Older Sibling Influences on Bilingual Children's Dutch Language Development and the Effect of Age Difference: the Closer the Better

Research report

Elsa Opheij[1], Anna Gorter[2]

[1]Research Master *Linguistics and Communication Sciences*, Radboud University Nijmegen
[2]Master *Language Variation and Multilingualism*, Radboud University Nijmegen

Abstract

This study investigates the effect of older siblings on their younger sibling's Dutch language development and proficiency (RQ1). Moreover, this study focuses on the age difference between the older and the younger sibling to investigate whether a smaller or larger gap leads to a higher language proficiency (RQ2). Dutch language proficiency was measured by using sentence repetition task-scores (SRTs). For the first research question the participants were children growing up with Dutch and another language (i.e., German, English, Greek, Turkish, Spanish). Only children with one older sibling or no siblings were included. For the second research question a subset of these data was used (only the children with siblings) and the influence of age difference was investigated. The results showed that having a sibling leads to lower SRT-Dutch scores (RQ1). However, these results might have been impacted by various factors, like the different languages spoken by a child. In addition, a smaller age difference between the two siblings led to higher SRT-Dutch scores (RQ2). Lines for future research include a more extensive analysis of factors that could affect older sibling influences on language development, resulting in a more comprehensive overview of how the presence of an older sibling in a bilingual home setting can affect the younger sibling's societal language development.

Keywords: child bilingualism, language development, language input, older siblings

1. Introduction

Throughout the day, children receive a lot of language input, which they can use to acquire a language. This input is usually not produced by one single person but by different people, like parents, siblings, and peers. Previous research has shown that language input from different input providers affects the bilingual child's language proficiency differently (Tsinivits & Unsworth, 2020). Regarding the input provided by siblings particularly, it has been found that the input bilingual children get from their older siblings is positively related to L2 proficiency scores (Duncan & Paradis, 2020). More specifically, Duncan and Paradis (2020) found a positive relation between the relative quantity of L2 (English) input offered by an older sibling and the younger sibling's L2 abilities on a range of linguistic subdomains. Another study found that children with older siblings are more advanced in English, but children without siblings are more advanced in the minority language spoken at home (Bridges & Hoff, 2012). The authors suggest that this is because older siblings bring the school language, in this case English, into the home. However, older studies looking at monolingual families have found an effect of the presence of an older sibling on the younger sibling's language input and output in mother-child interaction (Wellen, 1985). More specifically, mothers read stories to the younger sibling and asked questions about each story either in presence or in absence of the older sibling. The presence of an older sibling led to fewer an less informative answers on questions in the tasks.

In this study, we want to explore the role of older siblings on their younger sibling's proficiency in the majority language. Notably, the younger children are from varying language backgrounds, acquiring Dutch and another language, which allows for inspection of an older sibling's role in a rich sample regarding language background.

> ***RQ 1**: To what extent do children with older siblings have a higher Dutch language proficiency than children without older siblings?*

Although previous studies have found varying results, we hypothesise that having an older sibling has a positive effect on Dutch language proficiency scores in bilingual children, in line with previous research (e.g., Duncan & Paradis, 2020; Tsinivits & Unsworth, 2020). These studies have been conducted quite recently, and focussed on bilingual rather than monolingual children. The aim of this study is to investigate whether the positive effect of an older sibling on English proficiency scores found in previous studies (e.g., Duncan & Paradis, 2020; Bridges & Hoff, 2012) is

maintained for Dutch as well as for children from various language backgrounds. Moreover, this study investigates children from a different age group compared to previous studies (e.g., Tsinivits & Unsworth, 2020), namely children in elementary school instead of toddlers.

In a study on monolingual siblings, the age difference between siblings was found not to be associated with language development, contrary to predictions (Havron et al., 2019). However, the role of age difference on language development has not been subjected to much scientific research, and the influence of age difference between bilingual siblings in particular has not been studied before. Therefore, this topic will also be addressed.

