

TITLE: Predicting bilingual pre-schoolers' patterns of language development:
degree of non-native input matters

Sharon Unsworth
Radboud University, Nijmegen

Susanne Brouwer
Radboud University, Nijmegen

Elise de Bree
University of Amsterdam

Josje Verhagen
University of Amsterdam

Accepted for publication in *Applied Psycholinguistics* on March 4th, 2019

ABSTRACT

Whilst numerous studies have recently shown that variation in input quantity predicts children's rate of acquisition across a range of language skills, comparatively little is known about the impact of variation in input *quality* on (bilingual) children's language development. This study investigated the relation between specific quality-oriented properties of bilingual children's input and measures of children's language development across a number of skills whilst at the same time taking family constellation into account. Participants were bilingual pre-schoolers ($n=50$) acquiring Dutch alongside another language. Preschoolers' receptive and productive vocabulary and morphosyntax in Dutch were assessed. Parental questionnaires were used to derive estimates of input quality. Family constellation was first operationalised as presence of a native-speaker parent and subsequently in terms of patterns of parental language use. Results showed that proportion of native input and having a native-speaker parent were never significant predictors of children's language skills, whereas the degree of non-nativeness in the input, family constellation in terms of parental language use, and language richness were. This study shows that what matters is not how much exposure bilingual children have to native rather than non-native speakers, but how proficient any non-native speakers are.

Keywords ($n=5$): bilingualism, pre-schoolers, input quality, non-native input, degree of non-nativeness

INTRODUCTION

Differences in when, how and with whom bilingual children interact in each of their two languages entail that there is considerable variation in both the quantity and quality of bilingual experience from one child to the next. Numerous studies over recent years have shown that these differences in experience predict children's rate of acquisition across a range of language skills (see Unsworth, 2016 for review). The vast majority of these studies have focussed on variation in input *quantity*. Input quantity refers to the amount of exposure available to a child. This may be expressed as a relative or absolute value (e.g., number of hours per week or number of tokens; Grüter, Hurtado, Marchman & Fernald, 2014; De Houwer, 2014), and it may gauge the extent of exposure at the current time or cumulatively over time (e.g., Gutiérrez-Clellen & Kreiter, 2003; Thordardottir, 2011; Unsworth, 2013). Typically, input quantity is indexed as the percentage of a child's waking hours in which interlocutors speak a given language to the child, and is usually derived from parental questionnaire data (see Unsworth, 2018 for a review of such questionnaires).

Comparatively little is known about the role of variation in input *quality* on bilingual children's language outcomes. Input quality refers to the type of exposure available to a child. It typically involves some measure of diversity or "richness" (Jia & Fuse, 2007). There is evidence to suggest that various qualitative properties of bilingual children's input may impact on their rate of acquisition. These include, for example, the availability of speakers of the language in question (e.g., Gollan, Starr & Ferreira, 2015), the range of (early) literacy-related activities (e.g., Scheele, Leseman & Mayo, 2010), and whether children are exposed to language mixing (e.g., Byers-Heinlein, 2013). As with input quantity, measures of input quality may be absolute (e.g., the number of different conversational partners; Place & Hoff, 2011) or relative (e.g., the proportion of parental utterances containing code switches; Bail, Morini & Newman, 2015), and as noted by Paradis (2011, p. 217), they often incorporate some index of frequency, with the consequence that they are not exclusively qualitative in nature.

The aim of this study is to investigate which aspects of input quality best predict bilingual children's performance on a range of standardised language tests, with particular focus on the role of non-native input.

Effects of input quality on bilingual language development

The quality of input to which bilingual children are exposed may vary depending on a number of factors. These include the richness or intensity of the input, the context in which input is provided, variety in the number and type of sources providing input, as well as whether it is from native or non-native speakers.

A number of studies on the early child second language (L2) acquisition of English (Jia & Fuse, 2007; Paradis, 2011) have found that input richness, usually a composite measure comprising of the frequency and density of activities such as computer games, television, book-reading, and playing with friends, predicted children's scores on verbal morphology and vocabulary.

Input in a given language may be tied to a specific context or location. Hearing a language in different locations, and specifically at home compared to at school, may result in qualitative differences in the lexical and morphosyntactic

diversity and complexity of input in a child's two languages. Such differences may contribute to what has been referred to as the "distributed characteristic" of bilingual (lexical) knowledge (Bialystok, Luk, Peets & Yang, 2010; Oller, Pearson & Cobo-Lewis, 2007). They may also affect the (rate of) development of different types of linguistic and academic skills (e.g., Cummins, 1979).

Input from a variety of sources has been found to positively affect bilingual children's developing language skills. In one of the few studies to address this question, Place and Hoff (2011) observed that the number of different speakers providing input was a significant predictor of Spanish-English toddlers' vocabulary and grammar scores in English (see also Place & Hoff, 2016). Citing work by Richtsmeier, Gerken, Goffman and Hogan (2009) and Singh (2008), Place and Hoff (2011, p. 1835) suggest that hearing language from multiple speakers may help children to identify features relevant for the acquisition of (phonological) categories and words.

Input from a variety of sources may also increase the range of syntactic structures available in the input (Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010) as well as the functional significance of the language (Fishman, Cooper & Ma, 1971; Gathercole & Thomas, 2009). Indeed, variation in lexical and morphosyntactic diversity and complexity across individual speakers has been shown to predict variation in bilingual children's (rate of) acquisition. More specifically, it has been shown that children whose parents and teachers offer more diverse and complex input have a faster rate of acquisition (Bowers & Vasilyeva, 2011; Huttenlocher, Vasilyeva, Cymeran, & Levine, 2002). In studies on monolingual acquisition, variation in input quality in these terms has in turn been linked to differences in socio-economic status (Hart & Risley, 1995; Hoff, 2006).

Finally, the quality of bilingual children's input may vary as a result of who exactly is providing that input. Unlike most monolinguals, many bilinguals may hear input from native and non-native speakers (Fernald, 2006). Compared with (some) native speakers, non-native speakers are likely to have less diverse and less sophisticated vocabulary, less accurate and sophisticated morphosyntax and not be as phonologically accurate (e.g., Core & Hoff, 2014 qtd. in Place & Hoff, 2016). There is some evidence from monolingual children that in the early years, the quality of parental input in these terms predicts children's later language development. More specifically, using data from child-caregiver interactions at 18, 30 and 42 months, Rowe (2012) found that after controlling for input quantity, quality of caregiver input at 18 months, operationalised as lexical diversity (number of different words), lexical sophistication (number of "rare" words), and number of decontextualized utterances (i.e., utterances which refer to objects, events or people not present in the context), was a significant predictor of children's vocabulary scores at later ages.

If non-native-speaker parents provide input in their non-native language to children, then this could in principle have a similar effect. In fact, non-native input or exposure to a specific, non-native variety of the target language has been claimed to lead to incomplete acquisition or fossilisation in some bilingual child populations (Cornips & Hulk, 2008; Driessen et al., 2002). The proficiency level of the input-provider has also been argued to modulate the effect of input quantity: in a study on the early child L2 acquisition of vocabulary and tense morphology in English, Paradis (2011) found that overall amount of input at home was not a significant predictor of children's outcomes (see also Chondrogianni & Marinis,

2011). She argued that this was most likely due to the low proficiency level of the speakers providing this input. In other words, her study provides indirect evidence for the proficiency level of input-providers impacting on children's language outcomes.

The role of native versus non-native input in child bilingualism has been investigated more directly in two studies by Place and Hoff (2011, 2016). Both studies used the Language Diary method (De Houwer, 2011) to derive a number of measures of input quantity and quality, namely the number of single-language conversational partners, number of different speakers as sources of exposure, as well as proportion of language exposure from native speakers. These input measures were then used to predict bilingual English-Spanish toddlers' scores for vocabulary and grammar. In both studies, native input predicted children's scores in English, even after controlling for amount of input, although the amount of variance explained by native input was limited (e.g., between 4% and 5% in Place and Hoff, 2016). The authors speculate that this might be due to the relative high proficiency of the non-native speakers in their sample. They conclude that non-native input "is less beneficial [to language development] than native input" (p. 17; in line with e.g., Hammer, Komaroff, Rodriguez, Lopez, Scarpino, & Goldstein, 2012), although which specific properties of native input lead to these benefits remains unclear.

In sum, bilingual children's (rate of) language development has been shown to be affected by specific properties of the input, including its richness and the number of different speakers. However, previous results are mixed and effects are often limited in scope. There is some evidence that exposure to non-native vs. native input is related to children's rate of acquisition, although this relation may be indirect in the sense that expected effects of input quantity were not found when the proficiency level of the speakers providing said input was low. It remains unclear, however, whether it is the amount of native rather than non-native input which matters, or the quality of any non-native input. Previous studies have incorporated measures of one or the other, but not both at the same time. Furthermore, studies incorporating measures of parental proficiency have thus far focussed on maternal proficiency only. Given the heterogeneity in family constellations often present in studies on bilingual language development, it makes sense to consider the impact of non-native proficiency on children's developing skills more broadly.

The relation between family constellation, input quality/quantity and language development

Bilingual families come in many shapes and sizes. Patterns of parental language use vary. Sometimes parents consciously adopt a particular "language strategy", often as part of their own family's "language policy" (King & Fogle, 2017) whereas for other parents, the language or languages they use is more the result of happenstance than anything else. Irrespective of how parents' patterns of language use emerge, they offer us another way of tapping into the relation between properties of language input and children's language development.

The most well-known approach to raising bilingual children is the "one parent, one language" approach. In many cases, one of these two languages is the same as the majority language spoken by the wider community and consequently,

bilingual children raised within such a family constellation will typically hear more majority-language input than children raised in what is often referred to as a “minority language at home” situation, where majority language exposure is in principle restricted to outside the home. Needless to say, there is considerable variation within these two types of family constellation and other types exist, too, including, for example, constellations where one or both parents use two languages.

In a large-scale survey on the impact of parental language use on bilingual children’s development in their two languages, De Houwer (2007) observed that children were most likely to speak both the minority and the majority language when both parents spoke the minority language and at most one parent spoke the majority language at home. In their study on Spanish-English bilingual 25 month olds, Place and Hoff (2011) documented how family constellation (“native English mother + native Spanish father” vs. “native Spanish mother + native English father” vs. “native Spanish mother + native Spanish father”) impacted on the quantity and quality of input available to children. For example, they found that two thirds of the English input to children with a native English-speaking mother came from native speakers whereas when the father was a native speaker of English and the mother a native speaker of Spanish, native-speaker input in English was reduced to 28%. Furthermore, children’s language skills were also related to family constellation; for example, children without any native English-speaking parent had significantly smaller vocabularies than children with a native-speaker mother, and this finding was partially mediated by input quantity and the number of different speakers providing English input.

