Learning and Individual Differences 62 (2018) 1–11

Contents lists available at ScienceDirect

Learning and Individual Differences

journal homepage: www.elsevier.com/locate/lindif

Transfer of decoding skills in early foreign language reading

Marco van de Ven*, Marinus Voeten, Esther G. Steenbeek-Planting, Ludo Verhoeven

Behavioural Science Institute, Radboud University, The Netherlands

ARTICLE INFO

Keywords:
Reading
Early foreign language reading
Decoding fluency
Secondary education
Structural equation modeling

ABSTRACT

This longitudinal study investigated cross-linguistic transfer from native to foreign language decoding abilities in 787 Dutch first-year students in two differential tracks (high vs low) of secondary education. On two occasions, with a six months interval, we tested the students' word and pseudoword decoding fluency in their native language (Dutch) and their word decoding fluency in two foreign languages, English (L2) and French (L3). Our findings indicated that students' English word decoding development primarily depended on Dutch word decoding fluency. The development of their French decoding skills was mainly dependent on their Dutch pseudoword decoding fluency, and in the higher educational track also on their L2 word decoding ability. It is concluded that there is evidence of linguistic interdependencies in the development of word decoding in the native language and the subsequent development of word decoding in the foreign language. Theoretical as well as practical implications are discussed.

1. Introduction

A large body of research indicates that there is a strong correlation between native and non-native literacy development (e.g., Cheung, 1996; Durgunglu, Nagy, & Hancin-Bhatt, 1993; Geva, Wade-Woolley, & Shany, 1997; Lindsey, Manis, & Bailey, 2003; Melby-Lervåg & Lervåg, 2011; Van Gelderen et al., 2004). Less is known about language-specific, lower-order efficiency of information processing like word decoding in foreign language reading. When readers have substantial native-language reading experience and the languages have similar orthographic structures, a correlation is to be expected between native and non-native word decoding (Koda, 1996). A few studies indeed established a clear relationship between L1 and L2 word decoding development (e.g., Bernhardt, 2005; Commissaire, Duncan, & Casalis, 2011; Sparks, Patton, Ganschow, Humbach, & Javorsky, 2008). The study of non-native word decoding is important because of the relationship between word decoding efficiency and reading comprehension. A certain level of word decoding efficiency is necessary for the attainment of reading comprehension. L1 word decoding is a predictor of L1 reading comprehension ability, especially in less experienced readers (Braze et al., 2015; Freed, Hamilton, & Long, 2017; Protopapas, Mouzaki, Sideridis, Kotsolakou, & Simos, 2013), and L1 reading comprehension is a strong predictor of L2 (Van Gelderen et al., 2004; Van Gelderen, Schoonen, Stoel, De Glopper, & Hulstijn, 2007) as well as L3 reading comprehension ability (Van Gelderen et al., 2003). Yet, the development of word decoding is an important object of study in its own right as well (Koda, 1996). L1 word decoding fluency has been shown to continue to increase during the first half-year of secondary education (i.e., when adolescents are 12–13 years old; Van de Ven, Voeten, Steenbeek-Planting, & Verhoeven, 2017), the period that students (begin to) receive systematic foreign language instruction. The present study investigated whether the development of L2 English and L3 French (two foreign languages taught in Dutch secondary education) word decoding skills are associated with proficiency in Dutch (L1) decoding skills, and whether there is a relationship between L2 and L3 word decoding development during that very early phase of foreign language instruction. Further, we tested whether these relations are similar across different developmental stages of L2/L3 learning. Dutch students are allocated to educational tracks at the beginning of secondary education, based on their past performance on Dutch and mathematics (i.e. resulting from school decisions). These educational tracks allowed us to assess L2/L3 word decoding development across different stages of L2/L3 learning. Theoretical models suggest that the relationship between L1 and L2 development is a complex one, which may vary as a function of cultural/educational background, and may also be influenced by linguistic factors (which may be related; e.g., Cummins, 1979; Ganschow, Sparks, & Javorsky, 1998; Koda, 2005, 2008). First of all, the linguistic coding differences hypothesis (Sparks & Ganschow, 1991) proposed that the same underlying language learning mechanisms determine the level of L1 and L2 competence learners can attain. In a similar vein, the linguistic interdependence model (Cummins, 1979) claims that L1 and L2 literacy development are highly related. According to this model, L1 competence partially determines the level of L2 language competence that learners can obtain (i.e. the

* Corresponding author at: Radboud University Nijmegen, P.O. Box 9104, 6500 HD Nijmegen, The Netherlands.
E-mail address: Marco.vandeVen@pwo.ru.nl (M. van de Ven).

https://doi.org/10.1016/j.lindif.2018.01.005
Received 4 April 2017; Received in revised form 4 January 2018; Accepted 5 January 2018
1041-6080/

“developmental hypothesis”). Further, learners' L1 competence needs to exceed a first threshold to avoid subtractive bilingualism, and a second, higher threshold to foster additive bilingualism (i.e. the “threshold hypothesis”). In a more recent approach to study the relation between L1 and L2 learning, Koda (2005, 2008) developed the transfer facilitation model, which proposed that L2 learning can be enhanced by metalinguistic awareness skills acquired in the L1. The influence of L1 on L2 learning is moderated by L1-L2 orthographic distance, as well as by exposure to L2 visual (i.e. orthographic) input.

Research suggests that L2 word decoding fluency strongly predicts L3 word decoding development, because L2 decoding reflects general language-learning aptitude (e.g., Dörnyei, 2005). However, in L3 learning, the situation is somewhat more complicated. The typological primacy model (Rothman, 2010) proposed that, in beginning L3 learners, the magnitude of cross-linguistic transfer is moderated by the typological similarities between the languages involved, regardless of language status (L1 or L2). Moreover, in beginning L2/L3 learners, languages in the multilingual mental lexicon are strongly interconnected (e.g., De Bot, 2004; Sánchez, 2014; Wei, 2006), and bilinguals activate the L1 and L2 simultaneously, regardless of whether they use one or both languages (e.g., Bialystok, 2013; Jared & Kroll, 2001). Consequently, there are fundamental qualitative differences between L3 and L2 learning (e.g., Cenoz & Genesee, 1998). With respect to L3 reading, research suggests that bilingualism benefits L3 reading development if learners are sufficiently proficient in the L1 and L2, and these benefits appear to be (partly) due to enhanced metalinguistic awareness in proficient bilinguals (Rauch, Naumann, & Jude, 2011), in line with the transfer facilitation model (Koda, 2005, 2008).

In addition, we had specific predictions regarding the influences of L1 word and pseudoword decoding ability on the development of word decoding skills in a non-native language, depending on the stage of development, based on the Dual Route Cascaded model (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). The model distinguishes between the lexical (direct) and the non-lexical (indirect) route. Beginning L2 learners are predicted to primarily use the non-lexical route in the L2, in other words rely on grapheme-phoneme correspondence rules, during word decoding. The lexical traces of many L2 words in the beginning learners' mental lexicon are still weak and many words have not yet been fully lexicalized. In moderately advanced learners, these lexical traces are much stronger, and most frequent L2 words have been lexicalized. As a consequence, these learners are predicted to rely more heavily on the lexical route (lexical access, measured primarily by word decoding) than on the non-lexical route (measured primarily by pseudoword decoding) during L2 word decoding.

The present study aimed to investigate the relationships between Dutch word decoding ability and the developments of English (L2) and French (L3) word decoding fluency during Grade 7, the first year of secondary education in the Netherlands. At the onset of Dutch secondary education (ages 12–13), students are divided into tracks, based on their Dutch language and math skills. We distinguished two tracks of learners, where Track 1 contained students in pre-vocational education, and Track 2 those in higher level of secondary education and pre-university education. We distinguished these two educational tracks to compare development across different levels of L1, L2, and L3 proficiency. Admittedly, some students in the lowest educational track were not sufficiently competent to perform the French word decoding test by the time of the first test session, and therefore did not complete this test at Time 1.

The English and French orthography differ in their relation to the Dutch orthography. As shown by Seymour, Aro, and Erskine (2003), Dutch and English share a complex syllable structure, whereas French, in contrast, has a relatively simple syllable structure. On the other hand, Dutch and French have much more shallow orthographies than English does. Further, although English is a Germanic language, it has borrowed approximately 50% of its lexicon from French and Latin (e.g., Gray & Atkinson, 2003) and may therefore be regarded as lexically more similar to French than to Dutch. In terms of suprasegmental characteristics, however, English appears to be more similar to Dutch than to French (e.g., see Delattre, 1963; Domahs, Plag, & Carroll, 2014).

