



Universiteit  
Leiden

The Netherlands

## **Cross-language distance influences receptive vocabulary outcomes of bilingual children**

Blom, E.; Boerma, T.; Bosma, E.; Cornips, L.; Heuij, K. van den; Timmermeister, M.

### **Citation**

Blom, E., Boerma, T., Bosma, E., Cornips, L., Heuij, K. van den, & Timmermeister, M. (2019). Cross-language distance influences receptive vocabulary outcomes of bilingual children. *First Language*, 40(2), 151-171. doi:10.1177/0142723719892794

Version: Publisher's Version

License: [Creative Commons CC BY-NC 4.0 license](#)

Downloaded from: <https://hdl.handle.net/1887/3200197>

**Note:** To cite this publication please use the final published version (if applicable).

# Cross-language distance influences receptive vocabulary outcomes of bilingual children

First Language

2020, Vol. 40(2) 151–171

© The Author(s) 2019



Article reuse guidelines:

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

DOI: 10.1177/0142723719892794

[journals.sagepub.com/home/fla](https://journals.sagepub.com/home/fla)**Elma Blom** 

Department of Special Education, Utrecht University, The Netherlands

**Tessel Boerma**

Department of Languages, Literature and Communication, Utrecht University, The Netherlands

**Evelyn Bosma**

Leiden University Centre for Linguistics, Leiden University, The Netherlands

**Leonie Cornips**

Faculty of Arts and Social Sciences, Maastricht University NL-Lab, Humanities Cluster (KNAW), The Netherlands

**Kirsten van den Heuij**

Rotterdam University of Applied Sciences, The Netherlands

**Mona Timmermeister**

Department of Special Education, Utrecht University, The Netherlands

## Abstract

Various studies have shown that bilingual children score lower than their monolingual peers on standardized receptive vocabulary tests. This study investigates if this effect is moderated by language distance. Dutch receptive vocabulary was tested with the Peabody Picture Vocabulary Test (PPVT). The impact of cross-language distance was examined by comparing bilingual groups with a small (Close;  $n = 165$ ) and a large between-language distance (Distant;  $n = 108$ ) with monolingual controls ( $n = 39$ ). As a group, the bilinguals scored lower on Dutch receptive vocabulary than the monolinguals. The bilingual Distant

---

## Corresponding author:

Elma Blom, Department of Special Education, Utrecht University, Heidelberglaan 1, 3584 CS Utrecht, The Netherlands.

Email: [W.B.T.Blom@uu.nl](mailto:W.B.T.Blom@uu.nl)

group had lower receptive vocabulary outcomes than the bilingual Close and monolingual groups. No difference emerged between the monolinguals and the bilingual Close group. It can be concluded that bilingual children whose languages provide ample opportunities for transfer and sharing knowledge do not have any receptive vocabulary delays. The findings underscore that bilingual children cannot be treated as a homogeneous group and are important for determining which bilingual children are at risk of low vocabulary outcomes.

### **Keywords**

Child bilingualism, cross-linguistic overlap, language distance, Peabody Picture Vocabulary Test, receptive vocabulary

## **Introduction**

Many bilingual children score on vocabulary tasks within the normal range of variation for monolingual children (Bialystok, 2001; Hammer et al., 2014; Pearson et al., 1993), but in between-group comparisons they often have smaller vocabularies than their monolingual peers (Bialystok et al., 2010; Marchman et al., 2010; Scheele et al., 2010; Thordardottir et al., 2006; Vagh et al., 2009). The question arises if all bilingual children have relatively small vocabularies in one language or, more likely, if there are factors that moderate the vocabulary differences between bilingual and monolingual children and which factors these are.

In this study, we revisit the effect of bilingualism on vocabulary development by investigating the role of cross-language distance. The specific questions we address are whether there are significant differences between the Dutch receptive vocabulary outcomes of bilingual 6- or 7-year-old children whose other language is closely related or more distant to Dutch, and whether or not both these bilingual groups are outperformed by monolingual Dutch controls. The study builds on and complements previous research on cross-linguistic influence in the English vocabularies of bilingual toddlers (Floccia et al., 2018) and 3- to 10-year-old children (Bialystok et al., 2010). The findings are important for determining which specific bilingual children are at risk of low vocabulary outcomes. Understanding these risk factors is especially relevant for the domain of vocabulary, given that limited vocabulary indicates a risk for literacy development (August & Shanahan, 2006; Dickinson & Tabors, 2001; Oller & Pearson, 2002; Snow et al., 1998). The study demonstrates that bilingual children cannot be treated as one homogeneous group (Dixon et al., 2012), and shows that linguistic distance is an important individual difference factor in bilingual vocabulary development. As such, it contributes to a better understanding of the multifaceted character of bilingualism (Grosjean & Li, 2012; Luk & Bialystok, 2013).

### *Cross-linguistic influence and the role of language distance*

There is a vast amount of research showing that bilingual children differentiate their two languages from an early age (Bhatia & Ritchie, 1999; De Houwer, 1995; Meisel, 2001),

but there is also much evidence of interaction between the two languages (Blom et al., 2017; Paradis & Genesee, 1996; Serratrice, 2013). For example, in the domain of vocabulary, it has been found that it takes bilingual children more time to name pictures when both languages are highly active, as illustrated in research showing longer response times in mixed compared to single language conditions (Jia et al., 2006; Kohnert et al., 1999). This example illustrates that interaction between bilingual children's two languages leads to interference, which, in turn, can cause naming delays in online performance. At the same time, there is evidence that cross-language interaction facilitates bilingual children's lexical retrieval if words in their two languages are cognates (Poarch & Van Hell, 2012; Sheng et al., 2016). Cognates are words in different languages that have a shared origin and resemble each other semantically and phonologically.

Other research investigating cross-linguistic influences in bilingual children's vocabulary development has used offline measures of vocabulary size. This research suggests that two types of cross-language knowledge transfer can affect children's vocabulary development positively: (1) abstract knowledge that is relatively independent of language-specific encoding, and (2) language-specific knowledge that is shared across the two languages. Transfer of conceptual knowledge is an example of the first type of information, as the same concepts are encoded differently across languages. Conceptual knowledge is of direct relevance to vocabulary learning, as learning vocabulary includes matching a concept (word meaning) with a phonological representation (word form). Knowing many concepts in one language may facilitate vocabulary learning in the other language, as it increases the likelihood that a child is familiar with the concepts expressed in the other language and allows understanding the meaning of an unknown word in that language (Cummins, 2000). Research has shown positive correlations between children's vocabulary sizes in the two languages in contexts where the two languages are typologically very distinct (Prevoe et al., 2015; Scheele et al., 2010), suggesting that transfer on a conceptual level takes place even if words in the two languages have hardly any phonological overlap.

Carrying over of phonological knowledge exemplifies the second type of transfer. Evidence for the transfer of language-specific phonological knowledge has been demonstrated in a recent study on Dutch-speaking children who learn English as a second language (L2) in early foreign language education (Goriot et al., 2018). The children who participated in this study performed better on items of the English version of the Peabody Picture Vocabulary Test (PPVT-III; Dunn & Dunn, 1997), which is a standardized measure of receptive vocabulary, if they are phonologically similar: the closer the English word and its Dutch translation are, the easier the English items are for Dutch-speaking children. In the lexical domain, bilingual children can thus benefit from cognates, which are frequent in closely related languages and infrequent if the distance between their two languages is large. Several studies on cognates in a range of language pairs have confirmed that bilingual children score better on cognate items than on non-cognates (Bosch & Ramon-Casas, 2014; Bosma et al., 2019; Kelley & Kohnert, 2012; Malabonga et al., 2008; Schelletter, 2002).