There is thus evidence to suggest that both quantitative and qualitative aspects of bilingual children’s input may vary as a function of family constellation and – no doubt to a certain extent as a result of this variation – family constellation is also a predictor of bilingual children’s language outcomes.

This study

The aim of the present study was to examine the relation between specific quality-oriented properties of bilingual children’s input and measures of children’s language development across a number of skills. The properties in question were:

- the proportion of input from native speakers;
- the degree of non-nativeness in the input; and
- input richness.

The proportion of input from native speakers consisted of an estimation of the relative amount of target language input at home provided by native speakers as opposed to non-native speakers, whereas the degree of non-nativeness focussed on the proficiency level of any such non-native speakers. Input richness reflected the extent to which families engaged in language and literacy activities in the target language (following Jia & Fuse, 2007; Paradis 2011). Specific details concerning the operationalisation of these three variables are given below in the METHOD section.

Given that family constellation has been shown to affect the quantity and quality of input available to bilingual children, we first documented the influence

of family constellation on these three properties of dual language exposure (following Place & Hoff, 2011). Family constellation was operationalised in two ways. First, we divided children into groups based on whether their parents were native speakers (i.e., a more quality-based division, along the lines of Place & Hoff, 2011). Second, we divided children into groups based on the extent to which the parents used the two languages (i.e., a more quantity-based analysis, along the lines of De Houwer, 2007). Subsequently, we analysed the impact of the relative proportion of native vs. non-native input, the degree of non-nativeness in the input, and input richness on children's acquisition of the majority language, in this case Dutch, whilst taking family constellation into account. In both analyses, we used these three variables of interest as predictors whilst controlling for input quantity and for other background variables known to affect bilingual language acquisition, namely socio-economic status (e.g., Cobo-Lewis et al., 2002), working memory (e.g., Gangopadhyay, Davidson, Weisman and Kaushanskaya, 2016) and gender (e.g., Place & Hoff, 2016). Two other factors which have been shown to play a role in bilingual/early child L2 acquisition, namely L1 transfer (e.g., Blom & Baayen, 2013) and age of onset (e.g., Meisel, 2009) were not included here. This is because for the sample of children in this study, the considerable variation in other languages precluded any analysis of this variable and all children were exposed to the target language before the age of three, which meant that to the extent that a specific age of onset could be pinpointed, there was hardly any relevant variation for this variable.

We administered tests tapping into a number of language skills (i.e., receptive and productive one-word vocabulary, semantic fluency, receptive and productive morphosyntax). Our expectations were that more native input, non-native input from more proficient speakers, as well as richer input more generally would be associated with a faster rate of acquisition and hence would predict children's scores on these tests.

METHOD

Participants

Participants were 50 bilingual children aged 3 years ($M = 41$ months, $SD = 5.1$, range 31 - 49 months; 26 girls, 24 boys), recruited from preschools in the Netherlands. An additional three children were also tested but they were excluded from analyses because they failed to complete most of the tasks ($n = 2$) or because their parent did not participate ($n = 1$). These centres are attended by 2- to 3-year-old children, on average for four half-days a week (see e.g., Slot, 2014, for more information about and an evaluation of such centres and their educational programmes). Eligibility is determined by local authorities, usually via baby and toddler clinics. Children who attend are usually considered at risk of a language disadvantage, either because one or both of their parents speak a language other than Dutch and/or because their parents have a low level of education. By and large, this was also the case for the children in our sample, although they came from families from a whole range of socio-economic backgrounds, as measured by maternal education on a scale from 0, pre-primary, to 6, postgraduate degree ($M = 3.4$, $SD = 1.3$).

The bilingual children were all acquiring Dutch plus one of a range of other languages: Armenian (2), Berber (12), Chinese (1), English (3), Farsi (3), Greek (3), Indonesian (1), Italian (1), Kurdish (1), Polish (5), Russian (2), Serbian (1), Spanish (4), Surinamese (1), Turkish (9), Vietnamese (1). Three children were also exposed to a third language via their parents (Russian, Kurdish, Greek), four by other members of their family (Bulgarian, English, Kurdish, Greek) and one child had been exposed to Spanish at daycare (in addition to English/Dutch) until the age of 2;6. Parental report of home language skills suggested that there was a wide range of ability in the non-Dutch language(s) (using a modified version of the Alberta Language and Development Questionnaire, Paradis, et al., 2010; $M = 24$ (out of 33), $SD = 5.4$).

The main selection criterion for participation was that at least one parent was a non-native speaker of Dutch and used Dutch with the child at least some of the time. In the process of recruiting participants, non-native speaker was operationalised as being a speaker with an age of onset later than birth *and* (self-)reporting as having non-native proficiency in a detailed parent questionnaire – see *Language exposure and use* section below for further information. This resulted in a sample including family constellations which varied in terms of the number of non-native speakers, the level of Dutch language proficiency of these non-native speakers, and the amount of Dutch language input provided at home overall. The sample was thus heterogeneous in nature but ecologically valid in that it was typical for the type of early childhood education centre from which it was drawn.

Language tasks

Receptive vocabulary. A shortened version of the Dutch *Peabody Picture Vocabulary Test* (PPVT-III-NL; Dunn & Dunn, 2005) was used to assess receptive vocabulary (cf. Verhagen, de Bree, Mulder, & Leseman, 2017; Verhagen, Boom, Mulder, de Bree, & Leseman, in press). In this task, children were asked to select one out of four pictures that best matched an orally presented word, following the standard protocol. A shortened version was used to reduce testing time and fatigue. Specifically, items were removed from the original test on the basis of pilot data showing that they did not differentiate well across 3-year-old children (Verhagen et al., 2017; in press). To facilitate administration and scoring, the task was administered on a laptop and responses recorded with a button press by the experimenter. A fixed number of 24 items were presented to all children. Internal consistency of the task was sufficient ($\alpha = .73$, Verhagen et al., 2017). The dependent variable used in the analyses was total number of items correct.

Active vocabulary. The Active Vocabulary subtest of the *CELF Preschool-2-NL* (Wigg, Secord, Semel, & de Jong, 2012) was used to measure expressive vocabulary skill in Dutch. In this test, children name pictures in response to questions asked by the assessor such as “What is this?” or “What is the girl doing?”. The test contains 20 items, but testing is adaptive, such that administration is stopped when a child makes six consecutive errors. Raw scores rather than standard scores were used in the analysis, for the following reasons: first, the standard scores are normed for monolinguals only; second, some of the children in the current sample fell (just) under the age for which standard scores can be

calculated and would have to be excluded from analysis; third, using total (raw) scores allowed for the full range of variability to be included in the analysis. On this subtest, children were awarded 2 points when their answer was completely correct, and for 8 of the 20 items they received 1 point for responses which were approximating the right answer (e.g., *babykoe* 'baby cow' instead of *kalf* 'calf'). The test-retest reliability coefficient for 3-year-old children is excellent ($r = .90$; Wigg et al., 2012). The dependent variable used in the analyses was total number of items correct.

Semantic fluency. Semantic fluency or category fluency involves the ability to quickly generate items belonging to a certain semantic category within a given timeframe. It is thought to tap into lexical knowledge and retrieval and semantic (memory) organisation (Ardila, Ostrosky-Solís & Bernal, 2006) and has been found to correlate with general language proficiency (Bialystok, Craik & Luk, 2008). Two categories were used (*food* and *animals*) and children were given one minute to name as many items as they could think of (following the protocol used in Peña et al., 2003). Children's answers were recorded and scored afterwards by the experimenter following a pre-determined set of criteria as the total number of responses for each child within one minute, minus repetitions, words from another semantic category, and unintelligible responses. Scoring for all children was subsequently checked by the first author. Given that this task was included to measure lexical knowledge in Dutch, responses in the child's other language were excluded from the analysis. When children responded in a language other than Dutch, they were encouraged to produce the same (or another) word in Dutch. The dependent variable used in the analyses was the sum total of unique answers in both categories.

Sentence comprehension. Children's ability to understand spoken sentences in Dutch was assessed with the Sentence Comprehension subtest of the *CELF Preschool-2-NL* (Wigg, Secord, & Semel, 2012). In this test, children choose which one out of four pictures best matches a spoken sentence. The test contains 22 items, but testing is stopped when a child makes five consecutive errors. Again, for the reasons mentioned above, raw rather than standard scores were used, and the dependent variable used in the analyses was total number of items correct.. The test-retest reliability coefficient for 3-year-old children is acceptable ($r = .73$; Wigg et al., 2012).

Morphosyntax. The Word Structure subtest of the *CELF Preschool-2-NL* (Wigg, Secord, & Semel, 2012) was used to assess morphosyntax in Dutch. Specifically, this subtest assesses children's knowledge of subject-verb agreement, adjectival inflection, diminutives, noun plurals, and pronouns. The test contains 23 items, but testing is discontinued after seven consecutive errors. As for the other CELF subtests, the dependent variable used in the analyses was total number of items correct. The test-retest reliability coefficient for 3-year-old children is acceptable ($r = .74$; Wigg et al., 2012).

Other measures

Non-verbal working memory. Non-verbal working memory, or the ability to manipulate non-verbal information stored in memory, was assessed with the Hand Movements subtest of the Kaufman Assessment Battery for Children (KABC, Kaufman & Kaufman, 2004). In this test, children imitate a series of taps the assessor makes on the table with the fist, palm, or side of the hand. The test contains 12 items, divided into units of 3 or 4 items. It is adaptive, such that testing stops when children fail on all items in a unit or when children fail on one item after having provided correct answers to all items in the last unit intended for their age. The resulting score is the total number of items for which children imitated the research assistant's hand movements in the correct order. No information is available about the reliability of this subtest with this age group.