Although both English and French are being taught in Dutch secondary education, students' average amount of exposure to the two languages differs significantly. In the Dutch educational system, children receive an introductory English language course, approximately 15–20 min per week (Herder & De Bot, 2005), during grades 5 and 6 (the end of Dutch primary education), and an intensive English language course (approximately 2 h per week) and an introductory French language course (approximately 1.5 h per week) during the first three years of Dutch secondary education (Eurydice, 2005). Given the fact that children receive an introductory French language course starting at the beginning of secondary education, we measured early emergent French reading skills. In both educational tracks, instruction for both languages is largely based on the communicative approach, combined, especially in the higher track, with elements of language awareness (e.g., grammar). In addition, in the Netherlands, French is mostly solely learned in schools, whereas English is also learned outside the school setting, e.g., through the media. We assessed the influence of two L1 decoding skills, pseudoword decoding fluency (PWDF) and word decoding fluency (WDF), on the development of L2 and L3 word decoding fluency. In line with the Dual Route Cascaded model, PWDF reflects the success in using the non-lexical route (i.e., applying grapheme-phoneme correspondence rules), whereas WDF reflects the success in using the lexical route (i.e., lexical access). Our research questions can thus be summarized as

1. To what extent is there a positive relation between the level of English (L2) and French (L3) word decoding fluency on the one hand and Dutch (L1) word and pseudoword decoding fluency on the other?
   As predicted by the theoretical model by Sparks and Ganschow (1991), we hypothesized positive relations between foreign-language and native-language word decoding fluency. More specifically, we hypothesized that level of English WDF would be most strongly related to level of Dutch WDF (rather than PWDF), since our participants were elementary learners of English, with one or two years of experience; we expected them to use both the lexical and non-lexical route during English word decoding. On the other hand, we predicted that level of French WDF would show a stronger relation with level of PWDF (rather than WDF), given that our participants were still only beginning learners of French; we expected them to predominantly use the non-lexical route during French WDF.
   Furthermore, we expected all relationships to be the same for the two educational tracks.
2. To what extent is there a positive relation between the development during the school year of English (L2) and French (L3) word decoding fluency on the one hand and Dutch (L1) word and pseudoword decoding fluency on the other?
   We expected the development of foreign-language word decoding fluency to show similar relationships with native-language word and pseudoword decoding fluency as hypothesized in connection with the first research question. Based on Cummins' (1979) threshold hypothesis, we expected developmental differences between Tracks 1 and 2, especially for French. More specifically, we expected to find positive effects of L1 on L2/L3 development only for students who are sufficiently competent in the L1.
3. Is there an additional influence of English (L2) word decoding fluency on French (L3) word decoding fluency?
   We hypothesized that level of English WDF would predict the development of French WDF, because English WDF reflects general language-learning aptitude (e.g., Dörnyei, 2005). Based on Cummins' (1979) threshold hypothesis, any positive transfer effects of L2 on L3 decoding development may only occur for relatively

proficient L2 learners.

4. Are there average developmental differences in terms of English (L2) and French (L3) word decoding fluency between the two educational tracks?
   We expected the high track to outperform the low track and we expected the high track to show stronger average improvement than the low track.

With respect to these research questions, we distinguished between level of and change in word decoding fluency. For Dutch word and pseudoword decoding fluency, we had several measures, which were combined into two latent variables; one for word decoding fluency and another one for pseudoword decoding fluency. We defined latent change factors to capture the difference between Time 1 and Time 2 in each of the two latent factors. This enabled us to study the effect of both level of (the Time 1 latent factors) and change in Dutch pseudoword and word decoding fluency during the school year. The level of foreign-language word decoding fluency was determined at Time 1. The change during the school year was investigated by having foreign-language word decoding fluency at Time 2 as criterion variables while using the same measures at Time 1 as covariates. Since we only had one variable for both English and French word decoding fluency, we could not define latent variables for foreign language decoding fluency.

2. Method

2.1. Participants

Participants were 787 students from six schools in Dutch lower and intermediate level pre-vocational education ($n = 459$ students) and higher level of secondary education and pre-university education ($n = 328$ students); only students with Dutch as their native language were selected in the present study. There were 17 students missing completely the first measurement occasion and 18 students missing completely the second measurement occasion. Further missing values appeared, because 159 students (20%) did not complete all tests; especially the French word decoding test was missing at Time 1 for 153 students. The main reason for these missing values was that these students had not yet received sufficient instruction in French to take the test. These students scored on average lower on all other tests than did the students with a valid score on French. The missing data were dealt with by using full information maximum likelihood estimation (Enders, 2010), such that all available information of all students could be included in the analyses.

The mean age of the students during the first test session was 12 years and 11 months (SD 5.7 months) in the first track and 12 years and 9 months (SD 5.0 months) in the second. On average, students were older in the first track than in the second. This (small) age difference may be related to the fact that students with learning difficulties were more frequent in the first track than in the second. This is confirmed by the number of students diagnosed with dyslexia: 10.7% in the first track and 5.3% in the second track.

The first session took place in December (three months after the beginning of the school year) and the second in June (i.e., near the end of the school year) during the first year in secondary education. We selected December as the period for our first test session, because most students had by then received three months of French language instruction.

Before the first measurement occasion, the students had received approximately two years of English language instruction. English lessons in primary education focused primarily on basic oral and writing communication skills. In addition to English, students had received approximately three months of French language instruction. However, the latter was not true for almost 20% of the students. In the Netherlands, no formal IRB approval was required, though appropriate permissions and consent were obtained.

2.2. Instruments

We tested students' L1 word and pseudoword decoding skills, and their English and French decoding skills, all by means of paper and pencil tasks. All tests were administered twice, with a six months interval. Detailed descriptions of the tests are provided below.

2.2.1. (Pseudo)word decoding fluency

Word decoding skills were assessed with the (standardized) Word Decoding Test (Verhoeven, 2004). Participants were instructed to read aloud words from a card as quickly and accurately as possible. They read a total of eight cards, and they had 1 min for each card. The first four cards consisted of existing Dutch words. The first card was composed of 150 CVC (consonant-vowel-consonant) words, the second of 150 C(C)VC(C) words, the third of 120 disyllabic words and the last card of 120 trisyllabic and quadrisyllabic words. The last four cards consisted of Dutch pseudowords, and were composed of the same structures as the initial four cards. The score for each subtest was the number of (pseudo)words read correctly within 1 min. We concentrated on decoding fluency, because reading accuracy is rarely a problem in Dutch. The test-retest correlations for pseudoword and word decoding fluency appeared rather high for each card (0.88 to 0.92 for word reading and 0.87 to 0.91 for pseudoword reading).

2.2.2. English word decoding

We measured English word decoding by means of a standardized test (Steenbeek-Planting, Kleijnen, & Verhoeven, 2008a). Dutch-English cognates and homonyms were excluded from the test. Moreover, since Dutch and English share several orthographic rules, English words that could be pronounced correctly with Dutch orthographic knowledge, or with French orthographic knowledge (i.e. French loanwords; the test contained eight of these) only were excluded as well. Words in the English reading test increased in length and decreased in word frequency. Participants were instructed to read the words as quickly and accurately as possible. The test consisted of 108 words in total, and the score was the number of words read correctly within 1 min. The test assistants who were responsible for scoring this test (and the French decoding test below) were provided with digital recordings containing the appropriate pronunciations of the target words. In order to establish the inter-rater reliability for the seventeen test-assistants who scored the data, one independent test assistant re-rated 460 tests administered during the pre-test and 410 tests during the post-test. The inter-rater reliabilities of one test assistant during the pre-test and another test assistant during the post-test were below 0.80, and a different test assistant therefore rescored these data. The average inter-rater reliability of the final data was 0.91 for the pre-test (SD = 0.06) and 0.88 for the post-test (SD = 0.03). The test-retest correlation coefficient was 0.88.

2.2.3. French word decoding

We measured French word decoding skills by means of a standardized test (Steenbeek-Planting, Kleijnen, & Verhoeven, 2008b). Similar to the English word decoding test, Dutch-French cognates and homonyms were excluded, as well as any words that could be pronounced correctly using Dutch orthographic knowledge. Further, the French decoding test contained five English-French full cognates, although, again, none of the French words were pronounced the same as the English counterpart. The test was composed of 100 words, which increased in word length and decreased in word frequency. Participants were again instructed to read the words as quickly and accurately as possible. The inter-rater reliability of one test assistant during the pre-test was below 0.80, and a different test assistant therefore rescored these data. The average inter-rater reliability of the final data was 0.89 for the pre-test (SD = 0.06) and 0.89 for the post-test (SD = 0.05). The test-retest correlation coefficient was 0.72.