While there is robust evidence that cross-linguistic overlap explains differences between words and the rate at which they are learned, it is less clear whether cross-linguistic similarities also lead to vocabulary differences between children. Bialystok and

colleagues (2010) analyzed data from a large sample of children between ages 3 and 10, all learning English. In this sample, the bilingual children scored lower than monolinguals on the English PPVT, which is expected given bilingual children's distributed input (Marchman et al., 2010; Scheele et al., 2010; Thordardottir et al., 2006; Vagh et al., 2009). Based on a comparison of children whose non-English language was either East Asian or non-Asian, Bialystok et al. concluded that the observed difference between monolinguals and bilinguals did not change with language pair, suggesting that cross-linguistic overlap does not moderate the effect of bilingualism on vocabulary knowledge. However, a closer look at the composition of the two subsamples in the study by Bialystok et al. may explain why language pair did not have an effect: the East Asian languages (i.e., Cantonese, Japanese, Korean, Mandarin, Thai, and Shanghainese) in the first subsample have relatively few resemblances with English, as expected, but the same holds for many of the non-Asian languages in the second subsample (i.e., Amharic, Arabic, Croatian, Farsi, French, German, Greek, Gujarati, Hebrew, Hindi, Hungarian, Italian, Kannada, Macedonian, Marathi, Persian, Polish, Portuguese, Punjabi, Romanian, Russian, Serbian, Spanish, Tamil, Telugu, Turkish, Tagalog, and Urdu).

That the division in the study of Bialystok and colleagues is not optimally suited to investigate the role of language distance can be demonstrated with Levenshtein distances, which is a measure of language distance (Bakker et al., 2009; Wichmann et al., 2016; see below for further details). The average normalized Levenshtein distance for the non-Asian language group is 94.54 ( $SD = 6.71$ ), ranging between 68.62 for the distance between German and English to 102.20 for the distance between Tagalog and English. The average normalized Levenshtein distance for the East Asian language group is 99.66 ( $SD = 1.57$ ) ranging between 97.21 for the distance between (Khorat) Thai and English to 103.23 for the distance between (Suzhou Wu) Chinese, spoken in Shanghai, and English. Thus, the distances from English in the non-Asian and East Asian language groups overlap considerably. Consequently, the way in which language distance was operationalized by Bialystok and colleagues may not have allowed for detecting an effect of cross-language distance. If language pairs are more clearly distinct from each other in terms of distance from the target language, language distance may be found to influence the relation between bilingualism and vocabulary outcomes after all.

A second reason to revisit the impact of language distance on the vocabulary knowledge of bilingual children are the results of a recent study by Floccia and colleagues (2018), which shows effects of language distance on the vocabulary knowledge of bilingual toddlers. Floccia et al. compared the outcomes of the Communicative Development Inventory (CDI) of 372 2-year-old children learning British English, and one of 13 additional languages. The CDI is a widely used parent report instrument that provides information about children's vocabulary comprehension, production, gestures, and grammar. Linguistic distance was measured through phonological similarity (cognates), morphological complexity (ratio of morphemes to words), and word order typology (OV, VO, mixed OV/VO). It was found that a higher level of phonological overlap between languages was related to better vocabulary production scores, whereas higher degrees of morphological and word order overlap were related to better vocabulary comprehension. The relationships with structural language distance measures (morphology, syntax) suggest that carrying over word and sentence processing routines from one language to the

other facilitates young bilingual children's learning of vocabulary in the other language. These results confirm the role of language distance, and show, moreover, that language-specific knowledge that is transferred across languages is not limited to phonological knowledge, but also includes morphological and syntactic knowledge. It is unknown, however, whether language distance also impacts the vocabularies of older children.

### **Present study**

Whereas the study by Floccia and colleagues (2018) shows that language distance impacts on bilingual toddlers' vocabulary, the study by Bialystok and colleagues (2010) provides no support for effects of language distance on the vocabulary of older bilingual preschool and school-aged children. These contradictory results warrant further research, in particular research with bilingual child groups that differ clearly in language distance between the language pairs. The present study revisited the effect of language distance by investigating receptive vocabulary outcomes of bilingual and monolingual early school-aged children. The following research question guided our study: Do bilingual 6- or 7-year-old children have lower Dutch receptive vocabulary outcomes than monolingual age peers, and to what extent is this difference modulated by linguistic distance between bilingual children's languages?

Receptive vocabulary scores were collected with the Dutch version of the PPVT (PPVT-III-NL; Schlichting, 2005), the Dutch equivalent of the test used by Bialystok and colleagues. We assigned the bilingual children who participated to groups that are very clearly distinct in terms of linguistic distance from the target language (Dutch), expecting that this method would be better able to identify effects of language distance on receptive vocabulary development. Specifically, we compared the receptive vocabulary scores of monolingual Dutch children to those of bilingual children whose non-Dutch home language was a closely-related West-Germanic language (Frisian, Limburgish) or a more distant Turkic (Turkish), Afro-Asian (Tarifit, Moroccan-Arabic) or Slavic (Polish) language. Frisian is a regional language which is formally acknowledged as a second language by the government in the Dutch province of Fryslân in the north of the Netherlands. Limburgish is recognized as a regional language spoken in the south of the Netherlands. In terms of language distance, Dutch, Frisian, and Limburgish are close to each other and share many properties including lexical, morphological, and syntactic properties. For instance, many Dutch, Frisian, and Limburgish words are cognates, like the Frisian–Dutch pair *kâld* [kɑ:t] and *koud* [kaut] ('cold') and the Limburgish–Dutch pair *tandj* [tantʃ] and *tand* [tant] ('tooth'). All three languages are fusional inflecting languages, with morphological paradigms that are comparable in richness and the obligatory presence of a subject (no pro-drop), and share the same basic word order (mixed SOV/SVO and Adjective–Noun). Turkish, the Berber language Tarifit, Moroccan-Arabic, and Polish are morphologically richer languages than the three West-Germanic languages, and allow pro-drop. Moroccan-Arabic (in contrast to classical Arabic which is typically VSO) and Tarifit, however, display both VSO/SOV and N–A as basic word order, and allow for zero-copula. Moroccan-Arabic just like the Germanic varieties has a definite article whereas Tarifit and Turkish do not (Nortier, 1990, p. 43). Importantly, both in Moroccan-Arabic and Tarifit words are built on a basic consonantal skeleton

(root) which can be modified by combining the radicals with vowels and other consonants according to fixed patterns, as in Moroccan-Arabic *ktab* ‘book’, *kteb* ‘he wrote’, *ka-nketbu* ‘we are writing’, and *mektaba* ‘library’ (Nortier, 1990). In the remainder of the article children whose home language is Moroccan-Arabic and Tarifit are treated as one group and referred to as ‘Moroccan–Dutch children’ based on country of descent. Turkish is an agglutinative language with a highly flexible word order but its ‘unmarked’ order is SOV. The basic word order in Polish is SVO.

Our first expectation was that the bilingual group, as a whole, would be outperformed by the monolingual control group, consistent with research that found receptive vocabulary gaps between bilingual and monolingual school-aged children, and research that reports effects of distributed receptive vocabularies in bilingual children (Bialystok et al., 2010; Engel de Abreu et al., 2013). We also expected that more specific comparisons between monolinguals and bilingual subgroups would nuance this overall impression. Given the small linguistic distance between Dutch and Frisian and Dutch and Limburgish, we expected that the Frisian and Limburgish children would score similarly on the PPVT as their monolingual Dutch age peers (Francot et al., 2017), unlike bilingual children who are exposed to a more distant language, like the Polish, Moroccan or Turkish children in the Netherlands. These children are expected to be less well able to recognize words in Dutch because the migrant languages they are exposed to at home facilitate Dutch word recognition much less than the Dutch regional languages do.

