description of the sample for the dependent variable the Dutch Speaking Proficiency Test (criterion), and the explanatory variables, split out between learner and context characteristics (33 L1s, 35 L2s and 67 countries).

	Mean	SD	<i>n</i>
<i>Criterion:</i>			
Dutch Speaking Proficiency Test	519	23	8,754
<i>Learner characteristics:</i>			
Gender (female = 1)	.75		8,754
Age of arrival in the Netherlands	27.48	6.56	8,754
Length of residence (at first exam date)	4.83	4.88	8,754
Years of education	3.31	.85	8,754
<i>Context characteristics:</i>			
Gross enrolment rate in secondary schooling in 2005 (% of population of secondary education age)	85.76	26.01	67
Western countries (reference category in the analysis)	.57		5,060
(proportion of learners in the sample)			
Middle East	.14		1,188
South America and Central America	.08		705
Africa	.08		629
South Asia	.04		352
East Asia	.09		820
Morphological complexity L1	.77	.25	33
Morphological complexity L2	.74	.28	35

Notes. L1 = first language; L2 = second language.

Netherlands. The typical speaking test consists of 14 assignments. Candidates are urged to respond orally to prompts like: 'In Dutch television a lot of ads are made for all kinds of products, even in the middle of a program. What is your opinion about ads on TV?' These spoken elicitations are recorded on tape. The examination takes 30 minutes. Two independent expert raters evaluate the oral production on content and correctness criteria. Primary content criteria are the appropriateness of the content related to the task (about 30%) and vocabulary size (around 18%). The most important linguistic criteria are word and sentence formation (about 28%), and pronunciation (about 12%). The remaining 12% refer to fluency, rate of speech, coherence, word choice and register (for detailed information, see Schepens et al., 2016).

The degree of difficulty of the examinations was held constant over time, by applying a specific Item Response Theory (IRT) model, namely the One-Parameter Logistic Model (Verhelst et al., 1995): an advanced type of Rasch model. A decisive advantage of IRT models as compared to models based on Classical Test Theory is that the test scores of candidates who took the exam on different occasions are allocated to the same ability distribution, implying that their test results can be analyzed together. In order to achieve this, parts of earlier exams were used in new exams (though the actual design was much



more complex). The scores on the exam were standardized. A mark of 500 or higher means that the candidate had passed the exam and indicates that the examinee has a proficiency at the B2 level (independent user, vantage level) as defined in the Common European Framework (Council of Europe, 2001), equivalent to IELTS 5.5 (International English Language Testing System; Kuijper et al., 2004).

### 3 *Learner characteristics*

*a Gender.* Females were coded as 1, males as 0.

*b Age of arrival in the Netherlands.* On the basis of information on ‘Year of birth’ and ‘Date of arrival’, age at the time examinees arrived in the Netherlands was calculated. We excluded examinees who were younger than 18 years at the time of their arrival and examinees who were older than 65 years of age when arriving.

*c Length of residence.* On the basis of information on ‘Date of the exam’ and ‘Date of arrival’, length of residence in the Netherlands was calculated.

*d Years of education.* Examinees were asked how many years of regular daily education they had received from their sixth year onwards. Possible answers were: (1) up to 5 years; (2) between 6 and 10 years; (3) between 11 and 15 years; (4) 16 years or more.

All variables were centred around their grand mean to reduce the risks of multicollinearity in interaction and higher order terms (Kreft and de Leeuw, 1998).

### 4 *Context characteristics*

*a Morphological complexity L1.* In the present study, we apply an alternative, reduced way of computing morphological complexity. Previous studies (Schepens, 2015: 83–86; Schepens et al., 2013a, 2016) in fact showed that that morphological complexity only had an impact when the source language (the L1) was less complex than Dutch. Dutch has a positive complexity score on 16 of the 28 features (see Appendix 1; a score higher than 1), 8 of which had positive correlations with the measure of increasing morphological complexity of above .30 for the L1s (see Table 2). We used only these 8 features to compute a reduced measure of increasing morphological complexity; for a description of these features, see Appendix 2.

*b Morphological complexity L2.* Examinees were not just asked what their first language was, but also if they had command of other languages. If the answer was affirmative, they were asked to indicate which second language they spoke best. Based on this information, the morphological complexity of their second language relative to Dutch was computed.

Not all immigrants included in this study indicated to have command of an L2. The L2 complexity score of all monolingual language learners of Dutch was set at zero, signalling for this group of learners the complexity score of a language that is minimally complex compared to Dutch (for a detailed discussion, see also Schepens et al., 2016). These monolingual learners still have the complexity score of their L1.

**Table 2.** The correlations of the selected eight morphological features (WALS number) and speaking proficiency in La Dutch (BLUPS); the number of languages is 49, 32, and 33, respectively.

WALS number		Schepens et al. (2013a)	This study (Afrikaans excluded)	This study (Afrikaans included)*
29	Syncretism in verbal person	.71	.59	.41
100	Alignment of verbal person marking	.68	.57	.37
102	Person marking on verbs	.68	.57	.37
66	Past tense	.43	.35	.36
92	Polar question coding	.34	.62	.56
112	Coding of negation	.33	.51	.39
22	Inflectional synthesis of the verb	.33	.10	.11
26	Prefixing vs. suffixing in inflectional morphology	.31	.31	.09

Notes. WALS = world atlas of language structures (Dryer and Haspelmath, 2011). \* We thank Bertus van Rooy (personal communication, 30 January 2016) and Robbert van Sluijs (personal communication, 1 February 2016) for providing the feature values for Afrikaans of these eight morphological features (see also Appendix 1).

*c Educational accessibility.* The World Bank (2015) reports on education data in a wide number of countries around the world on a regular basis. We took the gross enrolment rate in secondary schooling per country in 2005 as an indicator for a country's educational accessibility.

*d Geographic region.* Previous research has shown that L2 learners from different geographic regions may differ in their La proficiency (Carliner, 2000). The reasons for this variation are largely unknown, but we concluded to control for this confound by including geographic region as a proxy. To capture these potential test score differences we made a distinction between the following geographic regions: Western countries, South and Central America, Africa, the Middle East, South Asia and East Asia. Western countries were used as the reference category.