2.3. Procedure

The tests were administered individually, during school hours, but not during lunch breaks. Students were seated in a quiet room. The English and French word reading tests were separated by an unrelated test that is not part of this study. The four Dutch word reading cards were always presented successively, while varying the order of the cards. The same holds for the four pseudoword reading cards. The tests were timed using a stopwatch. Each testing session lasted approximately 20 min. Students' responses were digitally recorded, in order to be able to check the interrater reliability. The inter-rater reliability (intraclass correlation) was minimally 0.80 for all tests (Kleijnen, Steenbeek-Planting, & Verhoeven, 2008).

2.4. Analytic method

Multiple-group Confirmatory Factor Analysis (CFA) was used to identify latent variables of Dutch word and pseudoword reading. For English and French word decoding ability, we only had one observed variable for each language at each time point; therefore, no latent variable could be constructed. Multiple-group analysis was used to account for the two educational tracks. A latent variable model was constructed using all eight cards of Dutch (pseudo)words at both time points.

Latent Difference Score (LDS) models (McArdle, 2009) were used to capture change in word and pseudoword reading between the two times of measurement during the first year of secondary education. A latent change score represents the difference between two measurements of a latent variable: a latent variable score at Time 2 minus the score on the same latent variable at Time 1: $\Delta\eta = \eta_2 - \eta_1$. An LDS can be defined in a structural equation model. The score on a latent variable, e.g., word decoding fluency, at Time 2 is represented as the score on word decoding fluency at Time 1 plus the latent change: $\eta_2 = \eta_1 + \Delta\eta$. Note that this specification does not impose restrictions on the CFA model. Consequently, defining an LDS as such does not influence the goodness of fit of a model. For model identification, the effects of η_1 on η_2 and of $\Delta\eta$ on η_2 were fixed to 1.0. The change that occurred between the two time points, $\Delta\eta$, may be related (negatively or positively) to the initial status. The regression of change on initial status was included in the model as a parameter to be estimated. Other parameters were the mean and the variance of the initial status and of the latent change. Measurement invariance (Little, 2013) across the two tracks and the two time points was tested. Measuring change presupposes that the same constructs are measured across time points.

The models were fitted with Mplus 7 (Muthén & Muthén, 1998–2012), using maximum likelihood estimation. The goodness of fit of models was evaluated by the chi-square statistic, the Root Mean Square Error of Approximation (RMSEA) and the Comparative Fit Index (CFI). Due to our large N, the chi-square goodness-of-fit test was highly sensitive to small deviations. Therefore, we also used other fit indices that were less sensitive to sample size. Conventional rules of thumb for RMSEA are that values below 0.08 present an acceptable fit and values below 0.05 a good fit (e.g., Little, 2013, p. 109). CFIs above 0.95 indicate a good fit (Little, 2013, p. 115).

3. Results

3.1. Descriptive statistics

Table 1 shows the means and standard deviations of all variables, separately per educational track and per time of measurement. For both tracks, mean decoding fluency was higher at the second than at the first time point, indicating that growth of decoding fluency took place in all variables; mean differences across time were approximately the same for both tracks. There was one exception, however: we found a higher average growth for French WDF in the high than in the low educational track. For all fluency variables, means were higher for the high than for the low track. The standard deviations were rather similar across tracks and across time for each variable. Only few students reached the end of a test card; even at Time 2 and in Track 2 there was hardly any ceiling performance in the data.

The accuracy measures show that very few errors were made in the Dutch word decoding and pseudoword decoding tasks by students at both levels. Most students obtained a close to perfect score, as is commonly the case for native speakers of Dutch. Accuracy was lower for reading di- and multisyllabic than for monosyllabic pseudowords. Likewise, more errors were made reading English and especially French words. For French words, the proportions correct varied even from 0.07 to 0.98. The accuracy of reading French words increased considerably during the school year, especially in the highest educational track. Except for English and French, the distributions of proportion correct were extremely skewed to the left. In the remainder of the article, only the fluency measures were analyzed.

Correlations among the Dutch reading measures were generally high, all above 0.63, with small track differences. The autocorrelations, which may be seen as test-retest reliabilities, even ranged from 0.83 to 0.92. The intercorrelations of the foreign-language decoding measures for the two educational tracks are shown in Table 2. For English WDF the stability or pretest-posttest reliability was 0.68 in the lower track and 0.83 in the higher track. These values were even lower for French WDF, only 0.58 in the lower track and 0.67 in the higher track. For the students in the sample, the development of French clearly lags behind the development of English, because learning French only just started for these students. Understandably, across-time and across-language correlations were lower than the stabilities. Except for the stabilities, track differences between correlations were not large, but the correlations in Track 2 tended to be higher than the correlations in Track 1. Nevertheless, correlations of English and French WDF with Dutch WDF tended to be lower in Track 2 than in Track 1 (not in the table[1]). For students in Track 1, Dutch word decoding tasks correlated more strongly with English (0.44 to 0.67) than with French (0.36 to 0.50) WDF. In Track 2, no such difference between English and French was observed. The correlations of English and French WDF with the Dutch pseudoword decoding tasks were roughly of the same order of magnitude as the correlations with Dutch real-word decoding tasks (0.39 to 0.60 for English and 0.37 to 0.53 for French in Track 1, and about the same in Track 2, namely 0.29 to 0.52 for English and 0.38 to 0.53 for French).

3.2. Measurement model and LDS model for Dutch word decoding fluency

The structure of the measurement model is shown in Fig. 1. This model was developed and tested in Van de Ven et al. (2017) and tested again on the present sub-sample of the original data, using multiple-group CFA. The model included one factor for *Dutch WDF* (abbreviated as W1 and W2 in Fig. 1) and one for *Dutch PWDF* (P1 and P2) for the eight Dutch decoding tasks at both the beginning and the end of the school year. Equality of factor loadings across time and across tracks was established. The model included a large number of error covariances (see Van de Ven et al., 2017), equal across time and across tracks. Likewise, the residual variances of the eight decoding tests were restricted to be equal both across time and across groups. Finally, the intercepts of the decoding tests were fixed to be equal across time and across tracks. For the polysyllabic words and pseudowords, however, track differences in the intercepts were allowed. In addition, in the lowest educational track, we had to allow differences across time for C(C)VC(C) words and for CVC pseudowords. The estimated intercepts in the lowest track pointed to a relatively larger change for CVC pseudowords in comparison with the other types of pseudowords, and

[1] Full correlation matrices are available upon request from the first author.

Table 1
Means and standard deviations (between parentheses) of word decoding fluency and accuracy in Dutch, English and French, tabulated by educational track and time of measurement.

	Time	Track 1		Track 2		Total	
		Fluency	Accuracy	Fluency	Accuracy	Fluency	Accuracy
Dutch words							
CVC	1	101.16 (16.739)	0.982 (0.024)	110.31 (14.773)	0.991 (0.014)	104.98 (16.567)	0.986 (0.021)
	2	105.61 (17.233)	0.985 (0.018)	114.63 (15.390)	0.992 (0.012)	109.33 (17.076)	0.988 (0.016)
C(C)VC(C)	1	90.99 (18.067)	0.979 (0.028)	101.34 (15.939)	0.991 (0.015)	95.31 (17.943)	0.984 (0.024)
	2	94.66 (18.872)	0.981 (0.029)	105.62 (15.722)	0.993 (0.012)	99.18 (17.076)	0.986 (0.024)
Disyllabic	1	76.21 (18.002)	0.975 (0.031)	90.03 (14.498)	0.989 (0.016)	81.98 (17.964)	0.981 (0.027)
	2	81.69 (17.984)	0.981 (0.026)	94.95 (14.163)	0.993 (0.011)	87.17 (17.749)	0.986 (0.022)
Tri- and quadrisyllabic	1	57.83 (14.942)	0.957 (0.045)	71.67 (13.267)	0.980 (0.025)	63.60 (15.810)	0.967 (0.040)
	2	62.84 (15.406)	0.968 (0.037)	76.39 (13.353)	0.987 (0.018)	68.43 (16.039)	0.976 (0.032)
Dutch pseudowords							
CVC	1	76.43 (17.371)	0.952 (0.053)	86.18 (16.725)	0.969 (0.033)	80.49 (17.758)	0.959 (0.047)
	2	82.37 (17.861)	0.956 (0.050)	91.69 (16.759)	0.976 (0.030)	86.21 (18.001)	0.965 (0.044)
C(C)VC(C)	1	61.01 (17.137)	0.932 (0.066)	72.09 (17.076)	0.960 (0.047)	65.63 (17.952)	0.944 (0.061)
	2	65.43 (17.834)	0.942 (0.060)	76.86 (17.402)	0.971 (0.037)	70.14 (18.522)	0.954 (0.053)
Disyllabic	1	39.82 (13.238)	0.842 (0.119)	49.13 (13.854)	0.906 (0.083)	43.70 (14.250)	0.869 (0.110)
	2	43.27 (13.710)	0.863 (0.107)	53.19 (14.222)	0.926 (0.071)	47.36 (14.747)	0.889 (0.098)
Tri- and quadrisyllabic	1	30.52 (10.657)	0.799 (0.134)	37.93 (10.946)	0.861 (0.106)	33.61 (11.373)	0.825 (0.127)
	2	33.45 (11.019)	0.815 (0.125)	41.54 (11.160)	0.891 (0.087)	36.79 (11.765)	0.846 (0.117)
English words	1	48.40 (9.017)	0.775 (0.083)	56.95 (8.443)	0.825 (0.084)	51.97 (9.739)	0.796 (0.087)
	2	52.59 (7.811)	0.797 (0.079)	62.01 (8.954)	0.859 (0.075)	56.48 (9.507)	0.823 (0.083)
French words	1	24.24 (11.190)	0.469 (0.178)	31.89 (11.733)	0.574 (0.173)	28.13 (12.082)	0.523 (0.183)
	2	29.46 (14.050)	0.534 (0.205)	42.73 (13.255)	0.708 (0.154)	34.96 (15.197)	0.606 (0.204)
N[a]	1	449		321		770	
	2	452		317		769	