Language exposure and use. Bilingual children's current and previous patterns of language exposure and use were estimated using a detailed parental questionnaire (Bilingual Language Experience Calculator, BiLEC; Unsworth, 2013). Following Gutiérrez-Clellen and Kreiter (2003), this questionnaire asks parents to indicate where and with whom the child spends time on an average day in the week and an average day at the weekend, for how long, and which language(s) each person uses when addressing the child and how well they speak that language, as well as time spent on extra-curricular activities and the language(s) in which these occurred. Using this information, we calculated a general measure of *Input Quantity*, namely the child's relative exposure to Dutch at the current time, including sources at home, preschool and elsewhere such as tv, tablets and friends, and a measure of their exposure over time (i.e., *cumulative length of exposure*, following Unsworth, 2013). In addition, we calculated the proportion of exposure at home which was from native (NS) vs. non-native speakers (*% NS input*), as well as the average proficiency level of any input provided by non-native speakers (NNS), on a scale from 0, no fluency, to 5, native fluency. The latter variable was calculated in two different ways, first by simply averaging the proficiency level of all input-providers at home aged 4 years and older (*Average Quality NNS input*), and second by weighting the relative contribution of each depending on the amount of time they spent with the child (*Weighted Quality NNS input*). For more details, see Unsworth, 2013 and the manual available for download via the IRIS online instrument repository at www.iris-database.org.

Data collected from non-native parents as part of the wider project indicated that the (self-)reported proficiency used to derive our predictor variables was valid. More specifically, when the data collected from non-native parents (n=33) using a film retell task (Dimroth, Andorno & Benazzo, 2010) were evaluated by eleven native-speaker judges with extensive experience in teaching Dutch as an L2, the inter-class correlation was in the excellent range (Cicchetti, 1994) for all variables (i.e., vocabulary, grammar, accent, fluency and overall proficiency), and importantly, for present purposes, there was a strong, positive correlation between the average rating by the native-speaker judges and the (self-)reported proficiency data collected using the BiLEC questionnaire ($r(33) = .69, p < .001$).

Participating parents were categorised as a native speaker if they met the following criteria: age of onset was below 4 years old (following the broad consensus in the literature – see e.g., McLaughlin, 1978; Meisel, 2009; Unsworth,

2013) and the (self-)reported proficiency provided in the BiLEC questionnaire was nativelike (i.e., 5 on the scale given above). For the most part, these two criteria coincided. There were, however, a number of parents ($n = 12$) who reported nativelike proficiency in the questionnaire but who had an age of onset older than four. In these cases, we allowed (self-)reported proficiency to “trump” age of onset for two reasons: first, there exist individual late learners – albeit not many – who achieve nativelike competence (Hyltenstam & Abrahamsson, 2008), and second, (self-)reported proficiency – albeit not entirely unproblematic – is probably a more accurate estimation of parents’ abilities than a more general premise based on the literature, and as such likely provides a better characterisation of input quality, the variable we were ultimately trying to operationalise. Native-speaker judgements, available for seven of the twelve cases, supported this decision. All but two of these parents had scores which placed them in the top 20% of our sample (i.e., > 6.0 , on a scale from 0 to 9, where the highest score obtained was 7.0); they and the five parents for whom no native-speaker judgement data were available, were classified as native speakers. The remaining two parents, who had considerably lower native-speaker ratings (i.e., 3.0 and 4.2) were classified as non-native speakers. The potential risk here is that there will be parents who were categorized as native speaker but who in fact were not. We return to this question in the discussion. This categorisation resulted in 14 mothers and 14 fathers being categorised as native-speakers of Dutch. The average age of onset to Dutch for the non-native speaker parents who spoke Dutch to their child was 18.1 years for the mothers ($SD = 8.0$; range: 0 to 32) and 12.3 years for the fathers ($SD = 11.1$; range: 0 to 36).

To gain a better understanding of the families’ language and literacy practices at home, especially those in Dutch, a second questionnaire was administered. This Daily Communication Questionnaire (Mayo & Leseman, 2006; Scheele et al., 2010) consists of 32 questions about the frequency of various language and literacy activities (on a scale from 0, never, to 5, daily) and the language in which such activities take place (as a proportion of the total amount of time dedicated to said activities). These activities include: watching (specific kinds of) television, reading/being read to, parent-child interactions, singing and storytelling and educational conversations (e.g., talking about shapes and colours). Cronbach’s alpha ($\alpha = .84$) indicated good internal consistency. Responses were averaged across all questions and subsequently weighted for the extent to which they took place in Dutch by multiplying the average score for a given activity by the proportion of time the activity was carried out in Dutch.

Procedure

Informed consent was obtained from all parents. Children were tested individually by trained research assistants in a quiet room at preschool or at their homes, depending on the family’s preference. Tests were administered in a fixed order (i.e., PPVT, CELF subtests, Kaufman Hand Movements, semantic fluency) and interspersed with a number of other tasks assessing children’s knowledge of specific grammatical phenomena (i.e., definiteness, word order) as part of a larger project. Children received a sticker after each task, and a small gift at the end of the session. Parents completed both the BiLEC questionnaire and the Daily Communication Questionnaire via an interview with a trained assistant. In

addition to the film retell task mentioned briefly above, parents also performed a number of other language tasks not reported here.

Analysis

To determine which properties of bilingual children's language experience predicted their receptive and productive skills in vocabulary and morphosyntax, we conducted a series of multiple linear regression analyses with children's accuracy scores as dependent variable using the *lm* function in the *lme4* package (Bates, Martin, Maechler, Bolker, & Walker, 2015) in R (R Core Team, 2017). For each task, we first entered the "baseline" predictors (i.e., *Age*, *Gender*, *Working Memory*, *SES*). Subsequently, we added our predictors of interest in several separate models. These were:

- the proportion of input from native (rather than non-native) speakers (*% NS input*);
- the degree of non-nativeness of any non-native input (either *Average quality NNS input* or *Weighted quality NNS input*);
- the extent to which families engaged in language and literacy activities in Dutch (*Input richness*);
- overall proportion of exposure to Dutch (*Input quantity*); and
- family constellation.

With respect to the latter variable, family constellation, two series of analyses were performed. In the first, family constellation was operationalised as the presence or absence of a Dutch native-speaker parent (*NS parent*). In the second, family constellation was based on parental language use (*Parental language use*), that is, whether one, both or neither parent mostly used Dutch as language of communication with their child. For degree of non-nativeness, analyses were conducted with both variables, but only one at a time.

Other variables that were potential predictors of interest, namely past exposure (indexed by cumulative length of exposure) and children's own language use, showed multicollinearity with *Input quantity* ($r = .48, p < .001$ and $r = .74, p < .001$, respectively) and hence only *Input quantity* was included in the analysis, as outlined above. We return to this issue of multicollinearity in the discussion.

In all models, orthogonal sum-to-zero contrast coding was applied to our categorical fixed effects (i.e., *SES*, *gender* and *family constellation*), and all continuous variables were centered around zero (Baguley, 2012, pp. 590 – 621). For *Gender* girls (coded as 0.5) were contrasted with boys (reference, coded as -0.5). For *SES*, there were two contrasts: the first contrast compared primary (reference, coded as 0) with secondary (-0.5) and university level education (0.5), and in the second contrast, secondary and university level were compared to each other. For *NS parent* (at least) "one NS parent" (coded as 0.5) was compared to "no NS parent" (reference, coded as -0.5).¹ Finally, for *Parental language strategy*, there were also two contrasts: in the first contrast the group where both parents mostly speak heritage language (i.e., "mostly HL"; the reference group and as such coded a 0) was compared with the group where one parent mostly speaks Dutch, the other mostly speaks HL (i.e., "HL + Dutch"; coded as -0.5) and the group where both parents mostly speak Dutch (i.e., "mostly Dutch"; coded as 0.5), and in the second, the latter two groups were compared with each other.

For each model, a stepwise variable selection procedure was conducted in which non-significant predictors were removed to obtain the most parsimonious model. Model complexity was increased by including interaction effects between the predictors of interest and the baseline predictors and between different predictors of interest. In order to compare models, likelihood ratio tests were performed that compared the goodness of fit using the *anova* function in the *base* package (R Core Team, 2017). In this way, the final model was selected by checking whether the *p*-value from the likelihood ratio test was significant.

RESULTS

Descriptives (All children)

Table 1 presents an overview of the experiential variables derived from the parental questionnaires. To provide a more complete picture of the sample, this overview includes more general measures of bilingual experience in addition to the variables used in the analyses below.

--- INSERT TABLE 1 ABOUT HERE ---

The data in Table 1 show that on average, the bilingual children in our sample heard and used Dutch at home more often than their other language, that about half of their Dutch input was from non-native speakers and that the average (self-)reported proficiency level of these non-native speakers was “quite fluent”.²

Table 2 provides the descriptive statistics for children’s scores on the five Dutch language tasks and the non-verbal WM task.

--- INSERT TABLE 2 ABOUT HERE ---

The standard deviations on all tasks were high, indicating considerable individual variation.

Family constellation based on parents’ native speaker status

Children were divided into two groups based on whether their parents were native or non-native speakers of Dutch: a group where neither parent was a NS (*n* = 20) and a group where at least one parent was a NS (*n* = 30). The characteristics of these two groups with respect to our predictors of interest are given in Table 3.

--- INSERT TABLE 3 HERE ---

The groups were comparable in terms of *Input quantity* ($t(32.0) = -1.33, p = .194$), although there was more variation in the group without any NS parents. (NB: The alpha level here was corrected (from .05 to .01) to prevent Type 1 errors resulting from the multiple *t*-tests (*n* = 5) carried out with the same groups.) Children in both groups were exposed to NNS input but unsurprisingly, the proportion of NS input was lower in the “No NS parent” group ($t(48) = 3.47, p = .001$). The average quality of any NNS input did not significantly differ across groups, however, irrespective of how this was measured (*Average quality NNS input*: $t(43) = -2.03,$

$p = .048$; *Weighted quality NNS input*: $t(43) = -2.07$, $p = .044$). There was also no significant difference in terms of *Input richness* ($t(48) = -0.81$, $p = .420$).

Table 4 provides the descriptive statistics for children's scores on the five Dutch language tasks for the two family constellation groups based on parents' native speaker status.

--- INSERT TABLE 4 ABOUT HERE ---

The most parsimonious regression models for each task are given in Table 5.

--- INSERT TABLE 5 ABOUT HERE ---

The results for the three vocabulary tasks were as follows. For the PPVT (Table 5A), the results showed that children of university-level-educated mothers had significantly higher scores than the children with mothers educated to secondary level. Furthermore, the difference between children of primary-level-educated mothers and children of mothers educated to at least secondary level was approaching significance. In addition, the degree of non-nativeness was also a significant predictor of children's scores. For the active vocabulary subtest of the CELF (Table 5B), age and input quantity were significant predictors of children's scores, with older children and children with more overall exposure to Dutch obtaining higher scores. Finally, on the semantic fluency task (Table 5C), older children and children with a richer input named significantly more words than younger children and children with a less diverse input. Note, however, that when *Input richness* was not included in the model for semantic fluency, *Input quantity* became a significant predictor (cf. Table A in the Appendix). This was most likely the result of the strong correlation between these two variables ($r(50) = .71$, $p < .001$).