## 5 The reduced morphological complexity measure

The eight remaining WALS features and their numbers are given in Table 2; their actual complexity values are given in Appendix 1. Based on the coding patterns found by Lupyan and Dale (2010), each feature of a given language can be categorized as less complex (0) or equal or more complex (1) than Dutch. On feature 22 (Inflectional Synthesis of the verb), for example, Afrikaans, Dutch, English, and German have 2–3 categories per word, Spanish has 4–5 categories per word, while Chinese has 0–1 categories per word. Therefore, with respect to feature 22, Chinese is morphologically less, and Spanish more complex than Dutch. Less complex returns the value 0, equal or more complex returns the value 1. The complexity score was calculated by dividing the

**Table 3.** Cross-classified multilevel model parameter estimations for measures of the Dutch Speaking Proficiency Test (standard errors in parentheses) per mother tongue (L1), second language (L2), and country of birth.

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	515.36*** (2.29)	519.61*** (2.14)	518.24*** (1.96)	518.60*** (1.94)	518.71*** (1.93)
Female	3.99*** (.49)	3.95*** (.49)	3.97*** (.49)	3.97*** (.49)	3.97*** (.48)
Length of residence (LoR)	.63*** (.04)	.63*** (.04)	.63*** (.04)	.63*** (.04)	.58*** (.04)
Age of arrival (AoA)	-.49*** (.03)	-.49*** (.03)	-.49*** (.03)	-.49*** (.03)	-.52*** (.03)
Years of education	2.90*** (.24)	2.90*** (.24)	2.90*** (.24)	2.90*** (.24)	2.87*** (.24)
Female*education	.84 (.52)	.81 (.52)	.83 (.52)	.83 (.52)	.83 (.52)
Gross enrolment in 2005		.11** (.04)	.11** (.04)	.11** (.04)	.11** (.04)
Africa		-9.16* (3.52)	-8.97** (3.44)	-9.01** (3.45)	-8.88** (3.45)
Middle East		-7.34* (2.92)	-7.43* (2.84)	-7.43* (2.85)	-7.58** (2.85)
East Asia		-10.52** (3.76)	-1.86 (4.14)	-1.82 (4.14)	-1.12 (4.14)
South Asia		-6.66* (3.03)	-6.86* (2.93)	-6.97* (2.94)	-6.81* (2.94)
Central and South America		-3.09 (3.25)	-4.27 (3.15)	-4.40 (3.15)	-4.36 (3.15)
Morphological complexity L1			29.45*** (6.33)	29.50*** (6.31)	28.93*** (6.28)
Morphological complexity L2				3.65 (2.32)	3.66 (2.28)
Morphological complexity L1*LoR					1.45*** (.19)
Morphological complexity L2*AoA					.26*** (.08)
Residual	347.01*** (5.28)	347.52*** (5.29)	347.42*** (5.29)	347.48*** (5.29)	344.92*** (5.25)
Variance L1	108.34*** (33.65)	59.82** (20.03)	33.26** (11.60)	32.70** (11.43)	32.28** (11.21)
Variance L2	14.71* (6.09)	14.04* (5.84)	14.66* (6.16)	12.26* (5.74)	11.65* (5.46)
Variance country	45.35*** (11.04)	21.49*** (6.12)	20.68*** (5.73)	20.88*** (5.76)	20.94*** (5.75)
-2 log likelihood	76,311	76,253	76,229	76,223	76,161

Notes. \* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$ .

number of equal or more complex features by the total number of features available for that particular language (the eight features were not weighted according to their importance). For Afrikaans, only three out of eight features are evaluated as equal or more complex compared to Dutch, so the complexity score of Afrikaans relative to Dutch is  $3/8 = .375$ . German returns the maximum value of 1 (for more details, see Appendix 1). However, not all these eight morphological features are available for every language included in our study, so we decided to calculate a score only in case information on at least half of the features for a language was available.

It turned out that this reduced measure of complexity covered the original complexity measure almost entirely since the correlation on the L1 level was  $r(33) = .973$  ( $p < .0001$ ). Also the correlation for the L2s is extremely high,  $r(35) = .968$  ( $p < .0001$ ). Our first research question was whether the correlations found in the Schepens et al. (2013a) study would be replicated in the present study and these extremely high correlations automatically return a positive answer.

To investigate the specific role of the eight features in the acquisition of La Dutch speaking proficiency we computed correlations between these features and speaking proficiency. We repeated the analyses from Schepens et al. (2013a), now using the 2010–14 data base. We calculated Best Linear Unbiased Predictors (or BLUP's) for all L1s using a model with L1, L2 and country of birth as random factors, and controlling for the individual characteristics: gender, education, age of arrival, length of residence, and secondary school enrolment, using the *ranef* function in the statistical package *R*. Schepens et al. (2013a) did not include Afrikaans, as the data are not available in the WALS database.

We computed the values for the eight features in Table 2 and included the Afrikaans data in our analysis of the 2010–14 data. The changes in the first two columns do not seem substantial. The set of features remains intact regardless which data base is being used. The average correlation between feature values and BLUPS being .48 for the Schepens et al. (2013a) and .45, for this study (Afrikaans excluded). The correlations under the column with Afrikaans included clearly give lower outcomes (with an average correlation of .33) compared to the second column. It shows the special status of Afrikaans in our data analysis on morphological complexity, because its low morphological complexity is counteracted by the many lexical similarities between Dutch and Afrikaans. We return to this outcome in the discussion section.

An important issue is of course why these eight features are particularly relevant. Most of them relate to the verb, a domain where Dutch is marked by a series of inherent and derivational morphological properties (see Booij, 2002). The interpretation is complex, as the role of the particular configurations is dependent on the correlations between these properties as well. We come back to this issue in the discussion.

## 6 The statistical analysis

We used cross-classified multilevel models (we used SPSS 21, Heck et al., 2010), to take full account of the multilingual reality, i.e. migrants from different countries may have the same mother tongue, while migrants from the same country may speak different mother tongues, and may or may not have command of a second language. By applying

cross-classification, we were able to test for effects at the learner level, the mother tongue level, the second language level, and the country of origin level. As a result, we can investigate the impact of morphological complexity of both the L1 and L2, while simultaneously controlling for confounding effects at context and individual levels.

First, we defined a null model (Model 0) for the Dutch speaking proficiency test with no predictive factors added. Learners, countries, mother tongues, and second languages were included as random factors (random intercepts). Next, we added the variables at the learner level in Model 1. In Model 2, we added the country level predictors. And finally, in Models 3 and 4, the language level predictors (morphological complexity of the L1 and L2) were included.