[a] For French words at Time 1 the N was 294 in Track 1 and 305 in Track 2.

Table 2
Intercorrelations of English and French word decoding fluency (WDF) at two time points separately for Track 1 (above diagonal) and Track 2 (below diagonal).

	English WDF 1	English WDF 2	French WDF 1	French WDF 2
English WDF1		0.68	0.56	0.39
English WDF 2	0.83		0.36	0.52
French WDF 1	0.54	0.41		0.58
French WDF 2	0.58	0.60	0.67	

to a relatively smaller change for C(C)VC(C) words than for the other types of words. Factor means and factor variances and covariances were left free to be estimated, except the means in the lowest track at Time 1, in order to fix the scales of the latent variables. Next to the two factors of single-word decoding fluency, the latent differences between these factors at the two time points were in the model (ΔW and ΔP), see Van de Ven et al. (2017).

The fit statistics for Model 1 were satisfactory, χ^2(df = 214, N = 787) = 396.92, RMSEA = 0.047 (90% C.I. 0.039–0.054), CFI = 0.991. The fitted model implies metric invariance, both across time and across educational tracks. Table 3 shows the standardized factor loadings for Model 1. The unstandardized loadings were equal for all educational tracks and at both time points. Nevertheless, standardized loadings differed slightly by educational track and by time. All standardized loadings appeared consistently high, between 0.83 and 0.95. A strong common factor was identified for both word and pseudoword decoding fluency. For comparing means, only partial scalar invariance was obtained (cf. Van de Ven et al., 2017).

The two factors were very highly correlated at Time 1, namely 0.93 in Track 1 and 0.91 in Track 2. The correlation between the two latent change factors was also high but much lower, namely 0.63 in Track 1 and 0.69 in Track 2. The regression of the two latent change factors on initial level (the two time-1 factors) was fixed to be the same for the two educational tracks. Change was negatively but not significantly related to the initial level of word or pseudoword decoding fluency. The intercepts of the latent change factors were estimated as 4.77 and 2.83 for word, respectively pseudoword decoding fluency in Track 1 and 4.44, respectively 3.18 in Track 2. These intercepts represent the mean changes in decoding fluency during one school year; all were statistically significant ($p < .001$). Thus, there was evidence of growth in WDF and PWDF in both tracks. The track differences in mean change were not statistically significant, $p = .41$ for the latent change in WDF and $p = .22$ for the latent change in PWDF. Thus, growth in both WDF and PWDF may be considered of about the same magnitude in both tracks. The two tracks, however, did differ significantly in initial mean level ($p < .001$); students in Track 2 on average scored 8.54 higher on WDF1 and 6.78 higher on PWDF1.[2]

3.3. Associations between Dutch, English and French decoding fluency

To study the associations of WDF in foreign languages with WDF and PWDF in Dutch, the students' mother tongue, a model was constructed to predict English and French WDF2 from level of and change in Dutch WDF and PWDF, including English and French WDF1 as covariates. The model included the measurement model and the LDS model of Fig. 1 to which the word decoding fluency tests for English and French at the two time points were added, see Fig. 2. Within the context of this model, we studied the associations of the level of foreign-language word decoding fluency at Time 1 with the level of word and pseudoword decoding fluency, also assessed at Time 1 (first research question). The model also allows us to study *change* in English and French WDF during the school year, by testing the effects of Dutch WDF1 and PWDF1 on English and French WDF2 while controlling for the pretest measures of these variables (second research question). We started with a model in which all regression relationships were the same across the two educational levels. The fit statistics for this Model 2 were χ^2(df = 334, N = 787) = 671.21, RMSEA = 0.051 (90% C.I. 0.045–0.056), CFI = 0.985. Subsequently, we tested whether the effects of Dutch WDF1 and PWDF1 on English and French WDF1 could be considered the same for the two tracks. The chi-square difference test

[2] Note that this analysis of mean latent differences may be biased to some (small) extent, because scalar invariance across time and across tracks was not fully satisfied.

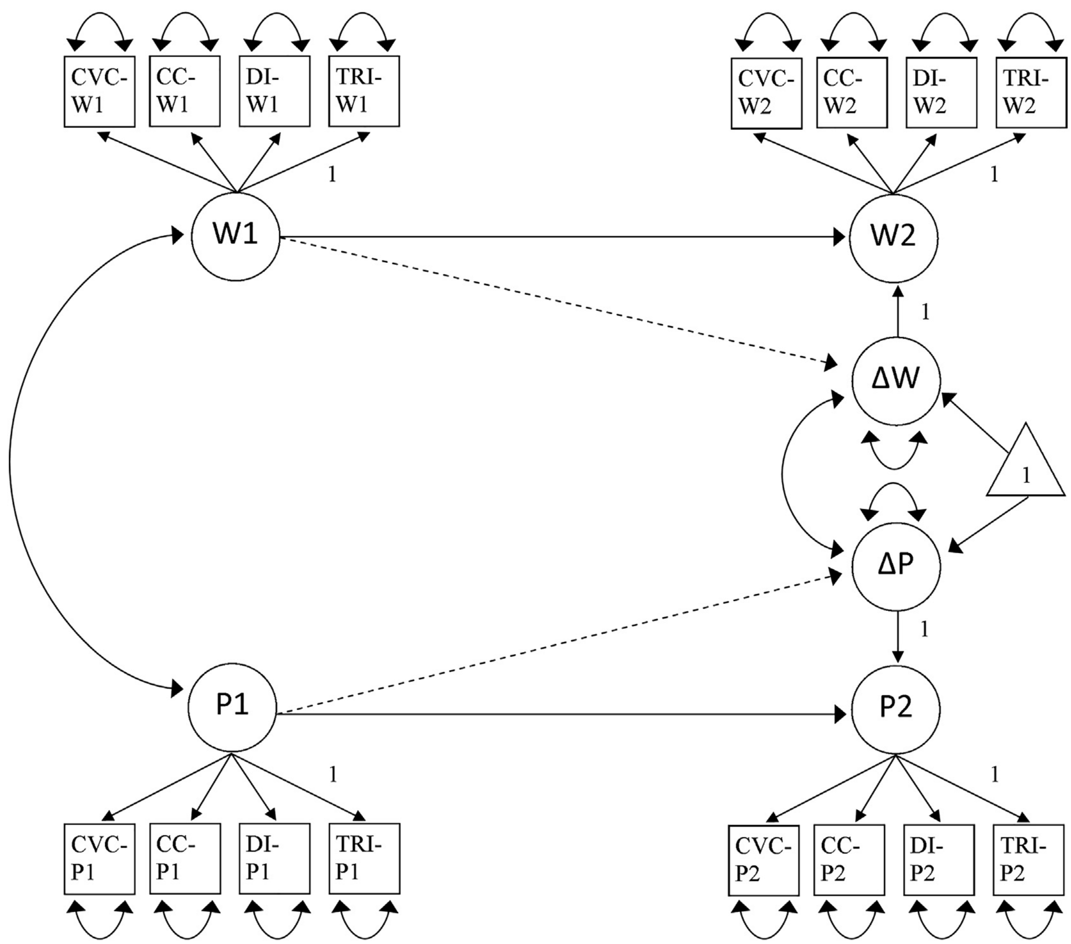

Fig. 1. Measurement model for Dutch word and pseudoword decoding fluency at two time points (Model 1). W1 and P1 refer to word and pseudoword decoding at Time 1. ΔW and ΔP refer to the latent change in word and pseudoword decoding fluency during one school year. W2 = W1 + ΔW and P2 = P1 + ΔP. Residual correlations between test cards omitted from the figure.