Turning to the two morphosyntactic tasks, on the word structure subtest (Table 5D) girls scored significantly higher than boys. In addition, non-verbal WM and the degree of non-nativeness (*Average quality NNS input*) were also significant predictors of children's scores on this subtest. The latter two variables were also significant predictors of children's scores on the sentence comprehension subtest (Table 5E), as was age.

Family constellation (i.e., *No NS parent vs. NS parent*) was not a significant predictor of children's scores for any of the five tasks. Furthermore, replacing *Average quality NNS input* with *Weighted quality NNS input* did not change the results, except for sentence comprehension, where the weighted measure was not significant; in other words, degree of non-nativeness remained a significant predictor for the PPVT and the word structure subtest irrespective of whether this was based on a simple average of all non-native input-providers at home, or whether the contribution of input-providers was weighted according to how much time they spent with the child.

To summarise, these results show that the amount of Dutch language input was a significant predictor of both productive vocabulary tasks. Input quality, when operationalized as the extent to which parents undertook language and literacy activities with their children (*Input richness*), was a significant predictor of children's semantic fluency when *Input quantity* was not included in the model. Neither the proportion of input at home from native speakers nor whether

children had (at least) one or no parents providing native input predicted children's scores on any of the skills tested. The degree of non-nativeness was however significantly related to children's scores on receptive vocabulary and the two morphosyntactic tasks. Finally, age, SES and non-verbal WM were significant predictors for several of the tasks. The models explained between 30% and 40% of the variance in children's scores

Family constellation based on parental language use

For the second analysis, children were divided into three groups based on whether their parents both mostly spoke the heritage language (HL) at home ($n = 13$), one parent mostly spoke the HL and the other mostly Dutch ($n = 12$), or whether both parents mostly spoke Dutch ($n=24$), where "mostly" was operationalised as $\geq 50\%$. The characteristics of these three groups with respect to our predictors of interest are given in Table 6.

--- INSERT TABLE 6 AROUND HERE ---

The groups by definition differed in terms of *Input quantity* ($F(2,46) = 17.14, p < .001$); as the SDs in Table 6 indicate, there was however considerable variation within and overlap between the *Mostly HL* group and the *HL + Dutch* groups. Again, somewhat unsurprisingly, the three groups differed in terms of *% NS input* ($F(2,46) = 3.45, p = .040$): when both parents mostly spoke the HL, most of Dutch input at home came from non-native speakers, whereas in the other two groups (i.e., where at least one parent mostly spoke Dutch), a comparable proportion of input in Dutch came from native speakers. Post-hoc (LSD) tests confirmed that children in the *Mostly HL* group heard significantly less NS input than children in the *HL + Dutch* ($MD = -14.85\%, p = .033$) and the *Mostly Dutch* ($MD = -33.34\%, p < .001$) groups. Note, however, that once again, there was considerable variation amongst families in all groups. The degree of non-nativeness was also significantly different across groups ($F(2,41) = 3.42, p = .042$), with more proficient non-native speakers providing input in the *Mostly Dutch* families when compared with the *Mostly HL* families ($MD = 0.66, p = .012$); the degree of non-nativeness in the *HL + Dutch* families did not however differ from either of the other two groups. When the same variable was weighted for the amount of time each non-native speaker spoke Dutch to the child, the mean values per group increased, suggesting that the more proficient non-native speakers provided more input. Furthermore, for *Weighted quality NNS input*, there were no statistically significant differences between family constellation groups ($F(2,41) = 1.03, p = .177$). For *Input richness* ($F(2,46) = 9.60, p < .001$), families where both parents spoke mostly Dutch engaged in more language and literacy activities in Dutch than families where both parents mostly spoke the HL ($MD = 0.88, p < .001$); once again, the children in *HL + Dutch* families were not significantly different from either of the other two groups.

Table 7 provides the descriptive statistics for children's scores on the five Dutch language tasks for the three family constellation groups based on parental language use.

--- INSERT TABLE 7 ABOUT HERE ---

The most parsimonious regression models for each task are given in Table 8.

--- INSERT TABLE 8 AROUND HERE ---

The results for the three vocabulary tasks were as follows. For the PPVT (Table 8A), SES and family constellation were the only significant predictors. Children of university-educated mothers scored significantly higher than children of mothers educated to secondary level only. Children in the *Mostly HL* group scored significantly lower than the children in the other two groups, who did not significantly differ from each other. Re-running the analysis such that the contrast between the *Mostly HL* and the *HL + Dutch* group was included confirmed that there was a significant difference between these two groups (see Table B in Appendix for complete model). Variation in input quantity and quality did not account for any additional variance.

Family constellation was also a significant predictor of active vocabulary (Table 8B); once again, the children in the *Mostly HL* group scored significantly lower than the children in the other two groups, who did not significantly differ from each other, and re-running the analysis with the contrasts between *Mostly HL* and the *HL + Dutch* group confirmed that there was a significant difference between these two groups (see Table C in Appendix for complete model). In addition to family constellation, age was also a significant predictor for active vocabulary.

For the final vocabulary task, semantic fluency (Table 8C), the predictors which were added to the model once again affected the results: in the model reported in Table 6C, older children and children whose families engaged in more language and literacy activities named significantly more words than younger children and children with less rich input. Given that this is in fact the same model as in the previous set of analyses (i.e., Table 5C), the same caveat with respect to *Input Richness* holds here: as a result of multicollinearity between this variable and *Input Quantity*, *Input Quantity* was a significant predictor when *Input Richness* was omitted from the analysis (see Table A in Appendix).

For word structure (Table 8D), there was a significant but slightly different effect of Family constellation: children in the *Mostly HL* group did not significantly differ from the other two groups but there was a significant difference between these two groups, with the children in the *HL + Dutch* group scoring significantly higher than the children in the *Mostly Dutch* group. Re-running the analysis such that the contrast between the *Mostly HL* and the *HL + Dutch* group was included confirmed that the children in the *HL + Dutch* group also scored significantly higher than the children in the *Mostly NL* group (see Table D in Appendix). The quality of any NNS input was also a significant predictor (either *Average quality NNS input* or *Weighted quality NNS input*), as was non-verbal WM, with *Input richness* approaching significance ($p = .054$). Quality of NNS input and non-verbal WM were also significant predictors of children's scores for sentence comprehension (Table 8E), as was age.

To summarise, family constellation, when operationalised as parental language use, was a significant predictor of children's scores on all tasks except sentence comprehension. Variation in *Input quantity* within the three family constellation groups did not capture any additional variance. Neither did the

proportion of input from native speakers (*% NS input*). Input quality, when operationalised as degree of non-nativeness, was a significant predictor of children's scores on the two morphosyntactic tasks only. Our other measure of input quality, *Input richness*, predicted children's scores on word structure, and when *Input Quantity* was not included in the model, on semantic fluency. Finally, age, SES and non-verbal WM were significant predictors for several tasks. The various models explained between 30% and 55% of the variance in children's scores

Table 9 provides an overview of the findings from both analyses.

--- INSERT TABLE 9 AROUND HERE ---

DISCUSSION

This study investigated the relationship between the quality of young bilingual children's language input and their developing productive and receptive skills in vocabulary and morphosyntax in the majority language (Dutch). More specifically, taking children's family constellations as our starting point, we examined the extent to which three different quality-oriented properties of children's input in Dutch would predict their language outcomes whilst taking into account their age, gender, non-verbal working memory and SES, as well as overall amount of Dutch input. The three properties were i) the proportion of input from native (rather than non-native) speakers, ii) the degree of non-nativeness in the input, and iii) input richness, measured as the extent to which families engaged in language and literacy activities in Dutch. Two sets of analyses were conducted on the same five tasks, a shortened version of the PPVT, a semantic fluency task, and the active vocabulary, word structure and sentence comprehension subtests of the CELF-2. In the first analysis, family constellation was operationalised as the number of native- vs. non-native-speaker parents, and in the second in terms of parental language use, whereby both parents mostly spoke the same language (i.e., Dutch or the heritage language) or different languages.

Native vs. non-native input

After controlling for the effect of the background variables age, gender, SES and WM, and the effects of input quantity, neither the presence of a native-speaker parent nor the proportion of Dutch input from native speakers were found to predict children's performance on any of the five tasks. Taken at face value, this result seems to suggest that exposure to (some) non-native input may matter less for acquisition than often thought. Evidence from the acquisition of American Sign Language and from artificial language learning suggests that children are indeed able to overcome inconsistencies in input (Singleton & Newport, 2004; Hudson Kam & Newport, 2005).

Our findings contrast with those of Place and Hoff (2011, 2016), who found that the proportion of exposure provided by native speakers *was* significantly related to bilingual toddlers' outcomes in the majority language, English. There are a number of possible reasons for both our null finding and for the discrepancy between our study and this earlier work. First, whilst there was quite some

variation in the proficiency level of the non-native speakers in our sample (see below), their level of proficiency was on the whole quite high (cf. Table 1). This meant that the categorical distinction made between native and non-native speaker in this study was based on a relatively small difference in proficiency; the extent of this difference may also have been smaller than in the Place and Hoff studies.³

Second, there are other sources of (potentially high quality) input (e.g., preschool teachers, tv) which were not included in the calculation of % NS input. Consequently, at least some children may have been exposed to more native-speaker input than reported. Third, the operationalisation of this variable across the two studies differed. We used a parental questionnaire to elicit information about who spent time with the child, for how long, and which language(s) they used, whereas Place and Hoff (2011, 2016) made use of a language diary (De Houwer, 2011) in which parents noted down who was interacting with the child during 30-minute blocks and whether they spoke English, Spanish or both languages; only the single-language blocks were included when calculating the proportion of NS input. How exactly these differences may have impacted on our respective results is hard to say but it is clear that whilst similar, the two measures were not exactly the same.