The model improvement is tested by means of the Log Likelihood ratio which follows a chi-square distribution. A model is considered to have a better fit than a preceding, more parsimonious model if the difference in the Log Likelihood ratio ( $-2LL$ ) is at least 3.84 against one degree of freedom. When the improvement was significant, we checked the direction and size of the effect parameters.

The null model of the speaking proficiency scores contains four random variance components (so-called intra-unit correlations). They showed that 19.2% of the variation in speaking proficiency scores is across L1s, 2.6% across L2s and 7.9% across countries. Summing these up (see e.g. Goldstein, 2011), we observed that 29.7% of the total variance is to be attributed to country and language characteristics. Accordingly, the remaining variance at the learner level was estimated at 70.3% of the total variance in speaking scores (the fourth component).

In the next steps, we added fixed level 1 (learner characteristics) and level 2 (context characteristics) explanatory variables to the cross-classified design of languages by countries (see Table 2). Given the sizeable amount of data used in this study, effects were tested at the  $\alpha = .01$  level.

## IV Results

### I Main effects

Initially, the learner characteristics gender, length of residence, age of arrival, and years of education were added to the null model. This reduced the likelihood ratio with 605, against 5 degrees of freedom which is highly significant. On average, women scored almost 4 points ( $B = 3.99$ ,  $SE = .49$ ,  $p < .001$ ) higher than men did. The effects of age of arrival ( $-.49$ ,  $SE = .03$ ,  $p < .001$ ), length of residence ( $.63$ ,  $SE = .04$ ,  $p < .001$ ), and years of education ( $2.90$ ,  $SE = .24$ ,  $p < .001$ ) were all highly significant, and in the predicted direction: the older language learners were at their arrival the lower their La Dutch proficiency scores; if length of residence increases La Dutch speaking proficiency increases; the more education they have received the higher their La Dutch speaking skills. Instigated by previous research (van der Slik et al., 2015) we also included the interaction term of education with sex, but we were unable to replicate the expected positive outcome since the effect is not significant ( $.84$ ,  $SE = .52$ ,  $p = .111$ ). These outcomes corroborate the findings found in previous studies (Schepens et al., 2013a, 2016; van der Slik, 2010; van der Slik et al., 2015) using a different database to a large extent.

In Model 2 of Table 2, country of origin characteristics have been included. The improvement of fit was significant with 58 against 6 degrees of freedom, again being highly significant. The effect of educational accessibility as measured by gross enrolment in 2005 ( $B = .11$ ,  $SE = .04$ ,  $p = .009$ ) was significant and indicates that immigrants from countries with secondary schooling being more accessible score higher on the Dutch speaking test than immigrants from countries with lower accessible secondary schooling. Immigrants from African countries ( $B = -9.16$ ,  $SE = 3.52$ ,  $p = .012$ ), South Asia ( $B = -6.66$ ,  $SE = 3.03$ ,  $p = .031$ ), South and Central America ( $B = -3.09$ ,  $SE = 3.25$ ,  $p = .344$ ), and the Middle East ( $B = -7.34$ ,  $SE = 2.92$ ,  $p = .015$ ) did not differ significantly from those originating from Western countries. East Asian immigrants, however, did score lower ( $B = -10.52$ ,  $SE = 3.76$ ,  $p = .009$ ),

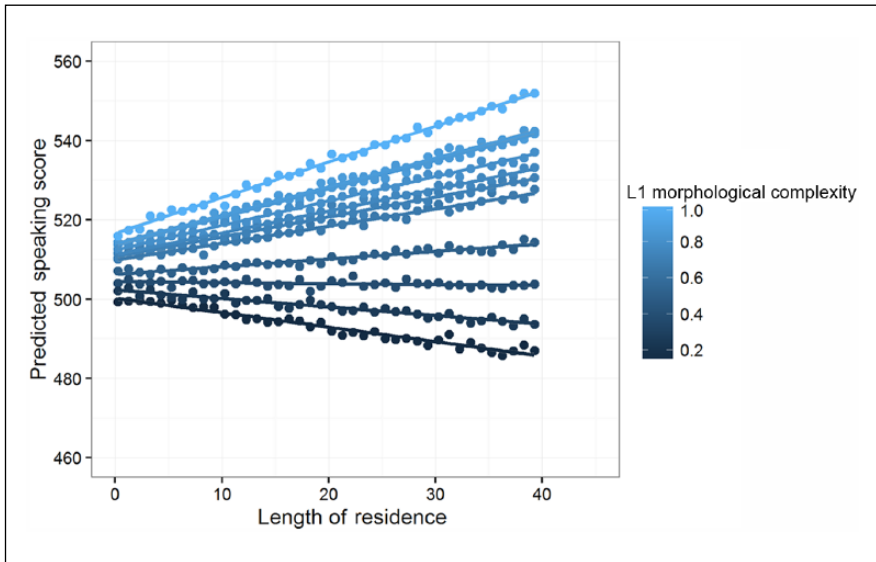
Next, the inclusion of the morphological complexity of the L1, Model 3, resulted in a significant reduction of the deviance with 24 against one degree of freedom. Immigrants with mother tongues that are morphologically less complex linguistically than Dutch had lower scores on speaking skills than immigrants whose first languages are equally or more complex morphologically ( $B = 29.45$ ,  $SE = 6.33$ ,  $p < .001$ ) in the sense that learners with L1s at least as complex morphologically as Dutch score almost 30 points higher than learners with L1s that are less complex as Dutch with regard to the eight morphological features studied here. It can also be observed that by including L1 morphological complexity, the relatively low La Dutch speaking scores of immigrants from East Asia (Model 2:  $B = -10.52$ ,  $SE = 3.76$ ) were levelled to a non-significant difference (Model 3:  $B = -1.86$ ,  $SE = 4.14$ ). Apparently, part of the regional differences in La Dutch speaking scores are to be ascribed to differences in L1 morphological complexity.

In model 4, we also included the measure of morphological complexity of the L2.<sup>1</sup> This posed a challenge however, since not all immigrants have indicated to have command of an L2. In order to overcome this issue, we, as already noted, decided to assign the value of zero for the L2 complexity scale for monolinguals (this approach defines the monolinguals as a separate group; see Schepens et al., 2016 for further explanation). The main effect of morphological complexity of the L2 did not prove to be significant, however ( $B = 3.65$ ,  $SE = 2.32$ ,  $p = .125$ ).