Table 3
Standardized factor loadings for words and pseudowords (Model 1).

Cards	Word reading	Pseudoword reading
CVC words	0.84	
C(C)VC(C)words	0.90	
Disyllabic words	0.92	
Tri- and quadrisyllabic words	0.89	
CVC pseudowords		0.86
C(C)VC(C)pseudowords		0.95
Disyllabic pseudowords		0.95
Tri- and quadrisyllabic pseudowords		0.92

Note: The standardized loadings differed slightly by educational track and by time of measurement; the largest difference was only 0.04.

gave a nonsignificant result, $d\chi^2(4) = 6.26$, $p = .18$. Hence, the effects of Dutch word decoding fluency on English and French word decoding fluency in the beginning of the school year appeared to be the same for both educational tracks, notwithstanding the mean differences between the tracks.

We then also tested group differences in the regressions of English and French WDF2 on the same variables at Time 1. These regressions appeared significantly different for the two tracks, $d\chi^2(4) = 15.31$, $p = .004$. Upon further exploration, this interaction effect seemed to be associated with differing effects of English WDF1, $d\chi^2(2) = 10.66$, $p = .005$, and not of French WDF1. The regression of English WDF2 on WDF1 showed a somewhat larger coefficient for high-track students than for low-track students, even though for both groups stability of English WDF was high. Likewise, the effect of English WDF1 on French WDF2 differed across tracks. In Track 1, there was no significant effect of English on French, but there was a significant positive effect in Track 2. Furthermore, there was no evidence of group differences in regressions of English or French WDF on Dutch WDF or PWDF, $d\chi^2(8) = 1.67$, $p = .99$. Hence, the final model (Model 3) includes two interactions with educational track, namely for the autoregression of English WDF and for the effect of English WDF1 on French WDF2. The fit statistics for Model 3 were $\chi^2(\text{df} = 332, N = 787) = 657.14$, RMSEA = 0.050 (90% C.I. 0.044–0.055), CFI = 0.986. The parameter estimates for Model 3 are shown in Tables 4a and 4b. These tables show the estimated unstandardized regression coefficients from the final model for predicting English and French WDF at the beginning (Table 4a) and the end (Table 4b) of the school year. These estimates were used to answer research questions 1 and 2, respectively.

Table 4a shows the effects of Dutch WDF1 on the level of English and French WDF1. There was no significant effect of Dutch WDF on French WDF, but there was a statistically significant positive effect of Dutch PWDF on French WDF. The level of English WDF at Time 1 appeared to be related to Dutch WDF and PWDF. Despite their very strong intercorrelation, both Dutch WDF and PWDF contributed to the prediction of the level of English WDF. There was a stronger effect of Dutch PWDF1 on French WDF1 than on English WDF1, Wald $\chi^2(1) = 8.12$, $p = .004$, which agrees with our expectations. Also, Dutch PWDF1 seemed a better predictor of French WDF1 than was Dutch WDF1, Wald $\chi^2(1) = 3.93$, $p = .047$. French and English WDF1 were positively correlated, 0.56 in Track 1 and 0.54 in Track 2 (see Table 2). The first research question may also be answered in a different model by using

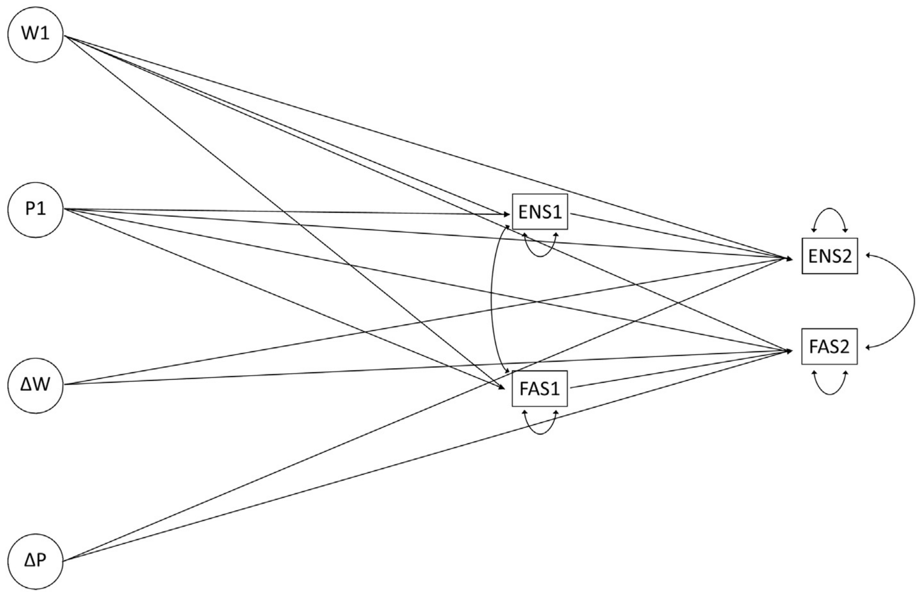

Fig. 2. Structural part of a model for predicting foreign-language word decoding fluency (ENS for English and FAS for French) at two time points from Dutch word (W) and pseudoword decoding (P) fluency. The four latent variables for Dutch are defined in the measurement model depicted in Fig. 1.

Table 4a
Regressions of English and French word decoding fluency (WDF) at Time 1 on Dutch word and pseudoword decoding fluency: Unstandardized regressions coefficients and their standard errors (between parentheses) Estimated from the final Model 3 (regressions the same in both tracks).

	English WDF1	French WDF1
Dutch WDF1	**0.256** (0.054)	0.072 (0.089)
Dutch PWDF1	**0.145** (0.073)	**0.475** (0.117)
R^2 in Track 1	**0.409**	**0.255**
R^2 in Track 2	**0.261**	**0.214**

Note. Statistically significant coefficients ($p < .05$) are in bold face type.

Table 4b
Regressions of English and French word decoding fluency (WDF) at Time 2 on Dutch word and pseudoword decoding fluency at Time 1 and Time 2, after controlling for English and French word decoding fluency at Time 1: Unstandardized regression coefficients and their standard errors (between parentheses) estimated from the final Model 3.

	English WDF2	French WDF2
English WDF 1 (in Track 1)	**0.693** (0.033)	0.097 (0.080)
English WDF1 (in Track 2)	**0.834** (0.035)	**0.325** (0.066)
French WDF1	**−0.078** (0.019)	**0.542** (0.041)
Dutch WDF1	**0.097** (0.038)	0.117 (0.084)
Dutch PWDF1	0.066 (0.051)	**0.212** (0.108)
Latent change in Dutch WDF	0.038 (0.076)	0.014 (0.172)
Latent change in Dutch PWDF	**0.283** (0.108)	**0.732** (0.249)
R^2 in Track 1	**0.708**	**0.445**
R^2 in Track 2	**0.727**	**0.610**

Note. Statistically significant coefficients ($p < .05$) are in bold face type.
Note. Regressions were the same in both educational tracks, except for the regressions of English and French WDF2 on English WDF1.

English and French at Time 2 as criterion variables, omitting English and French at Time 1. Results were highly similar.

Table 4b holds the parameter estimates of Model 3 for the prediction of English and French WDF at the end of the school year, using English and French WDF at the beginning of the school year as covariates to answer the second research question (see Fig. 2). English and French WDF2 depended in the same way on earlier Dutch WDF and PWDF in both tracks. As noted before, track differences, however, came to light for the effects of English WDF1. As expected, there were strong autoregressive effects, especially for English WDF in both educational tracks, but strongest in the highest track ($b = 0.693$ in Track 1 versus $b = 0.834$ in Track 2). French WDF showed less stability ($b = 0.542$) than English WDF, in both tracks. The students had less experience in reading French than in reading English, which may explain the lower stability of development in French than in English. These autoregressive coefficients imply that change in English or French WDF was negatively related with initial level. For instance, for French the regression coefficient of initial level on change by implication equals $0.542–1.0 = -0.458$. Students scoring high at the beginning of the school year gained less during the year than initially lower scoring students did.