Finally, our operationalisation of native speaker (AO < 4 years, (self-)reported nativelylike proficiency) may have been too conservative. As outlined in the *Method* section, where there was a discrepancy in the values for these two variables, we allowed (self-)reported proficiency to “trump” AO, and we used native-speaker judgements as arbiter wherever possible. For a handful of parents, however, no such native-speaker judgement data were available and hence, we essentially had to make an educated guess based on the patterns observed in the data we did have. It is possible that as a result, several parents were classified as native speakers whilst they should have in fact been classified as non-native speakers.

Rather than the *amount* of native or non-native input at home, it was the *degree* of non-nativeness which mattered. More specifically, the degree of non-nativeness in children’s input predicted their scores on receptive vocabulary and on the two productive morphosyntactic tasks (in line with Hammer et al., 2012). Following previous work (e.g., Core & Hoff, 2014 qtd. in Place & Hoff, 2016; Rowe, 2012), we speculate that this is likely due to more proficient non-native speakers providing more morphosyntactically complex and lexically diverse input. Future work examining the productive data we have from parents, reported on here in the context of the native-speaker judgements, will be able to address this question directly. Based on the present data alone, the question of whether continued exposure to low proficiency non-native input will lead to incomplete acquisition or fossilisation, as has been claimed by some (Cornips & Hulk, 2008; Driessen, van der Slik & De Bot, 2002), remains unclear.

Input richness

In addition to the two variables concerning non-native exposure, input quality was also examined in terms of input richness (i.e., the extent to which families engaged in language and literacy activities in Dutch). The finding that richer input from a range of different sources helps (second) language development is in line with

previous work (e.g., Jia & Fuse, 2007; Paradis, 2011), including research using the same parental questionnaire (Scheele et al., 2010). In particular, exposure to narratives written in the third person through, for example, shared-book reading not only provides children with more input but more concretely, to more types and tokens of the kind of verbal morphology elicited in the word structure task. It is important to note, however, that for this task, *Input richness* was only a marginally significant predictor ($p = .05$). Furthermore, for semantic fluency, a model almost comparable to the most parsimonious one contained *Input quantity* rather than *Input richness*. In other words, it is hard to know whether it is the relative amount of input in Dutch more generally or participation in language and literacy-related activities which is the decisive factor in predicting children's ability to name as many different objects as possible.

Patterns of parental language use

In our second set of analyses, family constellation was operationalised in terms of parental language use. Two different findings emerged. First, on the receptive and productive vocabulary, children with at least one (mostly) Dutch-speaking parent had higher scores than children whose parents mostly spoke the heritage language and children with one (mostly) Dutch-speaking parent and one (mostly) HL-speaking parent. The observation that having one or more parents who speak Dutch leads to better performance replicates the well-established finding that amount of exposure matters (e.g., Gathercole & Thomas, 2009; Hoff et al., 2012), although given the correlation with *Input richness*, it is impossible to say whether what matters is *how much* parents use a given language or *how* they use that language. Note that this problem is likely to hold for many studies, not just the this one.

Despite the observation that parental language use correlated (sometimes highly) with *Input richness* and *Input Quantity*, we believe analysing our data using this variable is insightful for two reasons. First, which language parent(s) mostly speak with their children is something more tangible than an overall measure of input quantity, and hence it can be more readily translated into recommendations for raising bilingual children (see below). Second, operationalising aspects of bilingual children's language experience in a range of different ways allows us to gain a better understanding of how these aspects (and the various terms used to describe them in the literature) relate to one other.

The second finding concerning parental language use was that children whose parents mostly spoke different languages (i.e., children in the *HL + Dutch* group) had significantly better scores on the word structure task than children whose parents mostly spoke the same language (i.e., children in the *Mostly HL* and *Mostly Dutch* groups). This finding cannot be accounted for in terms of amount of input. One possible explanation is that simultaneous exposure to both languages at home sensitizes children to the differences between the two languages and this, in turn, may be beneficial to learning (Kuo & Anderson, 2010).

The observation that – for word structure and sentence comprehension at least – the degree of any non-native input accounted for variance above and beyond that captured by parental language use emphasises the importance of including not only parents but other conversational partners at home, most

notably siblings, who have been shown to have both a direct and indirect impact on toddlers' developing language skills (Bridges & Hoff, 2014).

Input quantity

In the present study, input quantity was operationalised as the proportion of Dutch spoken to the child both inside and outside the home. Input quantity was found to predict bilingual children's active vocabulary scores, in line with earlier research showing input effects in this domain (e.g., Gathercole & Thomas, 2009; Hoff et al., 2012; Thordardottir, 2011). Furthermore, for receptive vocabulary this effect held after controlling for SES, confirming earlier research showing that input effects on bilingual vocabulary development exist independently of differences related to SES (Buac, Gross, & Kaushanskaya, 2014; Scheele et al. 2010).

It is important to note that the findings reported here as effects of input quantity could, as in many earlier studies, equally be effects of children's output (i.e., language use). The very nature of bilingual language interaction means that when bilingual children use one of their two languages more frequently, they likely elicit more input in that language; similarly, hearing more input in a given language may contribute to better proficiency and subsequently more output in that language (Pearson, 2007). In short, children's patterns of language input and language use, whilst not identical, are closely related to each other. Interestingly, input quantity was a significant predictor of children's vocabulary on the two productive tasks only. If this finding is in fact indicative of an effect of output rather than input, it could either reflect a modality-specific relationship between language use and expressive skills, or a more general influence of language use, affecting expressive skills more readily than receptive skills because the former are harder to achieve, as argued in a recent study by Ribot, Hoff and Burrige (2018). As previous research has shown, however, it is possible that input and output may relate to bilingual children's language outcomes differently (e.g., Bohman, Bedore, Pena, Mendez-Perez & Gillam, 2010; Ribot et al., 2018; Unsworth, 2015; Unsworth, Chondrogianni & Skarabela, 2018). Because of the multicollinearity between these two variables in the present dataset, it was not possible to investigate this possibility here.

Input quantity, as indexed by a general measure incorporating input inside and outside the home, was not related to children's performance on the two morphosyntactic tasks. This contrasts with previous research, where such a relationship has been observed (e.g., Chondrogianni & Marinis, 2011; Place & Hoff, 2016). This may be because the tasks were not sensitive enough to detect the relevant effects. This may hold for the sentence comprehension subtest of the CELF, which may be more sensitive to children's cognitive than their linguistic skills (see below), but it seems unlikely for the word structure subtest given that it targets many of the same linguistic structures (e.g., subject-verb agreement, adjectival inflection and noun plurals) as the tasks used in earlier studies where input effects *were* observed (e.g., Paradis et al., 2011; Thomas, Williams, Jones, Davies & Binks, 2014).

A second explanation for the lack of input quantity effects in the morphosyntactic domain is that many of the children in the present study were beyond the relevant input threshold, that is, their input in Dutch was enough for

variation to no longer matter. A number of studies have shown that once more than half their language exposure is in one language, bilingual children are likely to score as well as monolingual peers of that language (Bedore et al., 2012; Hoff et al., 2012). Whilst there are some differences between these studies in terms of the exact percentage at which bilingual-monolingual differences disappear, all report figures around the 50% to 70% mark. On average, the children in our sample were exposed to Dutch for 65% of the time ($SD = 22\%$), which means that input quantity, may for many children have nonetheless been at a level at which it is less likely to predict morphosyntactic development. At the same time, however, such an explanation does not square with the findings for vocabulary, given that the majority of the children in our sample would be beyond the relevant threshold for vocabulary (e.g., Thordardottir, 2011) and yet a significant relation was observed there.

The contribution of working memory and other background variables to language outcomes

One of the other factors which emerged as an equally and sometimes more important predictor of children's scores than input quality and quantity, was non-verbal WM. The observation that non-verbal WM is related to bilingual children's morphosyntactic development is in line with a recent study by Gangopadhyay et al. (2016), who found that non-verbal WM predicted bilingual children's ability to detect morphosyntactic violations whereas no such relationship was observed for the monolinguals. Whilst the children in that study were older (8 to 10 years old) than the children in the present study, and the task (grammaticality judgement) was different from the one used here, the authors' conclusion that limited proficiency may result in increased reliance on domain-general WM skills fits nicely with the findings from the present study (see also De Cat, 2018).

Most studies addressing the role of WM in bilingual language development or including WM as a covariate in their analysis focus on *verbal* rather than non-verbal WM, and whilst the results are mixed, there is some evidence to suggest that verbal WM is a predictor of children's sentence comprehension abilities (e.g., Engel de Abreu & Gathercole, 2012; McDonald, 2008; Verhagen & Leseman, 2016; see Kidd, 2013 for review). Interestingly, it has been suggested that the WM task used here, the Kaufman Hand Movements test, in part relies on verbal encoding strategies (Frencham, Fox & Mayberry, 2003) and as such, its relation with language outcomes observed here may to some extent reflect an effect of verbal rather than non-verbal WM. More specifically, when performing this task, adults have been found to create language-based labels to describe each hand movement and subsequently use these to recall the correct sequence (Frencham, et al., 2003). Whether children, especially those as young as the participants in the present study, make use of a similar strategy remains unknown.

In terms of other background variables, SES as indexed by maternal education was found to predict children's scores, although only on receptive vocabulary. This is in line with many earlier studies demonstrating an effect of SES on bilingual children's language development for vocabulary (e.g., Cobo-Lewis, Pearson, Eilers & Umbel, 2002; Hoff, 2003). Similarly, the effect of gender observed here for the CELF word structure test, albeit only in the first analysis, is in line with other studies which have observed more advanced language skills in girls than

boys (e.g., Place & Hoff, 2016), although it remains unclear why this effect should be found for this task only.

Implications

The findings of the present study provide further evidence that variation in input quantity and quality matters in bilingual language acquisition. More specifically, the extent to which simultaneous and early sequential bilingual children are exposed to input from non-native speakers impacts on their developing language skills. The present study provides new evidence that what matters is not necessarily the amount of non-native input relative to native input, but the degree of non-nativeness. Place and Hoff (2016, p. 17) note in their discussion that “the finding is not that non-native input is harmful; the finding is that it is less beneficial than native input”. Our findings are consonant with this claim. At the same time, our findings also show that non-native input from lower proficiency speakers is less beneficial than non-native input from higher proficiency speakers.

One implication of this finding is that the advice regularly given to immigrant parents that they should speak Dutch to their children may not be good advice. In fact, when parents do not speak Dutch well, our findings suggest that they can better not speak Dutch to the children but rather seek out opportunities for their children to interact with either more proficient non-native speakers or native speakers. Not only will investing valuable “language time” as a low proficiency non-native speaker likely not have the desired effect on a child’s development of the majority language, several studies have shown that it is likely to have a negative impact on the development of the heritage language (De Houwer, 2007; Paradis & Nicoladis, 2007).