## Method

### *Participants*

The data analyzed for this study were collected within three larger projects investigating the language and cognitive development of diverse groups of children in the Netherlands at Utrecht University, the Fryske Akademy, and Maastricht University. The children were 6 or 7 years old at time of testing (range 72–95 months). Data from the Turkish, Moroccan, Frisian, and monolingual Dutch children were part of longitudinal data sets comprising three waves of data collection with yearly intervals. In the current study, wave 2 data were included when most children were either 6 or 7 years old. Children with nonverbal intelligence scores, measured with the short version of the Wechsler Nonverbal Intelligence Scale (Wechsler & Naglieri, 2008), below 70 were excluded. Information on the language environment at home was collected with a parent questionnaire, the Questionnaire for Parents of Bilingual Children (PaBiQ; Tuller, 2015), described below in greater detail. Information obtained with the PaBiQ was used to assign children to the monolingual or one of the bilingual groups (Turkish–Dutch, Moroccan–Dutch, Polish–Dutch, Frisian–Dutch, Limburgish–Dutch). All bilingual children had at least one parent who interacted with them at home in the non-Dutch language, and had received education in Dutch for at least one year at time of testing.

To investigate the relation between cross-language distance and vocabulary outcomes, the bilingual children were assigned to two larger subgroups. The Limburgish–Dutch and Frisian–Dutch children were assigned to the Close group. The Polish–Dutch, Moroccan–Dutch, and Turkish–Dutch children were assigned to the Distant group. We

verified in two ways whether the binary split between Close and Distant languages was justified. First, we calculated normalized Levenshtein distances using a custom-made computer program that calculates the distances between pairs of languages by comparing words on a 40-item word list using an algorithm proposed by Levenshtein (Automated Similarity Judgment Program, asjp62; Bakker et al., 2009; Wichmann et al., 2016; see also Gampe et al., 2018, who used a similar procedure). The Levenshtein distance is an index of the least costly set of operations (changes and additions) needed to transform one transcription into another (Heeringa, 2004); the fewer manipulations are needed the closer two languages are, e.g., it takes fewer operations to get from Limburgish *tandj* to Dutch *tand* than to get from Dutch *tand* to Polish *zqb* and even more operations are needed to get to from Dutch *tand* to Turkish *diş*). Normalized Levenshtein distances are controlled for length of the strings. Levenshtein distances are a reliable tool to identify cognates across languages (Schepens et al., 2012), and a higher normalized Levenshtein distance implies fewer cognates than a lower lexical distance score hence fewer possibilities for children to use their lexical knowledge in the one language to recognize words in the other language. The average normalized Levenshtein distance for the Close language group is 50.41 ( $SD = 4.27$ ), ranging between 45.76 for the distance between Limburgish and Dutch to 54.15 for the distance between (Northern) Frisian and Dutch. The average normalized Levenshtein distance for the Distant language group is 99.12 ( $SD = 2.55$ ), ranging between 94.71 for the distance between Polish and Dutch to 101.96 for the distance between Turkish and Dutch. The clustering within the two bilingual subgroups coupled with a large difference between the two bilingual subsamples supports the binary split into linguistically close and distant to Dutch. Second, we asked for each language two proficient bilinguals for each language pair to indicate cognate items in the first 12 sets of the Dutch PPVT (higher sets are irrelevant for the age range investigated for the current study). The percentage of cognates for Limburgish and Frisian is above 80%, for Polish around 30%, and below 20% for Turkish, Moroccan-Arabic, and Tarifit.<sup>1</sup>

The characteristics of the monolingual, bilingual Distant and bilingual Close samples are listed in Table 1. Age did not differ significantly across the three groups,  $F(2, 309) = 1.73$ ,  $p = .18$ ,  $\eta_p^2 = .01$ . Nonverbal intelligence differed across the three groups,  $F(2, 308) = 3.68$ ,  $p = .03$ ,  $\eta_p^2 = .02$ . Bonferroni posthoc tests indicated that the nonverbal intelligence score was higher in the Close than in the Distant group ( $p = .02$ ). The other differences were not statistically significant. Socioeconomic status (SES) was indexed by parental education. Information on parental education was provided by the PaBiQ and measured the highest degree obtained on a nine-point scale ranging from 1 = no education to 9 = university degree. For the purposes of the present study, the average of the fathers' and mothers' educational levels was calculated. SES varied across the three groups,  $F(2, 307) = 5.57$ ,  $p = .004$ ,  $\eta_p^2 = .04$ . Bonferroni posthoc tests showed that SES was lower in the Distant compared to the Close group ( $p = .003$ ). The other differences were not significant.

Bilinguals are in general notoriously heterogeneous in terms of language environment, and the same holds for the sample in the current research. Most bilingual children were born in the Netherlands. These children were exposed to (some amount of) Dutch from birth and could be considered simultaneous bilinguals for this reason. Their degree of exposure to Dutch varied considerably, however, as indicated by the measures and

**Table 1.** Properties of the sample.

	<i>n</i>	Boys/Girls	Age months	NVIQ	SES
Monolingual Dutch	39	22/17	82 (6)	105 (14)	6.5 (2.0)
Bilingual Turkish–Dutch	19	11/8	83 (4)	101 (14)	4.9 (2.2)
Bilingual Moroccan–Dutch	40	17/23	80 (6.5)	93 (13)	4.8 (2.3)
Bilingual Polish–Dutch	49	22/27	82 (7)	110 (13)	7.3 (1.4)
<b>Bilingual Distant</b>	<b>108</b>	<b>50/58</b>	<b>81 (6)</b>	<b>102 (15)</b>	<b>6.0 (2.3)</b>
Bilingual Frisian–Dutch	108	55/53	82 (6.5)	107 (14)	6.8 (1.3)
Bilingual Limburgish–Dutch	57	31/26	84 (6)	107 (12.5)	6.6 (1.8)
Bilingual Close	165	86/79	83 (6.5)	107 (14)	6.7 (1.5)

Note: *n* = number of children; NVIQ = nonverbal intelligence quotient score with mean of 100; SES = socioeconomic status based on parental education measured on a nine-point scale (average of both parents). Standard deviations in parentheses.

**Table 2.** Dutch input in the home environment.

	Prior input	Current input	Richness of input	Proficiency mother	Proficiency father
<b>Bilingual Distant</b>	<b>44% (15)</b>	<b>47% (21)</b>	<b>.77 (.17)</b>	<b>2.97 (1.16)</b>	<b>3.44 (1.35)</b>
Bilingual Close	35% (22)	35% (26)	.66 (.19)	4.02 (.74)	3.94 (.87)

Note: Prior input = percentage Dutch input before age 4; Current input = percentage Dutch input at time of testing; Richness of input = estimation of frequency of different activities in Dutch on a scale from 0 to 1; Proficiency = Dutch proficiency measured on a six-point scale. Standard deviations in parentheses.

standard deviations in Table 2 for Prior Dutch input, Current Dutch input, Richness of Dutch input, and the self-rated proficiency in Dutch of both parents in the Distant and Close groups. In the Polish sample, 12 children were born abroad and would qualify for this reason as sequential bilinguals who were first exclusively exposed to Polish and at a later age also to Dutch. The questionnaire indicated that the majority (75%) of the children in the Moroccan group were exposed to Tarifit-Berber, the other Moroccan children were exposed to Moroccan-Arabic.

### Measures and procedures

Receptive vocabulary in Dutch was measured with the PPVT-III-NL (Schlichting, 2005). The PPVT is a standardized receptive vocabulary test designed for the age range from 2 years and 3 months to 90 years. It contains 204 items divided over 17 sets. The sets are ordered according to difficulty and each set consists of 12 items. In this task, a child hears a stimulus word and is asked to choose the correct referent out of four pictures. The PPVT was administered and scored according to the official guidelines. This means that the starting set was determined by a child's age and the task was terminated after a child produced nine or more errors within one set. Raw scores were converted to standardized scores (WBQ) based on age-corrected normative scores.