## 2 Interaction effects of morphological complexity

We examined all first-order interaction effects of morphological complexity with learner and country characteristics. We adopted the following strategy. First, we tested all interaction effects of the morphological complexity of the L1 and L2 with learner characteristics. Only the significant interaction terms were included in Model 5 and discussed.

The outcomes of Model 5 reveal that the morphological complexity of the L1 interacts significantly with length of residence ( $B = 1.45$ ,  $SE = .19$ ,  $p < .001$ ), while morphological complexity of the L2 interacts with age of arrival ( $B = .26$ ,  $SE = .08$ ,  $p = .001$ ). The interaction of length of residence with morphological complexity of the L1 signifies, as has been visualized in Figure 1,<sup>2</sup> that the effect of length of residence can range from hampering to beneficial, depending on the morphological complexity of the L1. This interaction effect qualifies both the main effect of LoR and morphological complexity. For morphological complexity, Figure 1 makes clear that the order always goes from less



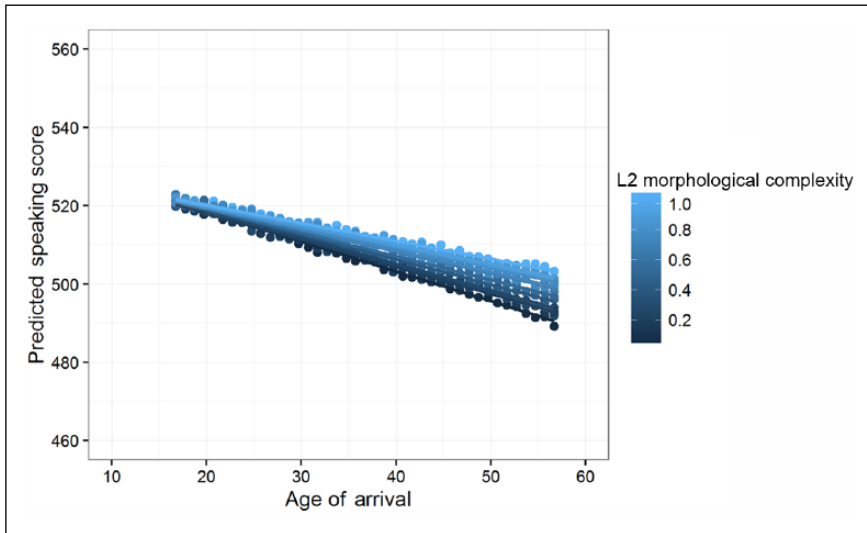
**Figure 1.** The interaction effect of length of residence and L1 morphological complexity in predicting La Dutch speaking proficiency.

to more complex. The character of LoR changes, because it turns out to be even negative for learners of less complex languages when they have a longer LoR. These learners didn't succeed in using the Dutch context to improve their level of speaking proficiency. Fossilization effects may play an adverse role here (Higgs and Clifford, 1982; Lardiere, 1998a, 1998b; Sims, 1989). We return to this in the discussion.

Finally we found, somewhat surprisingly as no interaction was found for the L1, the morphological complexity of the L2 to interact with the age of arrival on speaking proficiency in Dutch, as is shown in Figure 2. It can be observed that the negative effect of age of arrival is mitigated when adult language learners have a command of an L2 with higher morphological complexity, although the impact of the L2 is less strong than the impact of the L1. The overall L2 effect seems to have a comparable decline as the L1 effect, but it is not outspoken enough to be significant as a main effect. In sum, morphological complexity of both the L1 and the L2 affect the learnability of Dutch as an additional language, although these effects are moderated by AoA and LoR.

## V Discussion and conclusions

Our first two research questions address the application of the morphological complexity measure we used earlier. First, we found that the WALS morphological features, studied in the Schepens et al. (2013a) study, showed similar correlations with Dutch speaking test scores, when using a new data base. Thus, adult language learners whose L1 has a less complex morphological feature configuration compared to Dutch have more learning difficulties in acquiring Dutch the less complex their L1 is. Second, we



**Figure 2.** The interaction effect of age of arrival and L2 morphological complexity in predicting La Dutch speaking proficiency.

focused on a subset of eight morphological features to obtain a reduced morphological complexity measure and it turned out that this measure matches the performance of the measure that comprises and weights 16 WALS features. We did not lose predictive accuracy, as the correlation between the original measure and the reduced measure was .973 ( $N = 33$ ).

We included in our study the additional group of L1 Afrikaans learners of Dutch that we could not include in our earlier study, by using newly obtained typological information (see Appendix 1); the L1 morphological complexity score of Afrikaans is .25 (only two out of the eight features are equal or more (1) complex than Dutch); German scores 1.00 in comparison to Dutch (eight out of eight are equal or more (1) complex than Dutch). Including the feature values of Afrikaans in our analyses, and correlating the feature values with the aggregated speaking scores for the 33 L1s (thus including Afrikaans) resulted in lower correlations. Additional analyses, not presented here, showed that Afrikaans is an outlier as it combines low morphological complexity scores with high speaking skill scores. This does not have to come as a surprise if one realizes that the lexical overlap between Afrikaans and Dutch is more than 90%. It also signifies that other, additional measures of linguistic distance need to be used, like, for example, lexical (van der Slik, 2010; Schepens et al., 2013b) or phonological ones (Schepens, 2015) in order to explain the variance in speaking proficiency in an additional language. The special position of Afrikaans as a daughter language in fact confirms the special role of morphology, as Afrikaans speakers score lower than German speakers, despite their lexical and phonological advantages. The remaining morphological features behaved similarly as in Schepens et al. (2013a), i.e. they had only weak correlations with Dutch speaking proficiency scores.



Apparently, the eight specific morphological features possess discriminative power, i.e. can be used to predict if adult language learners will acquire higher or lower speaking skills in Dutch as an additional language. Does this imply that the eight morphological features uncovered here for Dutch as an additional language, will also be found when the proficiency scores of adult language learners of another target language than Dutch will be studied? This is impossible to tell without further empirical research. Conceivably, other target morphological target features may be present, but it is conceivable as well that specific subsets of features appear to be quite robust. It might perhaps be assumed that for target languages with substantial lexical and morphological overlap the same morphological features will prevail. In Appendix 1, we provide information on the feature values of six potential target additional languages. When larger databases for other target languages were available, our approach might be replicated relatively easily.