> ***RQ 2****: To what extent is Dutch language proficiency affected by age difference between siblings?*

Formulating a clear hypothesis for this research question is more challenging. Although Havron et al. (2019) did not find any relation between age difference and language development for monolinguals, several ideas can be posed to suspect a different relation for bilingual children. On the one hand, one could argue that a larger age difference means that the first-born sibling is relatively older, hence has acquired the Dutch language to a larger extent and can provide the younger sibling with better language input and feedback. Furthermore, a larger age gap results in less competition for parents' resources (Havron et al., 2019; Her et al., 2021).

On the other hand, the input and feedback received from an older sibling might be more attuned to the younger sibling's acquisition process when the siblings are closer together in age. Besides, siblings are more likely to be role models with a smaller age gap (Her et al., 2021), possibly also improving their language learning. Apart from the expectation that there is an effect of age difference, no specific hypothesis is formulated regarding the directionality of this effect. This is due to the low availability of research on this topic, so this research question is of a more exploratory nature.

2. Methods

2.1 Participants

Participants were selected based on information provided through a parental questionnaire on children's language experience (BiLEC; Unsworth et al., 2022). The BiLEC contains data from 346 children with various language backgrounds, but all

children are native speakers of Dutch and another language. For the current study, only bilingual children (raised with two languages; German-Dutch, English-Dutch, Greek-Dutch, Spanish-Dutch, or Turkish-Dutch) who had either no siblings or only one older sibling were selected. Including children with more than one sibling or with younger siblings would pose a problem. It would obscure studying the effect of age difference and explaining any potential difference between the two experimental groups, as it would be unclear what specific factors might have led to these results. Hence, this categorisation allowed for a sound comparison between the sibling groups. Only children who received at least 10% of the language input from their sibling in Dutch were included (as indicated by parental estimation of the percentage of language input provided in every language by every input provider for the BiLEC questionnaire) . In total, 250 children were excluded due to different reasons (missing SRT data for French-Dutch children, children with no siblings, children with more than one sibling, children with a younger rather than an older sibling, and children receiving less than 10% of Dutch input) leading to a total of 96 children, aged between 4;9 and 10;10 years. Table 1 provides an overview of the number of children per sibling group and language pair.

Table 1

Number of Children per Group for Each Language Pair

Language pair	**Sibling group** M_{age} = 7.49; SD_{age} = 1.67	**No-sibling group** M_{age} = 7.67; SD_{age} = 1.60
German-Dutch	18	22
English-Dutch	18	16
Greek-Dutch	6	4
Spanish-Dutch	5	4
Turkish-Dutch	2	1
Total	49	47

Children completed a sentence repetition task (SRT) and a productive vocabulary task (cross-linguistic lexical task, CLT) as measures of language proficiency. As the data from the Dutch SRT were collected for nearly all participants, which was less true for the CLT, SRT-Dutch scores were employed as a measure of Dutch language proficiency. In the sentence repetition task (LITMUS-SRep, see Marinis & Armon-Lotem, 2015, for a detailed description), children hear and repeat sentences of varying levels of complexity which are too long to retain in short-term memory. Hence, this task requires sufficient proficiency to process the sentence heard and repeat it accurately. For this task, an utterance was coded as a correct (i.e., verbatim) repetition if the child's utterance was identical to the target sentence. However, errors related to pronunciation were ignored, as long as the target word could still be recognised. Additionally, alternative forms in the spoken language variety a child speaks were coded as correct (e.g., the abbreviated form *gister* instead of *gisteren* 'yesterday' or *hun* 'their' as a subject pronoun instead of *zij*/*ze* 'they').