We expected an effect of English on (change in) French WDF (Research Question 3), since the acquisition of English preceded that of French (and possibly reflected experience with acquiring a non-native language). This was confirmed only for the highest track ($b = 0.325$), and not for the lowest. We did not expect an effect of French on English. Hence, the statistically significant negative coefficient for French in the equation for English ($b = -0.078$) came as a surprise, but the coefficient was rather low.[3] The same two effects appeared statistically significant, albeit weaker, in Model 3, after controlling for the English and French measures at the beginning of the school year. Likewise, statistically significant regression coefficients appeared of the latent change in Dutch PWDF on the change in both English ($b = 0.283$) and French WDF ($b = 0.732$). Stronger growth in Dutch pseudoword decoding fluency during the school year was related to higher gain in English and especially French word decoding fluency. The effect of Dutch PWDF on change in foreign-language WDF was larger on French than on English, but this difference was not statistically significant, Wald $\chi^2(1) = 1.98$, $p = .16$. Similarly, the effect of latent change in Dutch PWDF appeared stronger for French than for English, but this difference was not statistically significant either, Wald $\chi^2(1) = 3.49$, $p = .062$. The residuals of both foreign-language word decoding measures were correlated, both at Time 1 ($r = 0.32$ for Track 1 and $r = 0.41$ for Track 2) and Time 2 ($r = 0.29$ for Track 1 and $r = 0.27$ for Track 2).

In general, we found stronger effects of Dutch pseudoword decoding fluency on French than on English word decoding fluency. Further, the other way around, English word decoding fluency appeared more strongly related with decoding fluency of Dutch real words and not or

[3] The negative sign of the partial regression coefficient while all correlations were positive, resulted from the relatively large differences between correlations (see Table 2). Of course, English Word Decoding Fluency at Time 2 correlated much higher with the same variable at Time 1 than with French Word Decoding Fluency at Time 1. Moreover, the two foreign-language measures at Time 1 showed a substantial intercorrelation.

Table 5
Standardized regression coefficients (*β*) and correlations of the predictors with the criterion variables (English and French WDF at two time points), estimated from Model 3.

	β				*r*			
	Track 1		Track 2		Track 1		Track 2	
	English	French	English	French	English	French	English	French
Time 1								
Dutch WDF	**0.419**[a]	0.073	**0.316**	0.058	0.636	0.487	0.500	0.431
Dutch PWDF	**0.214**	**0.451**	**0.192**	**0.426**	0.617	0.510	0.483	0.462
Time 2								
English WDF1	**0.704**	0.054	**0.791**	**0.206**	0.808	0.453	0.831	0.586
French WDF1	**−0.112**	**0.432**	**−0.101**	**0.474**	0.391	0.595	0.413	0.697
Dutch WDF1	**0.178**	0.119	**0.125**	0.101	0.649	0.506	0.539	0.545
Dutch PWDF1	0.086	**0.151**	0.073	**0.155**	0.624	0.512	0.514	0.556
ΔDutch WDF[b]	0.021	0.004	0.017	0.004	0.000	0.034	0.011	0.040
ΔDutch PWDF	**0.113**	**0.161**	**0.084**	**0.145**	0.125	0.158	0.075	0.118

[a] Statistically significant regression coefficients ($p < .05$) are in boldface type. All correlations were statistically significant except the correlations with the latent difference score of Dutch WDF.
[b] Δ refers to the latent difference score.

less strongly with reading Dutch pseudowords. These differences were according to expectations. Dutch real word and pseudoword reading, however, were very highly correlated. Because of this multicollinearity problem, we present, in Table 5, zero-order correlations of the predictors with the criterion (cf. Courville & Thompson, 2001) next to the standardized regression coefficients from Model 3, separately for the two tracks. As noted above, most regression coefficients were the same in the two tracks. Nevertheless, standardized regression coefficients can differ between tracks because of differing variances; the same is true for correlations. The correlations have been calculated[4] on the basis of the estimated Model 3 (Tables 4a and 4b). The correlations show that Dutch real word and pseudoword decoding fluency were correlated to about the same extent with English and French word decoding fluency, despite the differing regression relationships. This results from their high intercorrelation. The correlations confirm that both level of and change in foreign-language word decoding fluency are related with Dutch word and pseudoword decoding fluency. Moreover, change in foreign-language decoding fluency during the first year of secondary education relates positively with change in Dutch pseudoword decoding fluency but not with change in Dutch real-word decoding fluency, as measured with the present set of test cards. This largely confirms the conclusions from the regression results. However, the differential effects of Dutch PWDF compared to Dutch WDF on French WDF were hardly visible in the correlations.

Finally, to answer Research Question 4, we looked at the intercepts and the track differences in intercepts (adjusted track means), estimated in Model 3. Note that the interactions with educational track (see Table 4b) implied that track differences in intercepts differed with the students' scores on English WDF1. Using grand-mean centering, we evaluated intercepts and their differences across tracks in the center of the distribution of English WDF1. For students scoring below (above) average, the intercepts were lower (higher) and the mean track differences were also lower (higher). These intercepts represented the mean changes corrected for initial achievement and for the regression on Dutch WDF1 and PWDF1 (for English 51.42 in Track 1 and 59.66 in Track 2; for French 28.45 in Track 1 and 37.98 in Track 2). These adjusted mean changes were statistically significant ($p < .001$), indicating that the scores on English WDF improved on average for both tracks. For average scoring students, this difference between the tracks was clearly statistically significant, Wald $\chi^2(1) = 490.71$, $p < .001$ for English and Wald $\chi^2(1) = 144.98$, $p < .001$ for French. Thus, the two tracks not only differed in their mean scores Time 1 but also in their average change during the school year. At the center of the distribution, students in Track 2 on average improved more than those in Track 1. This track difference became smaller, however, for students starting at a lower level at the beginning of the school year.

4. Discussion

The current study investigated the relation between the development of L1, L2, and L3 word decoding skills in the first year of Dutch secondary education, using multiple-group structural equation modeling, distinguishing two educational tracks. In the Netherlands, students receive an introductory English course during grades 5 and 6, and an intensive course from grade 7 onwards (i.e., the onset of secondary education). Children receive an introductory French course from the beginning of secondary education. Further, while French is mainly learned at school, English is also learned outside the school setting (through media exposure).

We compared the Dutch (pseudo)word decoding skills, and the English and French word decoding skills of 459 students in lower and intermediate level and 328 in higher level first-year secondary students at two points in time (December and June). Unsurprisingly, means for all fluency variables were higher in the high than in the low educational track. For both tracks, growth of decoding fluency was evidenced on all decoding measures. Only for French decoding, we found a higher average growth in the higher as compared to the lower educational track.

To answer the first research question, we investigated the relation between the level of English and French word decoding fluency on the one hand and Dutch word and pseudoword decoding fluency on the other (Table 4a), all measured at the beginning of the school year. We found that Dutch word and pseudoword decoding fluency can predict level of English word decoding fluency. For French, however, we found a positive relation between level of Dutch pseudoword decoding fluency, but not of word decoding fluency. This difference between English and French is consistent with the idea that the non-lexical route plays a predominant role during early reading development, since French word reading is in an earlier stage of development.

The second research question concerned the extent to which *change* in English and French word decoding fluency was associated with level of and change in Dutch word and pseudoword decoding fluency (between Time 1 and Time 2; Table 4b). We distinguished between level of and change in Dutch word and pseudoword decoding fluency by means of latent change modeling. Our results showed that level of Dutch word, but not pseudoword, decoding fluency can predict change in English word decoding fluency. Change in pseudoword decoding fluency, in

[4] Estimated correlation matrix of the latent variables presented in Technical Output 4 of Mplus.

contrast, only predicted change in (and not level of) English word decoding fluency. Hence, there is a connection between development of Dutch pseudoword reading ability and the development of English word decoding fluency. When focusing on change in French word decoding fluency, we again found an association with the level of Dutch pseudoword decoding fluency, but not with the level of Dutch word decoding fluency. Similar to English, change in French word decoding fluency was associated with change in Dutch pseudoword, but not word decoding fluency. Hence, there is also a connection between development of Dutch pseudoword and French word decoding fluency.

These results suggest that, compared to French, the development of English word decoding skills more strongly depends on lexical retrieval skills, which can be explained by the Dual Route Cascaded model (Coltheart et al., 2001). Dutch first-year secondary students are beginning learners of French, and therefore predominantly use the nonlexical route during French word decoding. In comparison, students are more proficient in English than in French due to more exposure, both through the media as well as at school. Hence, in addition to the nonlexical route, Dutch students also strongly relied on the lexical route during English word decoding. Furthermore, the positive relation of L1 decoding fluency with the level and development of both English and French word decoding skills is in line with the linguistic coding differences hypothesis (Sparks & Ganschow, 1991) and the linguistic interdependence model (Cummins, 1979). In addition, the relation between L1 and L2 learning corroborates previous findings (mostly based on concurrent data) on the relation between L1 and L2 decoding ability (e.g., Geva et al., 1997; Lindsey et al., 2003; Melby-Lervåg & Lervåg, 2011; Sparks et al., 2008).