Information on children's home language environment was obtained using the PaBiQ (Tuller, 2015). This questionnaire is the short version of a longer questionnaire piloted by research groups in several countries within COST Action IS0804 and provides information on Dutch input quantity (prior and current input), input richness, and the parents' self-rated Dutch proficiency. These factors are relevant to include while comparing different groups, because they influence the vocabulary development of monolingual and bilingual children (Place & Hoff, 2011; Scheele et al., 2010). Prior Dutch input is operationalized as the percentage of time a child was addressed in Dutch in the home environment before (s)he was 4 years old, as indicated by parents on a five-point scale ranging from 0 = never to 4 = always, corresponding to 0%, 25%, 50%, 75%, and 100%. Current Dutch input indicates the percentage of time a child was, around the time of testing, addressed in Dutch in the home environment. Current Dutch input was collected for the mother, father, other caretakers, and siblings on a five-point scale ranging from 0 = never to 4 = always, again corresponding to 0%, 25%, 50%, 75%, and 100%. These scores were added up and divided by the maximum score that was possible. This resulted in one value for the variable Current Dutch input. For example, a score of 0 (mother), 1 (father), 2 (sibling), would result in a score of  $(0+1+2)/(3*4) = .25$ , which would correspond to 25%. The language richness score indexed use of Dutch in reading activities, television and movies, and telling stories on a three-point scale ranging from 0 = never to 2 = every day, and use of Dutch with peers, and with friends of the family on a five-point scale ranging from 0 = never to 4 = always. These scores were added up and divided by the maximum score that was possible resulting in one score for Dutch Richness. For example, a score of 2 (reading activities), 1 (television and movies), 0 (telling stories), 2 (peers), and 3 (family friends) would result in a score of  $(2+1+0+2+3)/((3*2)+(4*2)) = .57$ . Self-rated Dutch proficiency was obtained for both parents on a six-point scale ranging from 0 = a few words to 5 = native proficiency.

This research was screened by the Standing Ethical Assessment Committee of the Faculty of Social and Behavioral Sciences at Utrecht University. Criteria were met and further verification was not deemed necessary. Parents of participants gave informed consent. All participants were tested individually in a quiet room at their school. The PPVT was administered by native or near-native Dutch assistants who had a high level of Dutch. In all groups, the PPVT was part of a larger test battery examining children's language and cognitive development. The PaBiQ was administered during an oral interview with one of the child's parents. The interview was conducted by bilingual assistants who were proficient in both Dutch and the home language and could therefore be carried out in the preferred language of the parent.

## Analyses

In order to replicate previous findings showing that bilingualism is related to lower vocabulary outcomes, a hierarchical regression analysis was performed with, in Model 1, nonverbal intelligence scores (NVIQ) and socioeconomic status (SES) as control variables. NVIQ and SES were included to control for differences between the groups in mental ability, a component of language aptitude that has been found to predict bilingual children's vocabulary scores in the majority language (Paradis, 2011), and linguistic

**Table 3.** Dutch receptive vocabulary (*SD*), range of vocabulary scores, and number of children with scores 1 *SD* below the norm sample mean.

	<i>n</i>	Mean vocabulary score ( <i>SD</i> )	Vocabulary score range	Number of children 1 <i>SD</i> below mean
Monolingual Dutch	39	111 (11)	87–138	0
Bilingual Turkish–Dutch	19	93 (15)	63–112	6
Bilingual Moroccan–Dutch	39 <sup>a</sup>	97 (13)	64–125	8
Bilingual Polish–Dutch	49	99 (14)	65–126	7
<b>Bilingual Distant</b>	<b>107</b>	<b>97 (14)</b>	<b>63–126</b>	<b>21</b>
Bilingual Frisian–Dutch	108	108 (10)	88–133	0
Bilingual Limburgish–Dutch	52 <sup>a</sup>	107 (8)	91–125	0
Bilingual Close	160	108 (9)	88–133	0

Note: *n* = number of children; Mean vocabulary score = quotient score with a mean of 100 and a standard deviation of 15; *SD* = standard deviation.

<sup>a</sup>Receptive vocabulary scores were not available for one and five children in the Moroccan and Limburgish samples respectively.

stimulation in the home environment (Hart & Risley, 1995; Hoff, 2006). In Model 2, the Group (monolingual versus bilingual) was added as a predictor, in addition to the two control variables.

The hierarchical regression analysis was followed by an ANCOVA with a three-level independent variable Group (monolingual, Distant, Close) as the between-subjects variable and NVIQ and SES as covariates to examine the role of cross-language distance and determine if the effect of bilingualism only holds for the Distant group. Whereas the hierarchical regression analysis looks at the overall effect of bilingualism, the ANCOVA compares the three different groups. The outcomes of an ANCOVA should be interpreted with caution as the groups differ significantly on the covariates NVIQ and SES (Field, 2013).<sup>2</sup> To strengthen the interpretation, we performed an ANOVA with matched groups, based on subsamples. Groups were matched on age, NVIQ, and SES. In matching, differences between the bilingual Distant and Close groups in Dutch input that could provide a rival explanation were included as well. We decided to include Dutch input by means of matching (instead of statistically controlling it in a regression analysis) because input variables tend to show high intercorrelations. High correlations between predictor variables results in multicollinearity, which renders the results of regression models unreliable. Further details of the matching procedure are in the Results section.

## Results

### *Comparing Dutch vocabulary across the three groups: full sample*

The vocabulary scores in Table 3 show that all groups score on average within the 1 *SD* range – standard score between 85 and 115 – that is, within the normal range of variation. There are, however, individual children who scored below 85, and this happened solely in the group of bilingual children whose non-Dutch language is distant from Dutch. To

**Table 4.** Summary of the hierarchical regression analysis of the variables predicting Dutch receptive vocabulary.

	$\beta$	$t$	$R$	Adj. $R^2$	$\Delta R^2$
Model 1			.38	.14	.14
NVIQ	.26	4.74***			
SES	.22	3.98***			
Model 2			.43	.18	.04
NVIQ	.26	4.89***			
SES	.22	4.07***			
Group	-.20	-3.90***			

Note: NVIQ = nonverbal intelligence quotient score with mean of 100; SES = socioeconomic status based on parental education measured on a nine-point-scale (average of both parents).

\*\*\*  $p < .001$ .

identify possible risk factors, we compared within the Distant group the NVIQ scores of the children who scored 1 *SD* below the mean on receptive vocabulary ( $M = 95$ ,  $SD = 16$ ) with the other children who scored within the normal range on receptive vocabulary ( $M = 104$ ,  $SD = 15$ ). This comparison indicated that the children in the first group had significantly lower NVIQ scores,  $F(1, 104) = 6.34$ ,  $p = .013$ ,  $\eta_p^2 = .06$ . Inspection of individual cases showed that there were seven children who scored 1 *SD* below the mean on receptive vocabulary but had above-average NVIQ scores ranging between 104 and 128, five of whom were sequentially (Polish–Dutch) bilingual children with relatively short exposure to Dutch. SES was also somewhat lower in the first group ( $M = 5.14$ ,  $SD = 2.25$ ) as compared to the second ( $M = 6.16$ ,  $SD = 2.25$ ), but this difference did not reach statistical significance,  $F(1, 104) = 3.45$ ,  $p = .066$ ,  $\eta_p^2 = .03$ .

Next, analyses were performed to test the effects of bilingualism and language distance. The hierarchical regression analysis demonstrated that Model 2 (which contained, besides the control variables, the variable ‘group’ that distinguished the bilinguals from the monolinguals) explained 18% of the variance (based on the adjusted  $R^2$ ), which was significant,  $F(3, 299) = 22.35$ ,  $p < .0005$ . Group predicted a significant amount of variance over and above the effects of control variables NVIQ and SES ( $\beta = -.20$ ,  $t = -3.90$ ,  $p < .0005$ ,  $\Delta R^2 = .04$ ,  $p < .001$ ), and confirmed that the bilingual children, as a group, were outperformed by the monolinguals on Dutch vocabulary. The summary of the regression models is presented in Table 4.