In our third research question, we hypothesized a main effect of L2 morphological complexity. We found a non-significant main effect of L2 morphological complexity on La Dutch learnability (in contrast to our previous study, Schepens et al., 2016), in the same direction as the L1 morphological complexity. We return to the L2 morphology effect in discussing the age-bound factors.

In our fourth research question, we hypothesized moderating effects of age of arrival (AoA) on the way L1 and L2 morphological complexity influence La learnability. We found a significant interaction effect of age of arrival with L2 morphological complexity on La Dutch learnability, but no interaction for L1 morphological complexity. The L2 by AoA interaction effects gives a stronger linear decrease (a steeper negative slope) the less complex the L2 is. A higher L2 complexity mitigates the AoA effect. It means that cognitive aging plays a role in second language acquisition bringing about a gradual declining process. The interaction effect found supports the validity of our morphological complexity measure, although we can only speculate why there was no interaction effect of L1 complexity and AoA. The most plausible explanation might be that the L1 complexity effect remains constant because the L1 is cognitively entrenched so early in life that no additional interaction effects show up.

We took LoR into account, to have the complete set of age-bound effects (LoR + AoA = the age of the participant). Length of residence (LoR) showed an outspoken interaction effect with morphological complexity of the L1. Generally, adult language learners' speaking skills in Dutch improve when they reside longer in the host country (there is a positive main effect for LoR). However, this conclusion is only warranted in case their L1's morphological complexity is high in comparison to Dutch. In case their L1's morphological complexity is low, the effect of length of residence even reverses and a higher LoR returns a negative impact on speaking skills in Dutch. We explained this by referring to fossilization effects. The weight of larger morphological differences increments when the time span between time of arrival and the actual taking of the speaking exam (defined as LoR) increases.<sup>3</sup> It might mean that during that period, immigrants with morphologically less complex L1s are more at risk of acquiring their own grammatically simplified or pidginized (Trudgill, 2008) version of Dutch than immigrants with morphologically more complex L1s. Although we do not have information on their target language learning experiences from the date of their arrival onwards, it seems plausible to assume that when immigrants have decided to actually take the exam, the duration of

having developed their own pidginized version Dutch will function as a threshold in acquiring Dutch morphosyntactic rules (Higgs and Clifford, 1982; Sims, 1989), and that threshold is thus assumed to become stronger the longer language learners have used their own grammatically simplified version. It is important to note that this interaction effect, independent from its explanation, corroborates the validity of our morphological complexity measure as it fits a theoretical framework in which the impact of morphological complexity unfolds over age. This process of a vigorous reduction of the (morphological) complexity of the language system also matches the Basic Variety as described by Klein and Perdue (1997). Crucially, they observe the emergence of this variety, simplified along a large number of linguistic parameters, in spontaneous adult language acquisition (outside of the classroom).

Nevertheless, we want to add another scenario that might explain the observed variation in the effect of length of residence in relation to the L1 morphological effect. Immigrants, once arrived, do not settle completely randomly across the host country. They typically concentrate where family and friends have migrated to earlier or near locations where migrant jobs are available (as already observed by Bartel, 1989). As a consequence, these close-knit groups of migrants perhaps feel less need to learn the language of their receiving country since they are able to continue their communication in their first language within their group to a large extent. The immigrants studied in the present study were not forced however to take the Dutch state exam; they took it voluntarily so it may safely be assumed they were well motivated to learn La Dutch. Despite their potentially strong motivation, we have detected that the effect of length of residence is moderated by the degree of morphological complexity of their L1.

The present study also contributes to what has been designated as the complexity, accuracy, fluency (CAF) literature (Housen and Kuiken, 2009). There is plethora of measures to operationalize these important concepts, but many of them are not successful (yet) (Housen et al., 2019). We prudently conclude that the way we handled the many morphological properties available through WALS gave us a concrete, restricted set of morphological features, relevant in adult learning of Dutch. We need to study in further detail how these features function in language learning, whether they signal deeper processing and performance problems or whether they fairly directly disclose direct difficulties. We have to add that we need comparative analyses on other language combinations as well. We sincerely hope that our approach is seen as promising enough to apply it to L2 English.

We conclude that our outcomes on adult La acquisition in relation to morphological complexity match the findings in the domains of Creole languages and language simplification processes through the influx of adult learners of a language. Our morphological complexity measure seems to be a tool worthwhile to be applied and tested in large-scale studies on other Las. The effects we found are too complex to be observed in small samples of learners, but narrowing down the morphological complexity of Dutch to eight features may obviously give direction to the type of morphological structures to be investigated in smaller studies on La Dutch. Our reduction of the morphological complexity measure in fact needs confirmation in fine-grained studies of actual acquisition processes. This might produce a thrilling exchange between acquisition studies based on big data and fine-grained acquisition studies on particular morphological properties. For Dutch, most of the eight morphological features were verb related, a domain with an interesting package of inflectional (contextual) and derivational (inherent) morphological

devices. The other three empirical studies in *Second Language Research* (De Clercq and Housen, 2019; Ehret and Szmrecsanyi, 2019; Palotti and Brezina, 2019) all produce evidence for the pivotal role of the verb in studying morphological complexity.

### Acknowledgements

We thank the two anonymous reviewers for their constructive comments and Theo Bongaerts for his editorial improvements of the text.

### Declaration of Conflicting Interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

### Notes

1. We also tested a model in which we made an unspecified distinction between monolingual and multilingual learners of Dutch. It proved that the model, presented here, fitted the data significantly better because the improvement in fit was 7 against 1 *DF*.
2. Figures 1 and 2 were produced by means of the *merMod* function in *R*.
3. Unlike for English, it is uncommon for Dutch to be acquired previous to arriving in the country (van der Slik, 2010).