Interestingly, compared to L2 word decoding, L1 word decoding appeared to be more predictive of L3 word decoding development, despite the fact that French may phonotactically and orthographically (Seymour et al., 2003), and lexically (Gray & Atkinson, 2003) be considered more similar to English than to Dutch. If we assume that the magnitude of transfer in foreign language development depends primarily on linguistic distance, as proposed in previous research (e.g., Koda, 2005, 2008; Rothman, 2010), it may be hypothesized, based on our findings, that if there is lexical similarity with suprasegmental disparity between the L2 and the L3, which holds for L2 English and L3 French, this creates (temporary) confusion in beginning learners. As a result, L2 English proficiency may have become a weaker predictor of French word decoding development (relative to L1 Dutch proficiency). In addition, this hypothesis could serve as an alternative explanation for the negative relation between French word decoding and English decoding development, discussed above. Further research is required to experimentally test this hypothesis.

The third research question addressed the association of English (L2) with French (L3) word decoding fluency. We found a positive relation of English (L2) word decoding ability with the development of French (L3) word decoding ability, but restricted to the higher educational track. Various explanations may be found in the literature to account for a positive relation between L2 and L3 development. For one thing, foreign (i.e., English) word decoding ability may reflect language-learning aptitude, and language-learning aptitude may predict ease of language learning (e.g., Dörnyei, 2005). In addition, previous research suggests that students may benefit from their experience with the acquisition of a foreign language (i.e., English) when they start learning French, because prior language-learning experience facilitates metalinguistic awareness (for a review, see Cenoz, 2003). Both factors may have contributed to our result. Our finding that this relationship between L2 and L3 word decoding is restricted to the higher educational track may be explained by the threshold hypothesis by Cummins (1979), which claims that benefit from cross-linguistic transfer only occurs when learners attain a particular level of proficiency in their L1 or L2. Moreover, if we assume that languages in beginning learners' multilingual mental lexicon are strongly interconnected (as suggested by e.g., De Bot, 2004; Sánchez, 2014; Wei, 2006), L2-induced development of L3 reading skills may only occur in more proficient learners, with stronger language control abilities.

Finally, we found a negative relation between French (L3) word decoding ability and the development of English (L2) word decoding skills, after controlling for English decoding fluency at Time 1, which suggests that the acquisition of French may (temporarily) interfere with the continued acquisition of English in the first year of secondary education. We speculate that since the L2 and L3 in beginning L3 learners with low L2 proficiency are strongly intertwined, as suggested in previous research (e.g., De Bot, 2004; Sánchez, 2014; Wei, 2006), this might cause temporary problems with language control, for example delayed inhibition or spurious cross-linguistic spreading of activation. However, the negative relation may also have resulted from a multicollinearity problem, namely the correlations between the predictors at Time 1 (see Table 2).

The present study has several limitations that warrant discussion. First of all, this study only comprised two points in time, which limits the possibilities to identify causal relations between Dutch decoding development and foreign language decoding development. For instance, our regression model for English and French at Time 1 (Table 4a) holds predictors and criterion variables from the same measurement occasion. This model is fully equivalent with a model that holds only the intercorrelations of all Time-1 variables, and thus gives exactly the same fit with the data. On the other hand, especially in the early phases of learning a foreign language, it is more plausible to assume an effect of native-language decoding fluency on foreign-language decoding fluency than the other way around. Including the regression of English and French word decoding fluency on Dutch (pseudo)word decoding fluency allowed to compare the native-language effects at two points in time. This revealed that these native-language effects were different for French, which started to develop later, than for English.

A second limitation was the substantial number of missing values for French WDF during the first test session in the low educational track, because students who were not sufficiently competent to perform the French word decoding test by that time did not obtain a score. Our method to deal with the missing values was full information maximum likelihood, which assumes that data are missing at random (MAR). These data, however, were probably *not* missing at random, since the students concerned all belonged to the lower educational track and should probably have scores in the lower tail of the distribution of French word decoding. When data fail to satisfy the MAR assumption, the method will result in biased parameter estimates, but "this bias tends to be isolated to a subset of the analysis model parameters" (Enders, 2010, p. 125). A method that adequately deals with the causes of missingness may be better. Such a model, however, will require untestable assumptions which, when violated, possibly result in serious biases (Enders, 2010, ch. 10).

A third limitation was the fact that our study only consisted of Dutch, English, and French decoding measures, and underlying language skills were not taken into account. Further research is required to establish which underlying processes are responsible for the connections established among L1, L2, and L3 decoding development.

Finally, there was substantial multicollinearity among the two latent variables of Dutch word and pseudoword decoding. Importantly, however, the results obtained were quite consistent across the different models, and we obtained effects of both word and pseudoword decoding on multiple occasions. Though the differential effects found of word and pseudoword decoding fluency were in agreement with our expectations, the correlations of the predictors with the criteria (Table 5) do not clearly show these differential effects. Because of the multicollinearity further research needs to indicate whether the findings concerning differential effects of word and pseudoword decoding fluency prove to be stable upon replication.

The findings presented in this study hold multiple implications for foreign language teaching. To begin with, our results imply that improving L1 (word) decoding fluency may enhance word decoding fluency in a non-native language. Moreover, our findings suggest that early gains in decoding fluency can be established primarily by practicing phonological decoding, in order to improve the non-lexical route, for example by means of flash card training (as suggested by e.g., Wentink, Van Bon, & Schreuder, 1997). More advanced language learners, in contrast, should aim to improve the non-lexical and lexical route in tandem, for instance by means of repeated reading (e.g., Taguchi, 1997).

In conclusion, this study shows that the development of word decoding skills in a foreign language depends on various factors. While beginning L2 learners especially rely on the non-lexical route, more advanced learners progress towards using the lexical route during word decoding, in line with the Dual Route Cascaded model (Coltheart et al., 2001). There is also a significant role of L2 decoding ability during L3 decoding development, which enhances the development of decoding skills for learners with prior foreign language learning experience.

Acknowledgements

We wish to express our gratitude to the children and the schools for participating in our study.

References

Bernhardt, E. (2005). Progress and procrastination in second language reading. *Annual Review of Applied Linguistics, 25*, 133–150. http://dx.doi.org/10.1017/s0267190505000073.

Bialystok, E. (2013). The impact of bilingualism on language and literacy development. In T. K. Bhatia, & W. C. Ritchie (Eds.). *The handbook of bilingualism and multilingualism* (pp. 624). New York: Wiley.

Braze, D., Katz, L., Magnuson, J. S., Mencl, W. E., Tabor, W., Van Dyke, J. A., ... Shankweiler, D. P. (2015). Vocabulary does not complicate the simple view of reading. *Reading and Writing, 29*(3), 435–451. http://dx.doi.org/10.1007/s11145-015-9608-6.

Cenoz, J. (2003). The additive effect of bilingualism on third language acquisition: A review. *International Journal of Bilingualism, 7*(1), 71–87. http://dx.doi.org/10.1177/13670069030070010501.

Cenoz, J., & Genesee, F. (1998). Psycholinguistic perspectives on multilingualism and multilingual education. In J. Cenoz, & F. Genesee (Eds.). *Beyond bilingualism: Multilingualism and multilingual education* (pp. 16–34). Clevedon: Multilingual matters.

Cheung, H. (1996). Nonword span as a unique predictor of second-language vocabulary learning. *Developmental Psychology, 32*(5), 867–873. http://dx.doi.org/10.1037/0012-1649.32.5.867.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*, 204–256. http://dx.doi.org/10.1037/0033-295x.108.1.204.

Commissaire, E., Duncan, L. G., & Casalis, S. (2011). Cross-language transfer of orthographic processing skills: A study of French children who learn English at school. *Journal of Research in Reading, 34*, 59–76. http://dx.doi.org/10.1111/j.1467-9817.2010.01473.

Courville, T., & Thompson, B. (2001). Use of structure coefficients in published multiple regression articles: β is not enough. *Educational and Psychological Measurement, 61*, 229–248. http://dx.doi.org/10.1177/00131640121971211.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49*(2), 222. http://dx.doi.org/10.2307/1169960.

De Bot, K. (2004). The multilingual lexicon: Modelling selection and control. *International Journal of Multilingualism, 1*(1), 17–32. http://dx.doi.org/10.1080/14790710408668176.

Delattre, P. (1963). Comparing the prosodic features of English, German, Spanish and French. *International Review of Applied Linguistics in Language Teaching, 1*(1), http://dx.doi.org/10.1515/iral.1963.1.1.193.

Domahs, U., Plag, I., & Carroll, R. (2014). Word stress assignment in German, English and Dutch: Quantity-sensitivity and extrametricality revisited. *The Journal of Comparative Germanic Linguistics, 17*(1), 59–96. http://dx.doi.org/10.1007/s10828-014-9063-9.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition.* Mahwah, NJ: Lawrence Erlbaum.