The outcomes of the ANCOVA in which the two different bilingual groups (Distant, Close) and monolinguals were compared indicated a main effect of group,  $F(2, 298) = 29.44$ ,  $p < .0005$ ,  $\eta_p^2 = .17$ , and of the covariates NVIQ,  $F(1, 298) = 19.56$ ,  $p < .0005$ ,  $\eta_p^2 = .06$ , and SES,  $F(1, 298) = 9.93$ ,  $p = .002$ ,  $\eta_p^2 = .03$ . Children with a higher NVIQ score and higher SES had higher Dutch receptive vocabulary outcomes. Bonferroni posthoc tests showed that the difference between the monolingual group and the Close group did not reach statistical significance ( $p = .11$ ) but that both the monolingual group ( $p < .0005$ ) and the Close group ( $p < .0005$ ) scored higher than the Distant group on Dutch receptive vocabulary.

**Table 5.** Comparison of the two matched groups based on Dutch input in the home environment, SES, age, and NVIQ.

	Prior input	Current input	Richness of input	proficiency mother	proficiency father	SES	Age	NVIQ
<b>Bilingual Distant (n = 45)</b>	<b>48% (15)</b>	<b>51% (22)</b>	<b>.82 (.15)</b>	<b>4.1 (.35)</b>	<b>4.0 (1.1)</b>	<b>6.5 (2.0)</b>	<b>82 (7)</b>	<b>102 (15)</b>
Bilingual Close (n = 45)	36% (22)	36% (25)	.67 (.18)	4.2 (.39)	3.9 (1.1)	6.7 (1.7)	83 (6)	103 (14)

Note:  $n$  = number of children; Prior input = percentage Dutch input before age 4; Current input = percentage Dutch input at time of testing; Richness of input = estimation of frequency of different activities in Dutch on a scale from 0 to 1; Proficiency = Dutch proficiency measured on a six-point-scale; SES = socioeconomic status based on parental education measured on a nine-point-scale (average of both parents); NVIQ = nonverbal intelligence quotient score with mean of 100. Standard deviations in parentheses.

### Comparing Dutch vocabulary across the three groups: matched groups

To exclude the role of confounding variables as much as possible, a second between-group analysis was performed in which the two bilingual groups were matched on Dutch input at home, in addition to NVIQ and SES.

For the Distant group, the Dutch input situation is rather favorable as they have relatively high Prior Dutch input, Current Dutch input, Richness of Dutch input, and father's proficiency. Only mother's proficiency was considerably lower in the Distant ( $M = 2.97$ ,  $SD = 1.16$ ) than the Close group ( $M = 4.02$ ,  $SD = .74$ ), as shown in Table 2. To exclude the possibility that input situation is a confounding factor, we created two matched groups, starting with only the children whose mothers indicated that they are proficient in Dutch (score 4–6). In the Distant and Close groups, 45 (42%) and 122 (74%) children had a mother who indicated to be proficient in Dutch, respectively. We did not match on Prior Dutch input, Current Dutch input or Richness of Dutch input. The reason is that if these variables have any effect, it can only have a positive effect on the Distant group because they score higher on these variables. Therefore, these variables do not provide an alternative explanation for the observed lower Dutch vocabularies of the Distant group. Next, we removed children from the Close group until the two groups were as close as possible in NVIQ by leaving out those children from the Close sample with an NVIQ score that was not present in the Distant sample. NVIQ was chosen as the starting point for matching because NVIQ correlated with the PPVT scores in both groups (Close:  $r(159) = .26$ ,  $p < .0005$ ; Distant:  $r(106) = .31$ ,  $p = .001$ ). Finally, as the mother's self-rated Dutch proficiency was still lower in the Distant group than in the Close group, we removed 15 children from the Close group with a score of 5 on mother's self-rated proficiency, using gender, which was unequally distributed in the two samples, as a secondary criterion.

These procedures resulted in two equally sized groups with 45 children each that both contained 24 girls and 21 boys. The Distant groups were comprised of 3 Turkish–Dutch, 21 Moroccan–Dutch (18 Tarifit, 3 Moroccan–Arabic), and 21 Polish–Dutch children. The Close group consisted of 26 Frisian–Dutch, and 19 Limburgish–Dutch children. The characteristics of the two matched groups are shown in Table 5. The two

resulting groups did not differ in NVIQ score, age, SES or father's proficiency and mother's proficiency.

The Distant group scored higher than the Close group on Prior Dutch input,  $F(1, 88) = 10.13, p = .002, \eta_p^2 = .10$ , Current Dutch input,  $F(1, 88) = 9.86, p = .002, \eta_p^2 = .10$ , and Richness of Dutch input,  $F(1, 87) = 18.42, p < .0005, \eta_p^2 = .18$ . Comparing the PPVT scores in the three groups (monolingual, Distant, Close) using an ANOVA analysis revealed an effect of group,  $F(2, 128) = 11.47, p < .0005, \eta_p^2 = .15$ . Bonferroni posthoc tests demonstrated no difference between the monolingual ( $M = 112, SD = 11$ ) and Close group ( $M = 107, SD = 8$ ) ( $p = .12$ ), but higher Dutch receptive vocabulary scores for the monolingual group compared to the Distant group ( $M = 101, SD = 12$ ) ( $p < .0005$ ) and higher Dutch receptive vocabulary scores for the Close group compared to the Distant group ( $p = .02$ ).

## Discussion and conclusions

Bilingual children tend to score lower on vocabulary tests than their monolingual peers if vocabulary is measured in one language (Bialystok et al., 2010; Scheele et al., 2010; Thordardottir et al., 2006; Vagh et al., 2009). The main aim of the present study was to determine if this effect depends on language distance. Some previous research suggests no effects of language distance (Bialystok et al., 2010), whereas other research indicates that a smaller linguistic distance predicts higher vocabulary outcomes (Floccia et al., 2018). Like Bialystok and colleagues (2010), we administered the widely used Peabody Picture Vocabulary Test (PPVT), and assigned children to a language group that is relatively close to the target language (Dutch) and a more distant group. We expected that our study would be better able to pick up any effects of linguistic distance on bilingual children's vocabulary knowledge: whereas the close language group in the study by Bialystok and colleagues was heterogeneous and included many languages that are very distinct from the target language, the close language group in our study was far more homogeneous consisting of regional languages that are typologically quite similar to Dutch.

The results showed that bilingualism predicted a significant amount of variance in Dutch receptive vocabulary scores over and above the effects of nonverbal intelligence and socioeconomic status, measured through the parents' level of education: the full bilingual group, not differentiated by language pair, scored on average lower on Dutch receptive vocabulary than monolingual Dutch controls. Children with more mental resources, who obtained higher nonverbal intelligence scores, knew more Dutch words, indicating that vocabulary development is driven by domain-general cognitive processes (Blom, 2019; Paradis, 2011). In addition, children whose parents were more highly educated had larger vocabularies. Different factors may contribute, and explain the effect of parental education, such as a greater lexical diversity in parental input (Hart & Risley, 1995), more home reading input (Prevoe et al., 2013), and richness in nonverbal communication in the preverbal phase of development (Rowe & Goldin-Meadow, 2009).