### References

- Arends J (2001) Simple grammars, complex languages. *Linguistic Typology* 5: 180–82.
- Bartel A (1989) Where do the new immigrants live? *Journal of Labor Economics* 7: 371–91.
- Bell B, Morgan G, Schoeneberger J, Loudermilk B, Kromrey J, and Ferron J (2010) Dancing the sample size limbo with mixed models: How low can you go? Paper 197–2010 presented at the SAS Global Forum, Seattle, WA. Available from: <http://support.sas.com/resources/papers/proceedings10/197-2010.pdf> (January 2017).
- Bialystok E and Hakuta K (1994) *In other words: The science and psychology of second language acquisition*. New York: Basic Books.
- Birdsong D (2005) Interpreting age effects in second language acquisition. In: Kroll JF and De Groot AMB (eds) *Handbook of bilingualism: Psycholinguistic perspectives*. New York: Oxford University Press, pp. 109–27.
- Birdsong D and Flege JE (2001) *Regular–irregular dissociations in L2 acquisition of English morphology*. In: Do AH-J, Domínguez L and Johansen A (eds) *BUCLD 25: Proceedings of the 25th annual Boston University Conference on Language Development*. Boston, MA: Cascadilla Press, pp. 123–32.
- Bongaerts T, Mennen S, and van der Slik FWP (2000) Authenticity of pronunciation in naturalistic second language acquisition: The case of very advanced late learners of Dutch as a second language. *Studia Linguistica* 54: 298–308.
- Bongaerts T, Planken B, and Schils E (1995) Can late learners attain a native accent in a foreign language? A test of the critical period hypothesis. In: Singleton D and Lengyel Z (eds) *The age factor in second language acquisition*. Clevedon: Multilingual Matters, pp. 30–50.
- Booij G (2002) *The morphology of Dutch*. Oxford: Oxford University Press.

- Carliner G (2000) The language ability of US immigrants: Assimilation and cohort effects. *International Migration Review* 34: 158–82.
- Chiswick BR and Miller PW (2008) A test of the critical period hypothesis for language learning. *Journal of Multilingual and Multicultural Development* 29: 16–29.
- College voor Toetsen en Examens (2016) Homepage of Staatsexamens Nederlands als tweede taal (NT2). Utrecht: College voor Toetsen en Examens. Available at: <https://www.staatsexamensnt2.nl> (accessed January 2017).
- Council of Europe (2001) *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- De Clercq B and Housen A (2019) The development of morphological complexity: A cross-linguistic study of L2 French and English. *Second Language Research* 35(1): 71–97. This article appears in the current special issue of *Second Language Research*.
- Dryer MS and Haspelmath M (eds) (2011) *The world atlas of language structures [WALS] online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available at: <http://wals.info> (accessed January 2017).
- Ehret K and Szmrecsanyi B (2019) Compressing learner language: An information-theoretic measure of complexity in SLA production data. *Second Language Research* 35(1): 23–45. This article appears in the current special issue of *Second Language Research*.
- Goldstein H (2011) *Multilevel statistical models*. 4th edition. Chichester: Wiley.
- Hakuta K, Bialystok E, and Wiley E (2003) Critical evidence: A test of the critical-period hypothesis for second-language acquisition. *Psychological Science* 14: 31–38.
- Harada CN, Natelson Love MC, and Triebel K (2013) Normal cognitive aging. *Clinics in Geriatric Med* 29: 737–52.
- Heck RH, Thomas SL, and Tabata LN (2010) *Multilevel and longitudinal modeling with IBM SPSS*. New York: Routledge.
- Higgs TV and Clifford R (1982) The push toward communication. In: Higgs TV (ed.) *Curriculum, competence, and the foreign language teacher*. Lincolnwood, IL: National Textbook Co., pp. 57–79.
- Housen A, De Clercq B, Kuiken F, and Vedder I (2019) Introduction. *Second Language Research* 35(1): 3–21. This article appears in the current special issue of *Second Language Research*.
- Housen A and Kuiken F (2009) Complexity, accuracy and fluency in second language acquisition. *Applied Linguistics* 30: 461–73.
- Johnson JJ and Newport EL (1989) Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21: 60–99.
- Kellerman E and Sharwood Smith M (1986) *Crosslinguistic influence in second language acquisition*. New York: Pearson College Division.
- Klein W and Perdue C (1997) The Basic Variety (or: Couldn't natural languages be much simpler?). *Second Language Research* 13: 301–47.
- Krashen SD, Scarcella RC, and Long MH (1982) (eds) *Child–adult differences in second language acquisition*. Rowley, MA: Newbury House.
- Kreft IGG and de Leeuw J (1998) *Introducing multilevel modeling*. London: Sage.
- Kuijper H, Bergsma A, and Bechger T (2004) *Staatsexamen NT2: Het gewenste niveau: Deel 1 Behoeftepijling Programma II [State exam Dutch as a second language: The desired level: Part I Needs investigation Program II]*. Arnhem: Cito.
- Kusters W (2008) Complexity in linguistic theory, language learning and language change. In: Miestamo M, Sinnemäki K, and Karlsson F (eds), *Language complexity: Typology, contact, change*. Amsterdam, PA: Benjamins, pp. 3–22.
- Lado R (1957) *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor, MI: Michigan University Press.
- Lardiere D (1998a) Case and tense in the ‘fossilized’ steady state. *Second Language Research* 14: 1–26.