2.2 Design

The SRT-Dutch scores were employed as a measure of Dutch language proficiency, which is the outcome variable for both research questions. These scores were computed as the percentage of correct verbatim items (out of 30). The following variables from the BiLEC were included as predictor variables:

- **Sibling**: the two participant groups entail one group of children with no siblings and one group of children with only one older sibling;
- **Age at testing**: participants' age at testing in years and months;
- **Cumulative LoE Dutch**: cumulative length of exposure to Dutch, i.e., the average percentage of exposure to Dutch at daycare/school and home over time, considering that one year of exposure for a bilingual child is not the same as for a monolingual child;
- **Language pair**: the child's language background (either German, English, Greek, Spanish, or Turkish and Dutch);
- **Age difference**: computed by calculating the difference in age (in months) between the sibling and the child tested;

- **Sibling input % Dutch**: the percentage of input provided by the sibling in Dutch;
- **Item**: included as a random effect to account for random variance caused by particular items in the SRT.

2.3 Data Analysis

The data analysis was conducted by fitting Linear Mixed Effects Models using R (version 3.4.0; R Core Team, 2023). R-package *lme4* (Bates et al., 2015) was used to perform multiple linear mixed-effects regression analyses for SRT-Dutch scores. Modelling commenced with a base model, including only SRT-Dutch scores and sibling and age difference as a predictor for RQ1 and RQ2, respectively. Afterwards, the remaining predictor variables were added to the base model one by one based on expected relevance to the specific research question. Each two subsequent models were compared using ANOVA comparisons of the package *lmerTest* (Kuznetsova et al., 2017) to test if the addition of another covariate improved the model fit. If AIC and BIC scores were significantly lower for the novel model, it could be assumed that including this covariate significantly improved the model fit. Finally, R-package *ggplot2* (Wickham, 2016) and *effects* (Fox, 2003) were used for visualising the data.

3. Results

The descriptive results and mixed effects model results for RQ1 are shown in Table 2 and Table 3, respectively.

Children without siblings (M_{score} = 73.83; SD_{score} = 24.09) scored significantly higher on the SRT-Dutch than children with siblings (M_{score} = 69.80; SD_{score} = 23.88), β = -3.40, SE = 0.72, $p < .001$. This effect of sibling group on SRT-Dutch scores, with higher SRT-Dutch scores for children without siblings, is visualised in Figure 1. Additionally, a significant effect of children's age at testing on SRT scores was found (β = 5.18, SE = 0.28, $p < .001$), indicating that older children scored higher on the SRT than younger children. Cumulative length of exposure was also found to have a significantly positive effect on SRT scores (β = 4.12, SE = 0.27, $p < .001$), meaning that the higher the (cumulative) exposure to Dutch, the higher the SRT score. Additionally, Dutch-German (functioning as the reference level in the mixed effects model, to which the other Language pair levels are compared) children scored significantly higher than

Table 2

Descriptive Results for all Fixed Factors

Sibling	**M/SD**	**SRT-Dutch score (in %)**	Age at testing	**Cumulative LoE Dutch**	**Age difference**	**Sibling input % Dutch**
Y	M	69.80	7.49	4.41	2.37	0.72
	SD	23.88	1.67	1.58	1.18	0.28
N	M	73.83	7.67	4.04	-	-
	SD	24.09	1.62	1.55	-	-