In subsequent analyses, we investigated if linguistic distance influenced the effect of bilingualism on children's Dutch receptive vocabulary outcomes. The Distant bilingual group (Polish–Dutch, Turkish–Dutch, Moroccan–Dutch) scored, on average, within the normal range of monolingual variation, similar to what has been found in other studies that compared the scores of bilinguals to monolingual norms (Hammer et al., 2014;

Pearson et al., 1993; Vagh et al., 2009). However, the normal range of variation is wide (Bialystok et al., 2010). Inspection of individual children revealed that 20% (21/107) of the children in the Distant group performed more than 1 *SD* below the mean of a monolingual normative sample. These children had had either relatively short exposure to Dutch or scored in the lower range on the nonverbal intelligence test. We expect the children with short exposure to catch up and reach monolingual norms eventually, but the children with fewer mental resources may show a different development. In this group, the vocabulary gap is expected to grow as children will lack the cognitive resources to counteract the Matthew effect (Stanovich, 1986), according to which ‘the rich get richer and the poor get poorer’, implying that children who know many words will also be better able to learn new words than children with a limited vocabulary. No child in the Close group (Frisian–Dutch, Limburgish–Dutch) scored below the  $-1$  *SD* threshold. The relatively good performance on Dutch receptive vocabulary in these two groups of regional language users is in line with the pattern found in the larger sample of 5- to 9-year-old Limburgish–Dutch children from which the current sample was drawn. In that study, we found that the Limburgish–Dutch sample scored significantly above the normative mean of 100 on the Dutch PPVT (Francot et al., 2017). Group comparisons confirmed this pattern: the Distant group had lower Dutch receptive vocabulary scores than the (very similarly performing) Dutch monolinguals and Close bilinguals. Subsequent analyses in which the Distant and Close groups were matched on nonverbal intelligence, age, socioeconomic status, and the parents’ self-rated Dutch proficiency demonstrated that despite the relatively larger quantity and richness of Dutch input at home in the Distant group, this group still scored lower on Dutch receptive vocabulary than the monolingual and the Close groups, confirming that linguistic distance plays a significant role.

Interestingly, the Distant group had had *more* Prior Dutch input, *more* Current Dutch input, and scored *higher* on richness of Dutch input than the Close sample with which they were matched. This could be due to a task-effect of reported behavior, as Dutch families with migration backgrounds are commonly subjected to discrimination and marginalization (Van de Weerd, 2019) and experience societal and political pressure to learn Dutch which, in turn, may lead to overreporting their use of Dutch. However, another study, based on the same sample, revealed that the parents’ report of amount and richness of Turkish and Tarifit at home correlates significantly with the Turkish–Dutch and Moroccan–Dutch children’s vocabulary scores in Turkish and Tarifit respectively (Blom, 2019), confirming the reliability of parental report of language use at home. The migrant languages of the Distant group may, in general, be less widely used outside the home environment than the regional languages in the Close group since these regional languages are also community languages. Consequently, the influence of Dutch, which is the nation’s dominant language, may be more dominant in the home environment in the Distant group than in the Close group, explaining why the children in the Distant group had had more Dutch input quantity and richness than the children in the Close group.

In addition to input differences, questions may arise regarding cultural differences and whether these may influence possibilities for conceptual transfer: is it likely that children in the Close group know more concepts assessed with the PPVT than the children in the Distant group because of cultural differences? Although we acknowledge that there are considerable cultural differences between the groups in the study, there are a number of

reasons why we think it is unlikely that cultural differences can explain all of the between-group differences. First, all children in the Close group were born in the Netherlands, but the same holds for nearly all children in the Distant group, increasing shared cultural experiences across the different groups. Second, especially within the Distant group, the three bilingual groups (Turkish–Dutch, Moroccan–Dutch, Polish–Dutch) have very different cultural backgrounds and vary widely in their experiences as immigrants, newcomers to Dutch society, in the neighbourhoods where they live, the extent of mobility, and religious and cultural identifications (Fought, 2006). If culture had been a major influence, or if vocabulary scores were indexical for migration experiences, we would have expected larger differences between their Dutch vocabulary scores.

Effects of cross-linguistic distance were expected beforehand based on several studies using a within-child design showing that bilingual children perform better on words that are very similar across the children's two languages (Bosch & Ramon-Casas, 2014; Bosma et al., 2019; Gampe et al., 2018; Kelley & Kohnert, 2012; Malabonga et al., 2008; Schelleter, 2002). Floccia and colleagues (2018) are, to our knowledge, the first to establish that linguistic distance has repercussions for between-child differences within the bilingual population. Our results show that their findings for toddlers learning English extend to 6- and 7-year-old children learning Dutch, and highlight that learning a closely related regional language is not associated with low vocabulary in the national language. Regarding the underlying mechanisms, we cannot provide any definite conclusions. However, given the robust evidence for the facilitating effects of cognates, including evidence for this effect in Frisian–Dutch children (Bosma et al., 2019), it is highly probable that bilingual children benefit directly from lexical overlap, and large numbers of cognates. Floccia et al. (2018) found that structural overlap predicted English receptive vocabulary in their study with toddlers, but lexical overlap did not. Kelley and Kohnert (2012) have argued that sensitivity to cognates is dependent on growing metalinguistic awareness. Possibly, the children's young age and their lack of metalinguistic awareness explain the absence of an effect in the study by Floccia and colleagues. As Frisian and Dutch and Limburgish and Dutch are structurally very similar, we assume that the Frisian–Dutch and Limburgish–Dutch children in our study benefited from structural overlap, in addition to lexical overlap. Structural knowledge contributes to vocabulary learning and provides information about the meaning of new words (Chemla et al., 2009; Gleitman, 1990; Samuelson & McMurray, 2016), as well as processing routines that facilitate input processing (Lew-Williams & Fernald, 2007). Evidence for significant cross-language cross-domain (lexicon, grammar) correlations is limited (Bedore et al., 2010; Simon-Cerejido & Gutiérrez-Clellen, 2009), but this research has not investigated any modulating effects of language distance.

A final point that we would like to discuss concerns the PPVT, a widely used standardized instrument. Goriot and colleagues (2018) warn researchers against the use of the PPVT as a measure for children's vocabulary in the L2 in case different first language (L1) groups are compared, as children whose L1 resembles the L2 benefit from cognates and will for this reason perform better on the test than children whose L1 has fewer cognates, invalidating cross-linguistic comparisons. We do not fully understand this warning as cognates are part of a bilingual child's vocabulary knowledge, both in the L1 and L2, and a child whose L1 resembles the L2 is more likely to have better lexical abilities in the L2 compared to a child whose L1 does not resemble the L2 because of this. In addition,

the alternative – a vocabulary measure that is insensitive to the benefits provided by cognate items and only includes non-cognate items – may be less representative of the target language than a vocabulary measure with cognates, in particular in cases of language pairs with many cognates such as Frisian–Dutch or Limburgish–Dutch.

### *Implications*

The outcomes are important for a better understanding of bilingualism and illustrate that, in addition to input, which has received much recent attention in the literature on bilingual children's vocabulary development, linguistic distance determines bilingual children's vocabulary knowledge. The between-subjects design highlights linguistic distance as an important individual difference factor, and the results emphasize the need for a nuanced view on bilingualism (Dixon et al., 2012; Grosjean & Li, 2012; Luk & Bialystok, 2013). The results of our research can help avoid unnecessary worries about bilingual children, and contribute to the identification of those factors that do constitute a risk. The identification of risk factors is important as low vocabulary may lead to impaired or delayed literacy development (August & Shanahan, 2006; Dickinson & Tabors, 2001; Oller & Pearson, 2002; Snow et al., 1998). Comparisons to monolingual norms and the finding that none of the Frisian–Dutch and Limburgish–Dutch children performed below monolingual norms show that growing up with multiple languages is not a risk factor if the children's two languages provide ample opportunities for transfer. Neither is a lack of such opportunities necessarily problematic as the majority of the bilingual Turkish–Dutch, Moroccan–Dutch, and Polish–Dutch children performed within monolingual norms. A large linguistic distance may become a risk factor when the effects of multiple risk factors accumulate, e.g., large linguistic distance coincides with low parental education and, particularly, nonverbal intelligence scores in the lower range.