- Lardiere D (1998b) Dissociating syntax from morphology in a divergent end-state grammar. *Second Language Research* 14: 359–75.
- Lupyan G and Dale R (2010) Language structure is partly determined by social structure. *PLoS ONE* 5: e8559.
- Marinova-Todd SH, Marshall DB, and Snow CE (2000) Three misconceptions about age and L2 learning. *TESOL Quarterly* 34: 9–34.
- McWhorter JH (2001) The world's simplest grammars are creole grammars. *Linguistic Typology* 5: 125–66.
- McWhorter JH (2002) What happened to English? *Diachronica* 19: 217–72.
- Miestamo M (2008) Grammatical complexity in a cross-linguistic perspective. In: Miestamo M, Sinnemäki K, and Karlsson F (eds) *Language complexity: Typology, contact, change*. Amsterdam, PA: Benjamins, pp. 23–42.
- O'Grady W (2006) The problem of verbal inflection in second language acquisition. In: Nakano M and Park K-J (eds) *Proceedings of the 11th Conference of Pan-Pacific Association of Applied Linguistics*. Chuncheon: Kangwon National University, pp. 169–90.
- Odlin T (1989) *Language transfer: Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.
- Pallotti G and Brezina V (2019) Morphological complexity in written L2 texts. *Second Language Research* 35(1): 99–119. This article appears in the current special issue of *Second Language Research*.
- Paquot M (2019) The phraseological dimension in interlanguage complexity research. *Second Language Research* 35(1): 121–145. This article appears in the current special issue of *Second Language Research*.
- Salthouse T (2004) What and when of cognitive aging. *Current Directions in Psychological Science* 13: 140–44.
- Salthouse T (2009) When does age-related cognitive decline begin? *Neurobiological Aging* 30: 507–14.
- Salthouse T (2010a) *Major issues in cognitive aging*. New York: Oxford University Press.
- Salthouse T (2010b) Selective review of cognitive aging. *Journal of the International Neuropsychological Society* 16: 754–60.
- Saville-Troike M (2012) *Introducing second language acquisition*. 2nd edition. New York: Cambridge University Press.
- Schepens JJ (2015) *Bridging linguistic gaps: The effects of linguistic distance on the adult learnability of Dutch as an additional language*. Published PhD Thesis. Utrecht: LOT. Available at: [www.lotpublications.nl/Documents/383\\_fulltext.pdf](http://www.lotpublications.nl/Documents/383_fulltext.pdf) (accessed January 2017).
- Schepens JJ, van der Slik FWP, and van Hout RWNM (2013a) Learning complex features: A morphological account of L2 learnability. *Language Dynamics and Change* 3: 218–44.
- Schepens JJ, van der Slik FWP, and van Hout RWNM (2013b) The effect of linguistic distance across Indo-European mother tongues on learning Dutch as a second language. In: Borin L and Saxena A (eds) *Approaches to measuring linguistic differences*. Berlin: Mouton de Gruyter, pp. 199–230.
- Schepens JJ, van der Slik FWP, and van Hout RWNM (2016) L1 and L2 distance effects in learning L3 Dutch. *Language Learning* 66: 224–56.
- Sims WR (1989) Fossilization and learning strategies in second language acquisition. *MinneTESOL* 7: 61–72.
- Snow C and Hoefnagel-Höhle M (1977) Age differences in the pronunciation of foreign sounds. *Language and Speech* 20: 357–65.
- Szmrecsanyi B and Kortmann B (2012) Introduction: Linguistic complexity: Second language acquisition, indigenization, contact. In: Kortmann B and Szmrecsanyi B (eds) *Linguistic*

- complexity: *Second language acquisition, indigenization, contact*. Berlin/Boston, MA: De Gruyter, pp. 6–34.
- Trudgill P (2001) Contact and simplification: Historical baggage and directionality in linguistic change. *Linguistic Typology* 5: 371–74.
- Trudgill P (2008) Linguistic and social typology. In: Chambers JK, Trudgill P, and Schilling-Estes N (eds) *The handbook of language variation and change*. Malden, MA/Oxford/Carlton: Blackwell, pp. 707–28.
- van der Slik FWP (2010) Acquisition of Dutch as a second language: The explanative power of cognate and genetic linguistic distance measures for 11 West-European first languages. *Studies in Second Language Acquisition* 32: 401–32.
- van der Slik FWP, van Hout RWNM, and Schepens JJ (2015) The gender gap in second language acquisition: Gender differences in the acquisition of Dutch among immigrants from 88 countries with 49 mother tongues. *PLoS ONE* 10: e0142056.
- van Rensburg MCJ (1983) Nie-standaardvorme, variasiepatrone en Afrikaans uit die vorige eeu [Nonstandard forms, variation patterns and the Afrikaans of the past century]. In: Claassen G and van Rensburg MCJ (eds) *Taalverskeidenheid: 'n blik op die spektrum van taalvariasie in Afrikaans [Language diversity: A view on the spectrum of language variation in Afrikaans]*. Pretoria/Cape Town/Johannesburg: Academica, pp. 134–61.
- van Sluijs R (2013) Afrikaans. In: Michaelis SM, Maurer P, Haspelmath M, and Huber M (eds) *The survey of pidgin and Creole languages: Volume 1: English-based and Dutch-based languages*. Oxford: Oxford University Press, 285–96.
- Verhelst ND, Glas CAW, and Verstralen HHFM (1995) *One-parameter logistic model OPLM*. Arnhem: Cito.
- Weinreich U (1963) *Languages in contact*. The Hague: Mouton.
- World Bank (2015) *World DataBank*. Washington, DC: World Bank. Available at: <http://data-bank.worldbank.org/ddp/home.do> (accessed January 2017).
- Wray A and Grace GW (2007) The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua* 117: 543–78.

## Appendix I

### *The morphological features*

In the table below we provide information on the Lupyan and Dale (2010) feature values of six potential target additional languages. The table explicates how morphological complexity works when using these features. In the right column, we have copied the findings from Lupyan and Dale (2010) for the 28 features selected from the WALS (world atlas of language structures; Dryer and Haspelmath, 2011) data base (the contents of the feature number can be found on the internet). This column ranks the complexity categories, from less to more complex, ranging from 1 to a maximum of 4. These numbers are given in the language columns.

For feature number 22 (Inflectional Synthesis of the verb), for example, Chinese, having the feature value 1, has been assigned the lowest complexity category, a value of 1. Afrikaans, Dutch, English, and German have a value of 2, and Spanish has a 3.

The remaining complexity scores for the other features have been assigned in the same way. As a result, Dutch and German, and English have similar morphological complexity scores, Spanish has the highest scores, while Afrikaans is less complex (there are many missing values though). Chinese is scoring the lowest. See the figures in the bottom rows.

It is important to note here again that the reduced morphological complexity measure we used is restricted to the eight features in bold face and is computed in relation to the Dutch values. It is  $1$ : the number of values less complex than Dutch/the number of features for which values are available. That gives the following outcomes: German = English = Spanish =  $1$ ; Afrikaans =  $.250$ ; Chinese =  $.143$ .

Feature value codes for six target languages, as compared with the WALS feature codes, found by Lupyan and Dale (2010). In bold the selected morphological features used for the reduced morphological complexity measure (and their means and SDs in bold).