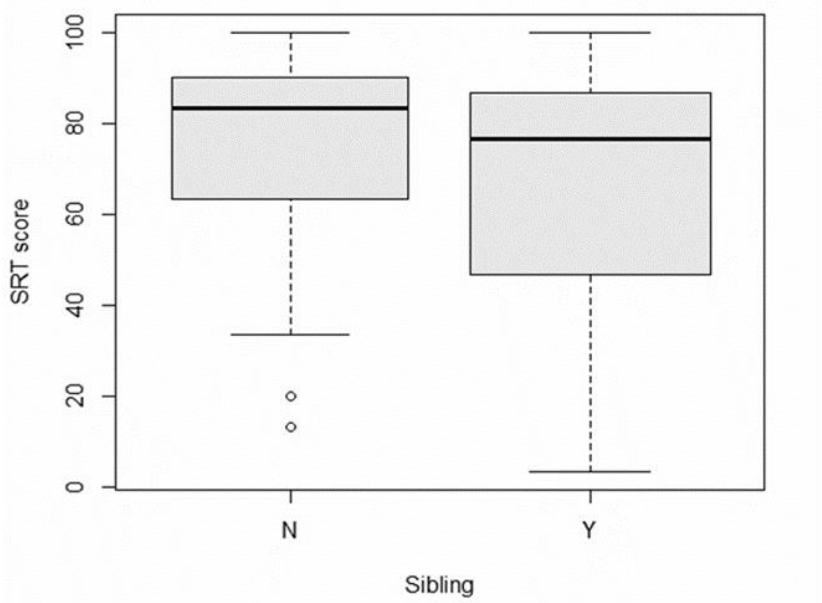
Table 3

Mixed Effects Model for RQ1

Fixed effects	***β***	***SE***	***t*-value**	***p*-value**
(Intercept)	22.26	2.10	10.59	<.001
Sibling: Yes	-3.40	0.72	-4.75	<.001
Age at testing	5.18	0.28	18.31	<.001
Cumulative LoE Dutch	4.12	0.27	15.04	<.001
Language pair: Greek-Dutch	-10.90	1.32	-8.25	<.001
Language pair: English-Dutch	-9.46	0.82	-11.60	<.001
Language pair: Spanish-Dutch	-3.86	1.36	-2.83	.0047
Language pair: Turkish-Dutch	-13.68	2.08	-6.57	<.001

Figure 1

Boxplot of SRT Scores Across the Two Sibling Groups



children in all other language pairs (Dutch-Greek: β = -10.90, *SE* = 1.32, $p < .001$; Dutch-English: β = -9.46, *SE* = 0.82, p < .001; Dutch-Spanish: β = -3.86, *SE* = 1.36, $p = .0047$; Dutch-Turkish: β = -13.68, *SE* = 2.08, $p < .001$).

The mixed effects model results for RQ2 are presented in Table 4. The results showed a significant effect of age difference on SRT scores (β = -8.30, *SE* = 0.74, $p < .001$), indicating that a smaller age gap between siblings was associated with higher proficiency scores. Several covariates showed significant results: age at testing (β = 3.42, *SE* = 0.34, $p < .001$) and cumulative exposure to Dutch (β = 7.43, *SE* = 0.42, $p < .001$) were both positively related to SRT scores, meaning that a higher age at testing and a higher exposure to Dutch led to higher scores. Including *percentage* of input provided in Dutch by the sibling as a fixed effect did not significantly improve the model fit. However, a significant interaction between age difference and *percentage* of input provided in Dutch by the sibling on SRT scores was found (β = 2.29, *SE* = 0.90, $p = .0115$), which indicates that the effect of age difference on score differed based on the *percentage* of Dutch input provided by the sibling. This interaction is visualised in Figure 2; the higher the age difference between siblings, the larger the positive effect of relatively more Dutch input given by the sibling on SRT-Dutch scores. This suggests that it mattered less how much input children receive from an older sibling when the age difference is small.

The fixed and random effects that were included in the final model per research question are presented in Table 5.

Table 4

Mixed Effects Model for RQ2

Fixed effects	***β***	***SE***	***t*-value**	***p*-value**
(Intercept)	27.14	2.26	12.00	<.001
Age difference	-8.30	0.74	-11.17	<.001
Age at testing	3.42	0.34	9.92	<.001
Cumulative LoE Dutch	7.43	0.42	17.64	<.001
Age difference:Sibling input % Dutch	2.29	0.90	2.53	0.0115

Figure 2

Plot of Interaction Between Percentage of Dutch Input Provided by the Sibling and Age Difference on SRT-Dutch Scores