### *Limitations and future research*

This study was focused on receptive vocabulary. Recent research has demonstrated that receptive and expressive vocabulary measure the same construct in preschoolers and school-aged children (Lonigan & Milburn, 2017), which may suggest that the outcomes generalize, at least to some extent, to productive vocabulary. Further research is recommended, however. In particular research that investigates if cognates, conceptual knowledge, and/or structural overlap impact differently on receptive versus expressive vocabulary could shed further light on the underlying mechanisms of linguistic distance effects (Floccia et al., 2018). Another limitation of our study is that we could not compare receptive vocabulary in both languages of the bilingual children, as Peabody Picture Vocabulary Tests are not available for the non-Dutch languages (Frisian, Limburgish, Polish, Moroccan-Arabic, Tarifit), or outdated (Turkish). Consequently, a full picture of the bilingual children's receptive vocabulary knowledge is lacking. It is expected that monolingual–bilingual differences in vocabulary will disappear if bilingual children's total vocabulary scores are taken into account (Engel de Abreu et al., 2013; Hoff et al., 2012; Pearson et al., 1993). In our study, language distance was treated as a binary variable with a large distance between the two values. The distance between languages is, however, gradual and the effects of language distance on bilingual children's ability to

recognize words are also expected to be gradual. Investigating the gradual nature of language distance was beyond the scope of our study, but we recommend this as an important venue for future research. Such research could combine gradualness as a between-subject factor that distinguishes between children with gradualness as a within-subject (and between-item) factor that distinguishes between different words that overlap more or less across languages. A between-subject design, as used in the present study, is typically threatened by confounding factors. We attempted to limit any confounding effects by exerting statistical control and through matching. In the current study, we did not conduct analyses that systematically compared performance across the different groups on cognate and non-cognate items. A reliable cross-language analysis would need to control for word difficulty because the likelihood of a cognate increases with difficulty of words (e.g., Bosma et al., 2019; Kelley & Kohnert, 2012), and the difficulty of translation equivalents is not necessarily the same across languages. In addition, such an analysis requires sufficient variation and balance in cognates and non-cognates for the different bilingual groups, which was not the case in our study. Combining a between-subject design with a within-subject design in a carefully designed experimental study that systematically manipulates cognate status, thereby controlling for word difficulty, would be a further step in excluding the effect of confounding factors.

### Acknowledgements

We are grateful to Ilias Alami, Marijke Andringa, Daria Boruta, Asli van der Baan-Ünlüsoy, Fatima Zohra Charki, Falco Cremers, Rianne Francot, Ben Hermans, Eric Hoekstra, Khalid Mourigh, Sukran Ulubas, and Paulina Wołoszyn for identifying Frisian–Dutch, Limburgish–Dutch, Moroccan Arabic–Dutch, Polish–Dutch, Tarifit–Dutch, and Turkish–Dutch cognate words in the PPVT.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: This study is funded with Grants from the Dutch Organization for Scientific Research (NWO Vidi, NWO Aspasia) awarded to Elma Blom, the Province of Fryslân, Road veur 't Limburgs, SWOL (Maastricht University/Chair Language culture in Limburg of Leonie Cornips), Department of Linguistics (Amsterdam University), and Meertens Institute Amsterdam.

### ORCID iD

Elma Blom  <https://orcid.org/0000-0001-8984-3650>

### Notes

1. The informants showed slight differences in their cognate judgments. The average percentages of cognates they indicated are 94%, 83%, 29%, 18%, 6%, and 5% for Limburgish, Frisian, Polish, Turkish, Moroccan-Arabic, and Tarifit respectively.
2. We decided to proceed despite violation of the assumption of independence as there is no better alternative. Dropping of the covariate and performing a one-way ANOVA would overstate the effect of Group. Transforming the continuous variables NVIQ and SES into dichotomous variables and performing a  $3 \times 2 \times 2$  ANOVA results in an unnecessary loss of information. The assumption of independence is moreover of particular relevance for experimental designs which crucially rely on random assignment, and less for an observational design, as used in the current research.

## References

- August, D., & Shanahan, T. (2006). *Report of the National Literacy Panel on Language-Minority Children and Youth. Developing literacy in second-language learners*. Erlbaum.
- Bakker, D., Müller, A., Velupillai, V., Wichmann, S., Brown, C. H., Brown, P., Egorov, D., Mailhammer, R., Grant, A., & Holman, E. W. (2009). Adding typology to lexicostatistics: A combined approach to language classification. *Linguistic Typology*, 13(1), 169–181.
- Bedore, L., Pena, E., Gillam, R., & Ho, T. (2010). Language sample measures and language ability in Spanish–English bilingual kindergarteners. *Journal of Communication Disorders*, 43(6), 498–510.
- Bhatia, T., & Ritchie, W. (1999). The bilingual child: Some issues and perspectives. In W. Ritchie, & T. Bhatia (Eds.), *Handbook of child language acquisition* (pp. 569–643). Academic Press.
- Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge University Press.
- Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, 13(4), 525–531.
- Blom, E. (2019). Domain-general cognitive ability predicts bilingual children's receptive vocabulary in the majority language. *Language Learning*, 69(2), 292–322.
- Blom, E., Cornips, L., & Schaeffer, J. (2017). *Cross-linguistic influence in bilingualism: In honor of Aafke Hulk*. John Benjamins.
- Bosch, L., & Ramon-Casas, M. (2014). First translation equivalents in bilingual toddlers' expressive vocabulary: Does form similarity matter? *International Journal of Behavioral Development*, 38(4), 317–322.
- Bosma, E., Blom, E., Hoekstra, E., & Versloot, A. (2019). A longitudinal study on the gradual cognate facilitation effect in bilingual children's Frisian receptive vocabulary. *International Journal of Bilingual Education and Bilingualism*, 22(4), 371–385.
- Chemla, E., Mintz, T., Bernal, S., & Christophe, A. (2009). Categorizing words using 'frequent frames': What cross-linguistic analyses reveal about distributional acquisition strategies. *Developmental Science*, 12(3), 396–406.
- Cummins, J. (2000). *Language, power and pedagogy: Bilingual children in the crossfire*. Multilingual Matters.
- De Houwer, A. (1995). Bilingual language acquisition. In P. Fletcher, & B. MacWhinney (Eds.), *The handbook of child language* (pp. 219–250). Blackwell.
- Dickinson, D. K., & Tabors, P. O. (2001). *Beginning literacy with language: Young children learning at home and school*. Paul Brookes Publishing.
- Dixon, L. Q., Wu, S., & Daraghme, A. (2012). Profiles in bilingualism: Factors influencing kindergartners' language proficiency. *Early Childhood Education Journal*, 40(1), 25–34.
- Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test* (3rd ed.). AGS Publishing.
- Engel, de Abreu, P., Baldassi, M., Puglisi, M., & Befi-Lopes, M. (2013). Cross-linguistic and cross-cultural effects on verbal working memory and vocabulary: Testing language-minority children with an immigrant background. *Journal of Speech, Language, and Hearing Research*, 56(2), 630–642.
- Field, A. (2013). *Discovering statistics using SPSS* (4th ed.). SAGE.
- Floccia, C., Sambrook, T. D., Delle Luche, C., Kwok, R., Goslin, J., White, L., Cattani, A., Sullivan, E., Abbot-Smith, K., Krott, A., Mills, D., Rowland, C., Gervain, J., Plunkett, K., Hoff, E., Core, C., & Patricia, J. (2018). Vocabulary of 2-year-olds learning English and an additional language: Norms and effects of linguistic distance. *Monographs of the Society for Research in Child Development*, 83(1), 7–29.
- Fought, C. (2006). *Language and ethnicity*. Cambridge University Press.