Limitations

There are a number of limitations to the present study. First, input quality was indexed using self-report only. Whilst this is by no means uncommon in the field, and available data from a film retell task from a subset of the non-native parents suggested self-report was a valid means of operationalizing proficiency level, including more objective measures in future analyses may allow for a more accurate evaluation of the impact of this variable. Second, the sample was small given the heterogeneity of the group, in particular the subgroups in the family constellation analyses, and this likely reduced the study’s power. Third, the heterogeneous nature of the sample in terms of home languages meant that it was not possible to account for any effects of language transfer.

CONCLUSION

Bilingual children’s language experience varies in a multitude of ways, not only in how much they hear (or use) their two languages but also in terms of the quality of this language input. The results of the present study show that the impact of this variation on bilingual children’s (rate of) language development is multifaceted in that not all linguistic domains are affected similarly, background variables such as SES and WM are sometimes equally good or better predictors of patterns of behaviour, and the effects of input quantity and quality are

intertwined. Which specific properties of non-native input are key in predicting bilingual children's outcomes and whether some are more important than others has yet to be determined. The findings of the present study suggest that the impact of non-native input in bilingual children's language is a matter of degree.

ACKNOWLEDGEMENTS

This research was supported by funding from the Education and Learning Sciences research impulse at Utrecht University, where all authors were previously employed. We are grateful to all participating preschools and their teachers, the children and their families, and to the various research assistants who helped with data collection and coding. Finally, we would like to acknowledge the constructive and informative comments provided by the reviewers.

NOTES

¹ As a result of allowing (self-)reported proficiency to “trump” AO in our operationalization of non-native proficiency, there were five children in the sample who had two NS parents; all of these parents were bilingual.

² The description of “quite fluent” used in the parental questionnaire was as follows: “speaks quite fluently and in addition to simple sentences sometimes uses more difficult/longer sentences, does not need to search very long for words. Can, for example, give someone directions”.

³ In Place and Hoff (2011, 2016), parents decided for themselves who counted as (non-)native when completing the language diary (E. Hoff, p.c., 25 January, 2019).

REFERENCES

- Ardila, A., Ostrosky-Solís, F., & Bernal, B. (2006). Cognitive testing toward the future: The example of Semantic Verbal Fluency (ANIMALS). *International Journal of Psychology, 41*, 324-332. doi: 10.1080/00207590500345542
- Bail, A., Morini, G., & Newman, R. S. (2015). Look at the gato! Code-switching in speech to toddlers. *Journal of Child Language, 42*, 1073-101. doi: 10.1017/S0305000914000695
- Baayen, R.H. (2008). *Analysing linguistic data: A practical introduction to statistics*. Cambridge: C.U.P.
- Baguley, T. S. (2012). *Serious stats: A guide to advanced statistics for the behavioral sciences*. New York: Palgrave Macmillan.
- Bates, D., Maechler, M., Bolker, B., Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48. doi: 10.18637/jss.v067.i01
- Bedore, L. M., Peña, E. D., Summers, C. L., Boerger, K. M., Resendiz, M. D., Greene, K., & Gillam, R. B. (2012). The measure matters: Language dominance profiles across measures in Spanish-English bilingual children. *Bilingualism: Language and Cognition, 15*, 616-629. doi: 10.1017/S1366728912000090
- Bialystok, E., Craik, I. F. M., & Luk, G. (2008). Lexical access in bilinguals: Effects of vocabulary size and executive control. *Journal of Neurolinguistics, 21*, 522-538. <https://doi.org/10.1016/j.jneuroling.2007.07.001>
- Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition, 13*, 525-531. doi: [10.1017/S1366728909990423](https://doi.org/10.1017/S1366728909990423)
- Blom, E., & Baayen, H. R. (2013). The impact of verb form, sentence position, home language, and second language proficiency on subject-verb agreement in child second language Dutch. *Applied Psycholinguistics, 34*, 777-811. <https://doi.org/10.1017/S0142716412000021>
- Bohman, T. M., Bedore, L. M., Pena, E. D., Mendez-Perez, A., & Gillam, R. B. (2010). What you hear and what you say: Language performance in Spanish-English bilinguals. *International Journal of Bilingual Education and Bilingualism, 13*, 325-344. doi: [10.1080/13670050903342019](https://doi.org/10.1080/13670050903342019)
- Bowers, E. P., & Vasilyeva, M. (2011). The relation between teacher input and lexical growth of preschoolers. *Applied Psycholinguistics, 32*, 221-241. doi:10.1017/S0142716410000354
- Bridges, K., & Hoff, E. (2014). Older sibling influences on the language environment and language development of toddlers in bilingual homes. *Applied Psycholinguistics, 35*, 225-241. doi:10.1017/S0142716412000379
- Buac, M., Gross, M. & Kaushanskaya, M. (2014). The role of primary caregiver vocabulary knowledge in the development of bilingual children's vocabulary skills. *Journal of Speech Language and Hearing Research, 57*(5), 1804-16. doi: 10.1044/2014_JSLHR-L-13-0055
- Byers-Heinlein, K. (2013). Parental language mixing: Its measurement and the relation of mixed input to young bilingual children's vocabulary size. *Bilingualism: Language and Cognition, 16*, 32-48.
- Chondrogianni, V., & Marinis, T. (2011). Differential effects of internal and external factors on the development of vocabulary, tense morphology and morpho-

- syntax in successive bilingual children. *Linguistic Approaches to Bilingualism*, 1, 318-342. <https://doi.org/10.1075/lab.1.3.05cho>
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6, 284-290.
- Cobo-Lewis, A. B., Pearson, B., Eilers, R. E., & Umbel, V. C. (2002). Effects of bilingualism and bilingual education on oral and written Spanish skills: A multifactor study of standardized test outcomes. In D. K. Oller & R. E. Eilers (Eds.), *Language and Literacy in Bilingual Children* (pp. 64-97). Clevedon: Multilingual matters.
- Cornips, L., & Hulk, A. (2008). Factors of success and failure in the acquisition of grammatical gender in Dutch. *Second Language Research*, 28, 267-296.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49, 222-251.
- Davis, M.J. (2010). Contrast coding in multiple regression analysis: Strengths, weaknesses, and utility of popular coding structures. *Journal of Data Science*, 8, 61-73.
- De Cat, C. (2018). What does the CELF-IV sentence structure test index in bilinguals? Poster presented at the *Child Language Symposium*, University of Reading, June, 2018.
- De Houwer, A. (2007). Parental language input patterns and children's bilingual use. *Applied Psycholinguistics*, 28, 411-424.
- De Houwer, A. (2011). Language input environments and language development in bilingual acquisition. *Applied Linguistics Review*, 2, 221-240. doi:10.1515/9783110239331.221
- De Houwer, A. (2014). The absolute frequency of maternal input to bilingual and monolingual children: A first comparison. In T. Grüter, & J. Paradis (Eds.), *Input and experience in bilingual development* (pp. 37-58). Amsterdam: John Benjamins.
- Dimroth, C., Andorno, C., Benazzo, S., & Verhagen, J. (2010). Given claims about new topics. How Romance and Germanic speakers link changed and maintained information in narrative discourse. *Journal of Pragmatics*, 42, 3328-3344. <https://doi.org/10.1016/j.pragma.2010.05.009>
- Driessen, G., van der Slik, F., & De Bot, K. (2002). Home language and language proficiency: A large-scale longitudinal study in Dutch primary schools. *Journal of Multilingual and Multicultural Development*, 3, 175-194. doi: 10.1080/01434630208666464
- Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test (PPVT-4)* (Fourth Edition ed.). Minneapolis, USA: Pearson.
- Engel de Abreu, P. M. J., & Gathercole, S. E. (2012). Executive and phonological processes in second-language acquisition. *Journal of Educational Psychology*, 104, 974-986. doi:10.1037/a0028390
- Fernald, A. (2006). When infants hear two languages: Interpreting research on early speech perception by bilingual children. In E. Hoff, & P. McCardle (Eds.), *Childhood Bilingualism. Research on Infancy through School Age*. (pp. 19-29). Clevedon, UK: Multilingual Matters.
- Fishman, J. A., Cooper, R. L., & Ma, R. (1971). Bilingualism in the barrio. *Language Science Monographs*, 7. Bloomington: Indiana University.
- Frencham, K. A. R., Fox, A. M., & Mayberry, M. T. (2003). The hand movement test

- as a tool in neuropsychological assessment: interpretation within a working memory theoretical framework. *Journal of the International Neuropsychological Society*, 9, 633-641. doi: [10.1017/S1355617703940033](https://doi.org/10.1017/S1355617703940033)
- Gangopadhyay, I., Davidson, M.M., Weismer, S.E. & Kaushanskaya, M. (2016). The role of nonverbal working memory in morphosyntactic processing by school-aged monolingual and bilingual children. *Journal of Experimental Child Psychology*, 142, 171-194. doi: 10.1016/j.jecp.2015.09.025
- Gathercole, V. C. M., & Thomas, E. M. (2009). Bilingual first-language development: Dominant language takeover, threatened minority language take up. *Bilingualism: Language and Cognition*, 12, 213-237. <https://doi.org/10.1017/S1366728909004015>
- Gollan, T., Starr, J., & Ferreira, V. S. (2015). More than use it or lose it: The number-of-speakers effect on heritage language proficiency. *Psychonomic Bulletin & Review*, 22, 147- 155.
- Grüter, T., Hurtado, N., Marchman, V., & Fernald, A. (2014). Language exposure and online processing efficiency: Relative versus absolute measures. In T. Grüter, & J. Paradis (Eds.), *Input and experience in bilingual development*. Amsterdam: John Benjamins.
- Gutiérrez-Clellen, V. F., & Kreiter, J. (2003). Understanding child bilingual acquisition using parent and teacher reports. *Applied Psycholinguistics*, 24, 267-288. doi:10.1017/S0142716403000158
- Hammer, C. S., Komaroff, E., Rodriguez, B. L., Lopez, L. M., Scarpino, S. E., & Goldstein, B. (2012). Predicting Spanish-English bilingual children's abilities. *Journal of Speech, Language and Hearing Research*, 55, 1251-1264. doi: [10.1044/1092-4388\(2012/11-0016\)](https://doi.org/10.1044/1092-4388(2012/11-0016))
- Hart, B., & Risley, T. (1995). *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore: Brookes.
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, 26, 55-88. <https://doi.org/10.1016/j.dr.2005.11.002>
- Hoff, E., Core, C., Place, S., Rumiche, R., Senior, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, 39, 1-27. doi:[10.1017/S0305000910000759](https://doi.org/10.1017/S0305000910000759)
- Hudson Kam C.L. & Newport E.L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, 1, 151-95. http://dx.doi.org/10.1207/s15473341ld0102_3
- Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology*, 45, 337-374.
- Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L.V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, 61, 343-364. doi: 10.1016/j.cogpsych.2010.08.002
- Hyltenstam, K. & Abrahamsson, N. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition*, 30, 481-509. doi:10+10170S027226310808073X
- Jia, G., & Fuse, A. (2007). Acquisition of English grammatical morphology by native mandarin speaking children and adolescents: Age-related differences. *Journal of Speech, Language and Hearing Research*, 50, 1280-1299. doi:

[10.1044/1092-4388\(2007/090\)](https://doi.org/10.1044/1092-4388(2007/090))

- Kaufman, A.S. & Kaufman, N.L. (1983). *Kaufman Assessment Battery for Children*. Pearson.
- Kidd, E. (2013). The role of verbal working memory in children's sentence comprehension: a critical review. *Topics in Language Disorders, 33*, 208-223. doi: 10.1097/TLD.0b013e31829d623e
- King, K.A. & Fogle, L.W. (2017). Family language policy. In T. McCarty & S May (Eds.) *Language Policy and Political Issues in Education. Encyclopedia of Language and Education* (3rd ed.) (pp. 315-327). Springer, Cham.
- Kuo, L.-J., & Anderson, R. C. (2010). Beyond cross-language transfer: Reconceptualizing the impact of early bilingualism on phonological awareness. *Scientific Studies of Reading, 14*, 365-385. <http://doi.org/10.1080/10888431003623470>
- Mayo, A. Y., & Leseman, P. P. M. (2006). *Dagelijkse Informele Educatie: Gezinsvragenlijst* [Daily Informal Education: Family Questionnaire]. Utrecht: Langeveld Instituut.
- McDonald, J. L. (2008). Grammaticality judgments in children: The role of age, working memory, and phonological ability. *Journal of Child Language, 35*, 247-268. doi: [10.1017/S0305000907008367](https://doi.org/10.1017/S0305000907008367)
- McLaughlin, B. (1978). *Second Language Acquisition in Childhood*. Hillsdale, N.J.: Erlbaum.
- Meisel, J. M. (2009). Second language acquisition in early childhood. *Zeitschrift Für Sprachwissenschaft, 28*, 5-34.
- Oller, D. K., Pearson, B. Z., & Cobo-Lewis, A. B. (2007). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics, 28*, 191-230. doi: [10.1017/S0142716407070117](https://doi.org/10.1017/S0142716407070117)
- Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism, 1*, 213-237. <https://doi.org/10.1075/lab.1.3.01par>
- Paradis, J., Emmerzael, K., & Sorenson Duncan, T. (2010). Assessment of English language learners: Using parent report on first language development. *Journal of Communication Disorders, 43*, 474-497.
- Paradis, J., & Nicoladis, E. (2007). The influence of dominance and sociolinguistic context on bilingual preschoolers' language choice. *International Journal of Bilingual Education and Bilingualism, 10*, 277-297. <https://doi.org/10.2167/beb444.0>
- Pearson, B. Z. (2007). Social factors in childhood bilingualism in the United States. *Applied Psycholinguistics, 28*, 399-410. doi:10.1017/S014271640707021X
- Place, S., & Hoff, E. (2011). Properties of dual language exposure that influence two-year-olds' bilingual proficiency. *Child Development, 82*, 1834-1849. doi: 10.1111/j.1467-8624.2011.01660.x
- Place, S. & Hoff, E. (2016). Effects and noneffects of input in bilingual environments on dual language skills in 2 ½-year-olds. *Bilingualism: Language and Cognition, 16*, 1023-101. doi:10.1017/S1366728915000322
- R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ribot, K. M., Hoff, E., & Burrige, A. (2018). Language use contributes to expressive language growth: Evidence from bilingual children. *Child Development, 89*,

- 929-940. <https://doi.org/10.1111/cdev.12770>
- Rowe, M. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development, 83*, 1762-1774. doi:10.1111/j.1467-8624.2012.01805.x
- 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*, 117-140. <https://doi.org/10.1017/S0142716409990191>
- Singleton, J. L. & Newport, E. L. (2004). When learners surpass their models: the acquisition of American Sign Language from inconsistent input. *Cognitive Psychology, 49*, 370-407. doi: [10.1016/j.cogpsych.2004.05.001](https://doi.org/10.1016/j.cogpsych.2004.05.001)
- Slot, P. L. (2014). *Early childhood education and care in the Netherlands: Quality, curriculum, and relations with child development*. PhD dissertation: Utrecht University.
- Thomas, E. M., Williams, N., Jones, L. A., Davies, S., & Binks, H. (2014). Acquiring complex structures under minority language conditions: Bilingual acquisition of plural morphology in Welsh. *Bilingualism: Language and Cognition, 17*, 478-494. doi:10.1017/S1366728913000497
- Thordardottir, E. (2011). The relationship between bilingual exposure and vocabulary development. *International Journal of Bilingualism, 15*, 426-445. <https://doi.org/10.1177/1367006911403202>
- Unsworth, S. (2013). Assessing the role of amount and timing of exposure in simultaneous bilingual development and ultimate attainment. *Bilingualism: Language and Cognition, 16*, 86-110. doi: 10.1017/S1366728912000284
- Unsworth, S. (2015). Amount of exposure as a proxy for dominance in bilingual language acquisition. In C. Silva-Corvalan & J. Treffers-Daller (Eds.), *Language dominance in bilinguals: Issues of measurement and operationalization* (pp. 156-173). Cambridge: Cambridge University Press.
- Unsworth, S. (2016). Quantity and quality of language input in bilingual language development. In E. Nicoladis & S. Montanari (Eds.), *Lifespan perspectives on bilingualism* (pp. 136-196). Mouton de Gruyter/APA.
- Unsworth, S. (2018). Quantifying experience in heritage language development. In M. Schmid & B. Köpcke (Eds.) (forthcoming), *The Oxford handbook of first language attrition*. Oxford: Oxford University Press.
- Unsworth, S., Chondrogianni, V. & Skarabela, B. (2018). Experiential measures can be used as a proxy for language dominance in bilingual language acquisition research. *Frontiers in Psychology* <https://doi.org/10.3389/fpsyg.2018.01809>
- Verhagen, J., & Leseman, P.P.M. (2016). How do verbal short-term memory and working memory relate to the acquisition of vocabulary and grammar? A comparison between first and second language learners. *Journal of Experimental Child Psychology, 141*, 65-82. doi: 10.1016/j.jecp.2015.06.015
- Verhagen, J., de Bree, E.H., Mulder, H., & Leseman, P.P.M. (2017). Effects of vocabulary and phonotactic probability on two-year-olds' nonword repetition. *Journal of Psycholinguistic Research, 46*, 507-524.
- Verhagen, J., Boom, J., Mulder, H., de Bree, E.H., & Leseman, P.P.M. (in press). Reciprocal relationships between nonword repetition and vocabulary during the preschool years. *Developmental Psychology*.

Wigg, E.H., Secord, W.A., Semel, E. & de Jong, J. (2012). *CELF preschool-2-NL: Clinical evaluation of language fundamentals – preschool, Dutch version*. Amsterdam: Pearson.

TABLES AND FIGURES

Table 1. *Descriptive statistics for experiential variables* (all children, n = 50)

Experiential variable	<i>M</i>	<i>SD</i>
% weekly input in Dutch [<i>Input quantity</i>]	65.28%	22.96%
% Dutch input at home from mother	56.50%	29.12%
% Dutch spoken by child to mother at home	70.80%	30.91%
% Dutch input at home from father	52.50%	36.44%
% Dutch spoken by child to father at home	65.10%	37.91%
Cumulative length of exposure	1.93 years	0.99 years
Average nativeness of Dutch at home (0 = no fluency, 5 = native fluency)	3.82	0.92
Proportion of input at home from native speakers [% NS input]	49.12%	33.65%
Degree of non-native input at home (0 = no fluency, 4 = very fluent) [Average quality NNS input]	2.96	0.76
Language and literacy practices in Dutch (0 = never, 5 = daily) [<i>Input richness</i>]	1.56	0.70

Table 2. *Children's (raw) scores on language tasks and working memory task: All children*

Task	<i>n</i>	<i>M (SD)</i>
PPVT (mean accuracy, %)	46	52.54 (18.69)
CELF active vocab (max. 20)	49	11.51 (7.97)
Semantic fluency (total)	48	5.60 (4.61)
CELF word structure (max. 23)	50	6.82 (4.81)
CELF sentence comprehension (max. 22)	50	7.94 (4.07)
Non-verbal working memory (max. 10)	48	2.81 (2.51)

Table 3. *Descriptive statistics for predictors of interest for family constellation groups based on parents' native speaker status*

Experiential variable	No NS parent (<i>n</i> = 20)		At least one NS parent (<i>n</i> = 30)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
% weekly input in Dutch [<i>Input quality</i>]	59.96%	25.77%	68.83%	18.63%
Proportion of input at home from native speakers [%NS input]	30.84%	30.62%	61.31%	30.26%
Degree of non-native input at home (0 = no fluency, 4 = very fluent)				
[Average quality NNS input]	2.17	0.67	3.15	0.74
[Weighted quality NNS input]	2.84	0.72	3.30	0.74
Language and literacy practices in Dutch (0 = never, 5 = daily) [<i>Input richness</i>]	1.47	0.72	1.63	0.70

Table 4. *Children's (raw) scores on language tasks and working memory task:
Family constellation groups based on parents' native speaker status*

Task	No NS parent (<i>n</i> = 20)		At least one NS parent (<i>n</i> = 30)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
PPVT (mean accuracy, %)	45.71	16.50	57.34	18.92
CELF active vocab (max. 20)	9.84	7.62	12.57	8.14
Semantic fluency (total)	5.12	4.18	5.93	4.91
CELF word structure (max. 23)	5.65	3.59	7.60	5.39
CELF sentence comprehension (max. 22)	7.25	3.70	8.40	4.30