Durgunglu, A. Y., Nagy, W. E., & Hancin-Bhatt, B. J. (1993). Cross-language transfer of phonological awareness. *Journal of Educational Psychology, 85*, 453–465. http://dx.doi.org/10.1037/0022-0663.85.3.453.

Enders, C. K. (2010). *Applied missing data analysis.* New York: The Guilford Press.

Eurydice (2005). The education system in the Netherlands 2005. Dutch Eurydice Unit, Ministry of Education, Culture and Science of the Netherlandshttps://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/rapporten/2005/12/23/education-system-in-the-netherlands/eurydice-en.pdf, Accessed date: 21 September 2017.

Freed, E. M., Hamilton, S. T., & Long, D. L. (2017). Comprehension in proficient readers: The nature of individual variation. *Journal of Memory and Language, 97*, 135–153. http://dx.doi.org/10.1016/j.jml.2017.07.008.

Ganschow, L., Sparks, R., & Javorsky, J. (1998). Foreign language learning difficulties: An historical perspective. *Journal of Learning Disabilities, 31*, 248–258. http://dx.doi.org/10.1177/002221949803100304.

Geva, E., Wade-Woolley, L., & Shany, M. (1997). Development of reading efficiency in first and second language. *Scientific Studies of Reading, 1*(2), 119–144. http://dx.doi.org/10.1207/s1532799xssr0102_2.

Gray, R. D., & Atkinson, Q. D. (2003). Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature, 426*(6965), 435–439. http://dx.doi.org/10.1038/nature02029.

Herder, A., & De Bot, K. (2005). *Vroeg Vreemdetalenonderwijs in Internationaal Perspectief. Literatuurstudie in Opdracht van het Europees Platform.* Expertisecentrum taal, onderwijs en communicatie, University of Groningen.

Jared, D., & Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of Memory and Language, 44*(1), 2–31. http://dx.doi.org/10.1006/jmla.2000.2747.

Kleijnen, R., Steenbeek-Planting, E. G., & Verhoeven, L. (2008). *Toetsen en interventies bij dyslexie in het voortgezet onderwijs: Nederlands en de moderne vreemde talen [Tests and interventions for dyslexia in secondary education. Dutch and modern foreign languages].* Nijmegen: Expertisecentrum Nederlands.

Koda, K. (1996). L2 word recognition research: A critical review. *The Modern Language Journal, 80*(4), 450. http://dx.doi.org/10.2307/329725.

Koda, K. (2005). Learning to read across writing systems: Transfer, metalinguistic awareness, and second language reading development. In V. J. Cook (Ed.). *Second language writing systems* (pp. 311–334). Clevedon: Multilingual Matters (2005).

Koda, K. (2008). Impacts of prior literacy experience on learning to read in a second language. In K. Koda, & A. M. Zehler (Eds.). *Learning to read across languages: Cross-linguistic relationships in first- and second-language literacy development* (pp. 68–96). New York: Routledge.

Lindsey, K. A., Manis, F. R., & Bailey, C. E. (2003). Prediction of first-grade reading in Spanish-speaking English-language learners. *Journal of Educational Psychology, 95*, 482–494. http://dx.doi.org/10.1037/0022-0663.95.3.482.

Little, T. D. (2013). *Longitudinal structural equation modeling.* New York/London: The Guilford Press.

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60*(1), 577–605. http://dx.doi.org/10.1146/annurev.psych.60.110707.163612.

Melby-Lervåg, M., & Lervåg, A. (2011). Cross-linguistic transfer of oral language, decoding, phonological awareness, and reading comprehension: A meta-analysis of the correlational evidence. *Journal of Research in Reading, 34*(1), 114–135. http://dx.doi.org/10.1111/j.1467-9817.2010.01477.x.

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Protopapas, A., Mouzaki, A., Sideridis, G. D., Kotsolakou, A., & Simos, P. G. (2013). The role of vocabulary in the context of the simple view of reading. *Reading & Writing Quarterly, 29*(2), 168–202. http://dx.doi.org/10.1080/10573569.2013.758569.

Rauch, D. P., Naumann, J., & Jude, N. (2011). Metalinguistic awareness mediates effects of full biliteracy on third-language reading proficiency in Turkish–German bilinguals. *International Journal of Bilingualism, 16*(4), 402–418. http://dx.doi.org/10.1177/1367006911425819.

Rothman, J. (2010). On the typological economy of syntactic transfer: Word order and relative clause high/low attachment preference in L3 Brazilian Portuguese. *International Review of Applied Linguistics in Language Teaching, 48*(2–3), 245–273. http://dx.doi.org/10.1515/iral.2010.011.

Sánchez, L. (2014). L2 activation and blending in third language acquisition: Evidence of crosslinguistic influence from the L2 in a longitudinal study on the acquisition of L3 English. *Bilingualism: Language and Cognition, 18*(02), 252–269. http://dx.doi.org/10.1017/s1366728914000091.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*(2), 143–174. http://dx.doi.org/10.1348/000712603321661859.

Sparks, R. L., & Ganschow, L. (1991). Foreign language learning differences: Affective or native language aptitude differences? *The Modern Language Journal, 75*(1), 3–16. http://dx.doi.org/10.1111/j.1540-4781.1991.tb01076.x.

Sparks, R. L., Patton, J., Ganschow, L., Humbach, N., & Javorsky, J. (2008). Early first-language reading and spelling skills predict later second-language reading and spelling skills. *Journal of Educational Psychology, 100*(1), 162–174. http://dx.doi.org/10.1037/0022-0663.100.1.162.

Steenbeek-Planting, E. G., Kleijnen, R., & Verhoeven, L. (2008a). Woordleestoets Engels [English word reading test]. In R. Kleijnen, E. G. Steenbeek-Planting, & L. Verhoeven (Eds.). *Toetsen en interventies bij dyslexie in het voortgezet onderwijs: Nederlands en de moderne vreemde talen [Tests and interventions for dyslexia in secondary education. Dutch and modern foreign languages].* Nijmegen: Expertisecentrum Nederlands.

Steenbeek-Planting, E. G., Kleijnen, R., & Verhoeven, L. (2008b). Woordleestoets Frans [French word reading test]. In R. Kleijnen, E. G. Steenbeek-Planting, & L. Verhoeven (Eds.). *Toetsen en interventies bij dyslexie in het voortgezet onderwijs: Nederlands en de moderne vreemde talen [Tests and interventions for dyslexia in secondary education. Dutch and modern foreign languages].* Nijmegen: Expertisecentrum Nederlands.

Taguchi, E. (1997). The effects of repeated readings on the development of lower identification skills of FL readers. *Reading in a Foreign Language, 11*, 97–119.

Van Gelderen, A., Schoonen, R., De Glopper, K., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2004). Linguistic knowledge, processing speed, and metacognitive knowledge in first- and second-language reading comprehension: A componential analysis. *Journal of Educational Psychology, 96*(1), 19–30. http://dx.doi.org/10.1037/0022-0663.96.1.19.

Van Gelderen, A., Schoonen, R., de Glopper, K., Hulstijn, J., Snellings, P., Simis, A., &

Stevenson, M. (2003). Roles of linguistic knowledge, metacognitive knowledge and processing speed in L3, L2 and L1 reading comprehension. *International Journal of Bilingualism, 7*(1), 7–25. http://dx.doi.org/10.1177/13670069030070010201.

Van Gelderen, A., Schoonen, R., Stoel, R. D., de Glopper, K., & Hulstijn, J. (2007). Development of adolescent reading comprehension in language 1 and language 2: A longitudinal analysis of constituent components. *Journal of Educational Psychology, 99*(3), 477–491. http://dx.doi.org/10.1037/0022-0663.99.3.477.

Van de Ven, M., Voeten, M., Steenbeek-Planting, E. G., & Verhoeven, L. (2017). Post-primary reading fluency development: A latent change approach. *Learning and Individual Differences, 55*, 1–12. http://dx.doi.org/10.1016/j.lindif.2017.02.001.

Verhoeven, L. T. W. (2004). *Screeningstest voor Taal- en Leesproblemen [Screening test for language and reading problems].* Nijmegen/Arnhem, The Netherlands: Instituut voor Orthopedagogiek RU/Citogroep.

Wei, L. (2006). The multilingual mental lexicon and lemma transfer in third language learning. *International Journal of Multilingualism, 3*(2), 88–104. http://dx.doi.org/10.1080/14790710608668390.

Wentink, H. W., Van Bon, W. H., & Schreuder, R. (1997). Training of poor readers' phonological decoding skills: Evidence for syllable-bound processing. *Reading and Writing, 9*(3), 163–192. http://dx.doi.org/10.1023/A:1007921805360.