- Francot, R. J., van den Heuij, K., Blom, E., Heeringa, W., & Cornips, L. (2017). Inter-individual variation among young children growing up in a bidialectal community: The acquisition of dialect and standard Dutch vocabulary. In I. Buchstaller, & B. Siebenhaar (Eds.), *Language variation – European perspectives VI: Selected papers from the eight international conference on language variation in Europe (ICLaVE 8)* (pp. 85–98). John Benjamins.
- Gampe, A., Kurthen, I., & Daum, M. M. (2018). BILEX: A new tool measuring bilingual children's lexicons and translational equivalents. *First Language, 38*(3), 263–283.
- Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition, 1*(1), 3–55.
- Goriot, C., van Hout, R., Broersma, M., Lobo, V., McQueen, J., & Unsworth, S. (2018). Using the Peabody picture vocabulary test in L2 children and adolescents: Effects of L1. *International Journal of Bilingual Education and Bilingualism, 1*–23.
- Grosjean, F., & Li, P. (2012). *The psycholinguistics of bilingualism*. John Wiley & Sons.
- Hammer, C. S., Hoff, E., Uchikoshi, Y., Gillanders, C., Castro, D., & Sandilos, L. E. (2014). The language and literacy development of young dual language learners: A critical review. *Early Childhood Research Quarterly, 29*(4), 715–733.
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday education of young American children*. Paul Brookes Publishing.
- Heeringa, W. (2004). *Measuring dialect pronunciation differences using Levenshtein distance* [PhD Thesis, University of Groningen].
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review, 26*(1), 55–88.
- Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language, 39*(1), 1–27.
- Jia, G., Kohnert, K., & Collado, J. (2006). Action naming in Spanish and English by sequential bilingual children and adolescents. *Journal of Speech, Language, and Hearing Research, 49*(3), 588–602.
- Kelley, A., & Kohnert, K. (2012). Is there a cognate advantage for typically developing Spanish-speaking English-language learners? *Language, Speech and Hearing Services in Schools, 43*(2), 91–204.
- Kohnert, K. J., Bates, E., & Hernandez, A. E. (1999). Balancing bilinguals lexical-semantic production and cognitive processing in children learning Spanish and English. *Journal of Speech, Language, and Hearing Research, 42*(6), 1400–1413.
- Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science, 18*(3), 193–198.
- Lonigan, C. J., & Milburn, T. F. (2017). Identifying the dimensionality of oral language skills of typically developing preschool through fifth-grade children. *Journal of Speech, Language, and Hearing Research, 60*(8), 2185–2198.
- Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology, 25*(5), 605–621.
- Malabonga, V., Kenyon, D. M., Carlo, M., August, D., & Louguit, M. (2008). Development of a cognate awareness measure for Spanish-speaking English language learners. *Language Testing, 25*(4), 495–519.
- Marchman, V. A., Fernald, A., & Hurtado, N. (2010). How vocabulary size in two languages relates to efficiency in spoken word recognition by young Spanish–English bilinguals. *Journal of Child Language, 37*(4), 817–840.
- Meisel, J. (2001). The simultaneous acquisition of two first languages: Early differentiation and subsequent development of grammars. In J. Cenoz, & F. Genesee (Eds.), *Trends in bilingual acquisition* (pp. 11–42). John Benjamins.

- Nortier, J. (1990). *Dutch–Moroccan Arabic code-switching among Moroccans in the Netherlands*. Foris Publications.
- Oller, D. K., & Pearson, B. Z. (2002). Assessing the effects of bilingualism: A background. In D. K. Oller, & R. E. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 3–21). Multilingual Matters.
- Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism*, 1(3), 213–237.
- Paradis, J., & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or independent? *Studies in Second Language Acquisition*, 18(1), 1–25.
- Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning*, 3(1), 93–120.
- Place, S., & Hoff, E. (2011). Properties of dual language exposure that influence two-year-olds' bilingual proficiency. *Child Development*, 82(6), 1834–1849.
- Poarch, G. J., & Van Hell, J. G. (2012). Cross-language activation in children's speech production: Evidence from second language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology*, 111(3), 419–438.
- Prevo, M., Malda, M., Mesman, J., Emmen, R., Yeniad, N., van IJzendoorn, M. H., & Linting, M. (2013). Predicting ethnic minority children's vocabulary from socioeconomic status, maternal language and home reading input: Different pathways for host and ethnic language. *Journal of Child Language*, 41(5), 963–984.
- Prevo, M., Malda Emmen, R., Yeniad, N., & Mesman, J. (2015). A context-dependent view on the linguistic Interdependence Hypothesis: Language use and SES as potential moderators. *Language Learning*, 65(2), 449–469.
- Rowe, M. L., & Goldin-Meadow, S. (2009). Differences in early gesture explain SES disparities in child vocabulary size at school entry. *Science*, 323(5916), 951–953.
- Samuelson, L. K., & McMurray, B. (2016). What does it take to learn a word? *Wiley Interdisciplinary Reviews: Cognitive Science*, 8(1–2), e1421. <https://doi.org/10.1002/wcs.1421>
- Scheele, A. F., Leseman, P. P. M., & Mayo, A. Y. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics*, 31(1), 117–140.
- Schelleter, C. (2002). The effect of form similarity on bilingual children's lexical development. *Bilingualism: Language and Cognition*, 5(2), 93–107.
- Schepens, J., Dijkstra, T., & Grootjen, F. (2012). Distributions of cognates in Europe as based on Levenshtein distance. *Bilingualism: Language and Cognition*, 15(1), 157–166.
- Schlichting, L. (2005). *Peabody Picture Vocabulary Test-III-NL*. Harcourt Test Publishers.
- Serratrice, L. (2013). Cross-linguistic influence in bilingual development: Determinants and mechanisms. *Linguistic Approaches to Bilingualism*, 3(1), 3–25.
- Sheng, L., Lam, B., Cruz, D., & Fulton, A. (2016). A robust demonstration of the cognate facilitation effect in first-language and second-language naming. *Journal of Experimental Child Psychology*, 141, 229–238.
- Simon-Cerejido, G., & Gutiérrez-Clellen, V. (2009). A cross-linguistic and bilingual evaluation of the interdependence between lexical and grammatical domains. *Applied Psycholinguistics*, 30(2), 315–337.
- Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. National Academy Press.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21(4), 360–406.

- Thordardottir, E., Rothenberg, A., Rivard, M. E., & Naves, R. (2006). Bilingual assessment: Can overall proficiency be estimated from separate measurement of two languages? *Journal of Multilingual Communication Disorders*, 4(1), 1–21.
- Tuller, L. (2015). Clinical use of parental questionnaires in multilingual contexts. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Methods for assessing multilingual children: Disentangling bilingualism from language impairment* (pp. 301–330). Multilingual Matters.
- Vagh, S. B., Pan, B. A., & Mancilla-Martínez, J. (2009). Measuring growth in bilingual and monolingual children's English productive vocabulary development: The utility of combining parent and teacher report. *Child Development*, 80(5), 1545–1563.
- Van de Weerd, P. (2019). 'Those foreigners ruin everything here': Interactional functions of ethnic labelling among pupils in the Netherlands. *Journal of Sociolinguistics*, 23(3), 244–262.
- Wechsler, D., & Naglieri, J. A. (2008). *Wechsler Nonverbal Scale of Ability*. [www.pearsonclinical.nl/wnv-nl-wechsler-non-verbal](http://www.pearsonclinical.nl/wnv-nl-wechsler-non-verbal)
- Wichmann, S., Holman, E. W., & Brown, C. H. (Eds.). (2016). *The ASJP Database* (version 17). <http://asjp.clld.org/>