Feature number	Dutch	German	Afrikaans	English	Spanish	Chinese	WALS increasing complexity*
20	2	2	2	2	2	–	2, 4 > 1 > 5–7
<b>22</b>	<b>2</b>	<b>2</b>	<b>2</b>	<b>2</b>	<b>3</b>	<b>1</b>	1 > 2 > 3–7
<b>26</b>	<b>2</b>	<b>2</b>	<b>1</b>	<b>2</b>	<b>2</b>	–	1 > 2 > 3–6
28	2	2	–	1	2	1	1, 2 > 3 > 4
<b>29</b>	<b>2</b>	<b>2</b>	<b>1</b>	<b>2</b>	<b>2</b>	<b>1</b>	1 > 2 > 3
34	2	2	–	2	2	3	3 > 6 > 2, 4, 5
36	1	2	2	1	1	2	4 > 1–3
38	2	2	–	2	2	–	5 > 1, 2, 4 > 3
41	2	1	2	2	3	2	1 > 2 > 3–5
48	1	1	–	1	1	1	1, 2 > 3
49	1	2	–	2	1	1	1 > 2–8
57	1	1	–	1	1	–	4 > 2
59	1	1	–	1	1	1	1 > 2 > 3, 4
65	2	2	2	2	1	1	1 > 2
<b>66</b>	<b>2</b>	<b>2</b>	<b>2</b>	<b>2</b>	<b>2</b>	<b>1</b>	4 > 1 > 2 > 3**
67	1	1	1	1	2	1	2 > 1
70	2	4	–	1	3	1	5 > 4 > 1 > 2, 3
73	1	1	–	1	1	1	2 > 1
74	1	1	–	1	1	1	2, 3 > 1
75	1	1	–	1	1	1	1, 3 > 2
76	1	1	–	1	1	1	1, 3 > 2
77	2	2	2	1	1	1	1 > 2, 3
<b>92</b>	<b>2</b>	<b>2</b>	<b>1</b>	<b>2</b>	<b>2</b>	<b>1</b>	1–4 > 5
98	1	1	–	1	1	1	1, 2 > 3, 4
<b>100</b>	<b>2</b>	<b>2</b>	<b>1</b>	<b>2</b>	<b>2</b>	<b>1</b>	1 > 2,3 > 4,5,6
101	1	1	–	1	2	1	1, 4, 5 > 2, 3
<b>102</b>	<b>2</b>	<b>2</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>1</b>	1 > 2 > 3, 5 > 4
<b>112</b>	<b>3</b>	<b>3</b>	<b>2</b>	<b>3</b>	<b>3</b>	<b>3</b>	1, 4 > 6 > 2 > 3, 5
Mean(Total)	1.61	1.71	1.57	1.54	1.75	1.25	
SD (Total)	.56	.71	.51	.58	.75	.61	
<b>Mean</b>	<b>2.13</b>	<b>2.13</b>	<b>1.38</b>	<b>2.13</b>	<b>2.38</b>	<b>1.29</b>	
<b>SD</b>	<b>.35</b>	<b>.35</b>	<b>.52</b>	<b>.35</b>	<b>.52</b>	<b>.76</b>	

Notes. \* The feature value lowest in complexity (utmost left) is assigned the code 1, next code 2, and so on (Lupyan and Dale, 2010). \*\* We did not follow Lupyan and Dale (2010) in this case because, according to the pattern they observed, 'past/ non-past, no remoteness distinction' (1) would be morphologically less complex than 'no grammatical marking' (4). We switched these categories.

## Appendix 2

### *Description of the eight morphology features*

The table below briefly describes the features used in Table 2, as adapted from <http://wals.info/chapter/#> where # is the world atlas of language structures (WALS; Dryer and Haspelmath, 2011) feature number/chapter.

- 
- |     |  |
|-----|--|
| 22  | <p>Inflectional synthesis of the verb: Grammatical categories like tense, voice, or agreement can be expressed either by individual words or by affixes attached to some other word (or the stem of a word). If a word combines with affixes, the resulting construction is said to be synthetic; if not, it is said to be analytic.<br/>Feature values: (1) 0–1 category per word, (2) 2–3 categories per word, (3) 4–5 categories per word, (4) 6–7 categories per word, (5) 8–9 categories per word, (6) 10–11 categories per word, (7) 12–13 categories per word</p> |
| 26  | <p>Prefixing vs. suffixing in inflectional morphology: The overall extent to which languages use prefixes versus suffixes in their inflectional morphology.<br/>Feature values: (1) Little or no inflectional morphology, (2) Predominantly suffixing, (3) Moderate preference for suffixing, (4) Approximately equal amounts of suffixing and prefixing, (5) Moderate preference for prefixing, (6) Predominantly prefixing.</p>  |
| 29  | <p>Syncretism in verbal person / number marking: This describes instances of syncretism in the inflectional marking of subject person in verbs.<br/>Feature values: (1) No subject person/number marking, (2) Subject person/ number marking is syncretic, (3) Subject person/number marking is never syncretic.</p>   |
| 66  | <p>Past tense: The basic dichotomy here is between languages that mark the past/non-past distinction grammatically (including marking by periphrastic constructions) and those that do not.<br/>Feature values: (1) Past/non-past distinction marked; no remoteness distinction, (2) Past/non-past distinction marked; 2–3 degrees of remoteness distinguished, (3) Past/non-past distinction marked; at least 4 degrees of remoteness distinguished, (4) No grammatical marking of past/non-past distinction.</p>   |
| 92  | <p>Polar questions: Polar questions are ones that elicit the equivalent of a ‘yes’ or ‘no’ response. They do not include interrogative affixes on verbs.<br/>Feature values: (1) Question particle at end of sentence, (2) Question particle in second position in sentence, (3) Question particle with other position, (4) Question particle in either of two positions, (5) No question particle.</p>  |
| 100 | <p>Alignment of Verbal Person: Marking Languages differ not only in regard to the number and nature of the arguments displaying verbal person marking, but also in the alignment of the person markers.<br/>Feature values: (1) Neutral alignment, (2) Accusative alignment, (3) Ergative alignment (4) Active alignment, (5) Hierarchical alignment, (6) Split alignment.</p>   |
| 102 | <p>Verbal person markings can depend on the agent, the patient, both agent and patient, or either one.<br/>Feature values: (1) No person marking of any argument, (2) Person marking of only the A argument, (3) Person marking of only the P argument, (4) Person marking of the A or P argument, (5) Person marking of both the A and P arguments.</p>   |
| 112 | <p>Coding of negation: Expresses the nature of morphemes signalling clausal negation in declarative sentences. By clausal negation is meant the simple negation of an entire clause or the negative clause corresponding to the affirmative clause.<br/>Feature values: (1) Negative affix, (2) Negative particle, (3) Negative auxiliary verb, (4) Negative word, unclear if verb or particle, (5) Variation between negative word and affix, (6) Double negation.</p>  |
-