Table 5. Predictors of total scores identified through a multiple regression analysis in which family constellation group is based on parents' native speaker status

<i>A) Predictors of total scores on (shortened) PPVT</i>				
	Coefficient	Std. Error	<i>t</i> -value	<i>p</i>
Intercept	2.48	3.16	0.79	.437
SES: primary vs. secondary and university	-14.46	7.50	-1.93	.062
SES: secondary vs. university	19.30	7.18	2.69	.011
Average quality NNS input	9.83	3.71	2.65	.012
$R^2 = .36$, adjusted $R^2 = .30$, $F(3,35) = 5.83$, $p < .01$				
<i>B) Predictors of total scores on CELF active vocabulary subtest</i>				
	Coefficient	Std. Error	<i>t</i> -value	<i>p</i>
Intercept	-0.14	0.94	-0.15	.880
Age (months)	0.88	0.19	4.69	<.001
Input Quantity	10.44	4.44	2.35	.023
$R^2 = .35$, adjusted $R^2 = .32$, $F(2,46) = 12.26$, $p < .001$.				
<i>C) Predictors of total scores on semantic fluency</i>				
	Coefficient	Std. Error	<i>t</i> -value	<i>p</i>
Intercept	-0.12	0.52	-0.23	.818
Age (months)	0.52	0.10	5.06	<.001
Input richness	2.62	1.04	2.51	.016
Input quantity	0.23	3.37	0.07	.946
$R^2 = .44$, adjusted $R^2 = .40$, $F(3,44) = 11.33$, $p < .001$.				
<i>D) Predictors of total scores on CELF word structure subtest</i>				
	Coefficient	Std. Error	<i>t</i> -value	<i>p</i>
Intercept	-0.22	0.56	-0.38	.704
WM	0.89	0.27	3.25	.002
Gender	2.76	1.16	2.39	.022
Average quality NNS input	2.31	0.84	2.75	.009
$R^2 = .49$, adjusted $R^2 = .44$, $F(4,38) = 9.16$, $p < .001$.				
<i>E) Predictors of total scores on CELF sentence comprehension subtest</i>				
	Coefficient	Std. Error	<i>t</i> -value	<i>p</i>
Intercept	-0.15	0.51	-0.29	.777
WM	0.61	0.25	2.47	.018
Age (months)	0.24	0.11	2.24	.031
Average quality NNS input	1.67	0.76	2.20	.034
$R^2 = .38$, adjusted $R^2 = .33$, $F(3,39) = 7.86$, $p < .001$.				

Table 6. *Descriptive statistics for predictors of interest for family constellation groups based on patterns of parental language use*

Experiential variable	Mostly HL (<i>n</i> = 13)		HL + Dutch (<i>n</i> = 12)		Mostly Dutch (<i>n</i> = 24)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
% weekly input in Dutch [<i>Input quality</i>]	44.79%	21.02%	59.64%	15.86%	78.13%	15.86%
Proportion of input at home from native speakers [%NS input]	29.77%	32.04%	61.95%	34.35%	54.37%	31.44%
Degree of non-native input at home (0 = no fluency, 4 = very fluent)						
[Average quality NNS input]	2.55	0.66	2.98	0.48	3.21	0.82
[Weighted quality NNS input]	2.77	0.80	3.16	0.57	3.28	0.80
Language and literacy practices in Dutch (0 = never, 5 = daily) [<i>Input richness</i>]	1.02	0.56	1.41	0.77	1.90	0.51

Table 7. *Children's (raw) scores on language tasks: Family constellation groups based on patterns of parental language use*

Experiential variable	Mostly HL (<i>n</i> = 13)		HL + Dutch (<i>n</i> = 12)		Mostly Dutch (<i>n</i> = 24)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
PPVT (mean accuracy, %)	37.23	9.13	60.23	18.54	56.46	18.74
CELF active vocab (max. 20)	5.25	5.48	15.52	7.32	13.08	7.52
Semantic fluency (total)	3.25	2.77	7.17	5.10	6.26	4.72
CELF word structure (max. 23)	3.92	3.17	9.75	4.54	7.17	4.83
CELF sentence comprehension (max. 22)	5.77	3.22	9.42	4.64	8.42	3.94

Table 8. Predictors of total scores identified through a multiple regression analysis in which family constellation group is based on patterns of parental language use

<i>A) Predictors of total scores on (shortened) PPVT</i>				
	Coefficient	Std. Error	t-value	p
Intercept	54.03	2.90	18.66	< .001
SES: primary vs. secondary and university	-10.56	7.42	-1.42	.163
SES: secondary vs. university	17.68	6.26	2.83	.007
Family constellation: Mostly HL vs. HL + Dutch and Mostly Dutch	-20.79	6.22	-3.34	.002
Family constellation: HL + Dutch vs. Mostly Dutch	-0.35	5.71	-0.06	.951
$R^2 = .39$, adjusted $R^2 = .33$, $F(4,38) = 6.19$, $p < .001$.				
<i>B) Predictors of total scores on CELF active vocabulary subtest</i>				
	Coefficient	Std. Error	t-value	p
Intercept	10.9	0.94	11.59	< .001
Age (months)	0.71	0.19	3.79	< .001
Family constellation: Mostly HL vs. HL + Dutch and Mostly Dutch	-7.59	2.12	-3.58	< .001
Family constellation: HL + Dutch vs. Mostly Dutch	0.20	2.29	0.09	.931
$R^2 = .43$, adjusted $R^2 = .39$, $F(3,44) = 11.09$, $p < .001$.				
<i>C) Predictors of total scores on semantic fluency</i>				
	Coefficient	Std. Error	t-value	p
Intercept	-0.12	0.52	-0.23	.818
Age (months)	0.52	0.10	5.06	< .001
Input richness	2.62	1.04	2.51	.016
$R^2 = .44$, adjusted $R^2 = .40$, $F(3,44) = 11.33$, $p < .001$.				
<i>D) Predictors of total scores on CELF word structure subtest</i>				
	Coefficient	Std. Error	t-value	p
Intercept	6.89	0.54	12.86	< .001
WM	0.79	0.25	3.10	.004
Gender	1.99	0.11	1.21	.234
Family constellation: Mostly HL vs. HL + Dutch and Mostly Dutch	-2.09	1.35	-1.56	.129
Family constellation: HL + Dutch vs. Mostly Dutch	-3.31	1.37	-2.42	.021
Input richness	1.79	0.90	2.00	.054
Average quality NNS input	2.00	0.79	2.54	.016
$R^2 = .63$, adjusted $R^2 = .55$, $F(7,34) = 8.10$, $p < .001$.				

<i>E) Predictors of total scores on CELF sentence comprehension subtest</i>				
	Coefficient	Std. Error	<i>t</i> -value	<i>p</i>
Intercept	7.79	0.51	15.22	< .001
WM	0.61	0.25	2.47	.018
Age (months)	0.24	0.11	2.24	.031
Average quality NNS input	1.67	0.76	2.20	.034

$R^2 = .38$, adjusted $R^2 = .33$, $F(3,39) = 7.86$, $p < .001$.

Table 9. Summary of significant predictors

Task	Family constellation based on parents' native speaker status	Family constellation based on parental language strategy
PPVT	Quality NNS input SES	Mostly HL < HL + Dutch = Mostly Dutch SES
Active vocabulary	Input quantity Age	Mostly HL < HL + Dutch = Mostly Dutch Age
Semantic fluency	Input richness <i>or</i> Input quantity Age	Input richness <i>or</i> Input quantity Age
Word structure	Quality NNS input WM, Gender	HL + Dutch > Mostly Dutch HL + Dutch > Mostly HL Quality of NNS input (Input richness) WM
Sentence comprehension	Quality NNS input Age, WM	Quality NNS input Age, WM

NNS = non-native speaker; *WM* = non-verbal working memory; *SES* = socio-economic status; *HL* = heritage language

Appendix

Table A. *Predictors of total scores on semantic fluency identified through a multiple regression analysis in which family constellation groups is based on parents' native speaker status (cf. Table 5C)*

	Coefficient	Std. Error	<i>t</i> -value	<i>p</i>
Intercept	-0.11	0.55	-0.20	.842
Age (months)	0.51	0.11	4.69	< .001
Input quantity	6.03	2.59	2.33	.024

Table B. Predictors of total scores on PPVT identified through a multiple regression analysis in which family constellation groups is based on parental language use (cf. Table 8A)

	Coefficient	Std. Error	<i>t</i> -value	<i>p</i>
Intercept	1.49	2.90	0.51	.061
SES: primary vs. secondary and university	-10.56	7.42	-1.42	.163
SES: secondary vs. university	17.68	6.26	2.83	.007
Family constellation: Mostly HL and HL + Dutch vs. Mostly Dutch	10.13	4.86	2.08	.044
Family constellation: Mostly HL vs. HL + Dutch	-20.97	7.26	-2.89	.006
$R^2 = .39$, adjusted $R^2 = .33$, $F(4,38) = 6.19$, $p < .001$.				

Table C. Predictors of total scores on CELF active vocabulary subtest identified through a multiple regression analysis in which family constellation groups is based on parental language use (cf. Table 8B)

	Coefficient	Std. Error	<i>t</i> -value	<i>p</i>
Intercept	-0.58	0.94	-0.62	.540
Age (months)	0.71	0.19	3.79	< .001
Family constellation: Mostly HL and HL + Dutch vs. Mostly Dutch	3.95	1.81	2.18	.035
Family constellation: Mostly HL vs. HL + Dutch	-7.49	2.62	-2.86	.006

$R^2 = .43$, adjusted $R^2 = .39$, $F(3,44) = 11.09$, $p < .001$.

Table D. Predictors of total scores on CELF word structure subtest identified through a multiple regression analysis in which family constellation groups is based on parental language use (cf. Table 8D)

	Coefficient	Std. Error	<i>t</i> -value	<i>p</i>
Intercept	0.07	0.54	0.13	.897
WM	0.79	0.25	3.10	.004
Gender	1.99	0.11	1.21	.234
Family constellation: Mostly HL and HL + Dutch vs. Mostly Dutch	-1.43	1.20	-1.19	.241
Family constellation: Mostly HL vs. HL + Dutch	-3.75	1.53	-2.45	.020
Input richness	1.79	0.90	2.00	.054
Average quality NNS input	2.00	0.79	2.54	.016
$R^2 = .39$, adjusted $R^2 = .33$, $F(4,38) = 6.19$, $p < .001$.				