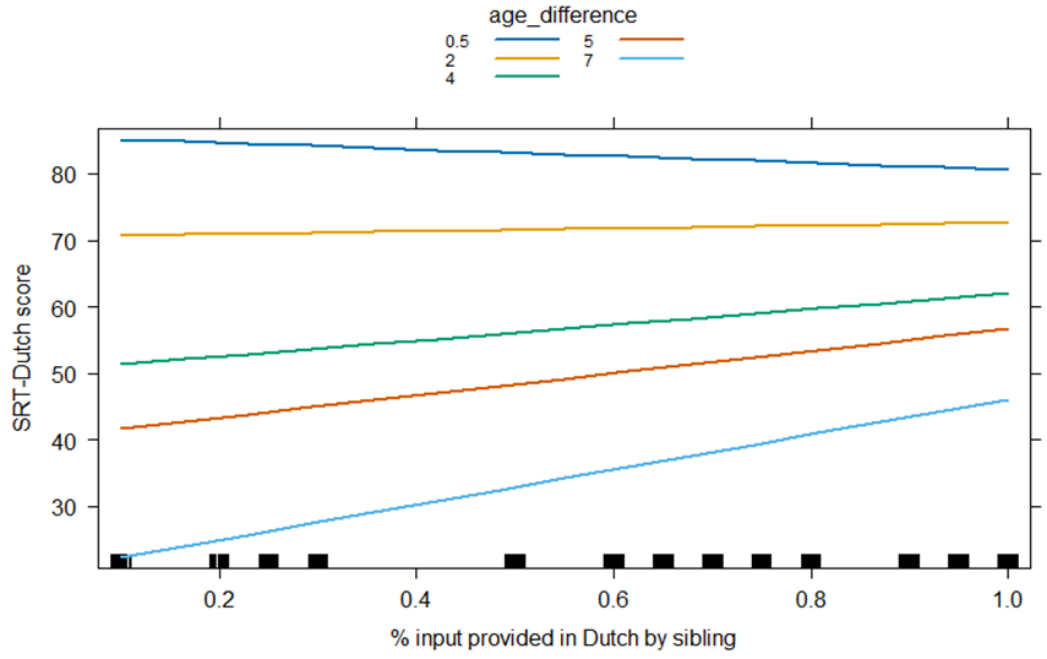


Table 5

Outcome and Predictor Variables for the Best Model per Research Question

Variables	**RQ1**	**RQ2**
Outcome	SRT-Dutch score	SRT-Dutch score
Predictor	<u>Fixed effects</u> • Sibling • Age at testing • Cumulative LoE Dutch • Language pair	<u>Fixed effects</u> • Age difference • Age at testing • Cumulative LoE Dutch • Age difference*Sibling input % Dutch
	<u>Random effects</u> • By-item	<u>Random effects</u> • By-item

4. Discussion and Conclusion

The aims of this study were twofold. Firstly, we investigated the effect of having an older sibling versus having no siblings on bilingual children's Dutch language proficiency, using performance on a sentence repetition task as a measure of language proficiency. Participants with different language backgrounds, acquiring Dutch and another language, were selected from the BiLEC dataset (Unsworth et al., 2022).

Mixed effects models showed that having a sibling led to significantly lower proficiency scores than having no siblings, which was not in line with our hypothesis based on previous studies on bilinguals (Bridges & Hoff, 2012; Duncan & Paradis, 2020). However, it is in line with Wellen's (1985) study on monolinguals. To control for a potential influence of the distribution across language pairs per group (sibling versus no-sibling), language pair was included in the analysis. Although SRT-Dutch

scores were higher for Dutch-German children than for all other language pairs, including language pair still revealed that SRT-Dutch scores for children without a sibling were significantly higher than for children with an older sibling. A possible explanation for these results might be that children with an older sibling have to compete for their caregivers' attention, which might result in less language input to the younger sibling, hence leading to lower proficiency scores (e.g., Havron et al., 2019).

Secondly, the role of age difference between siblings was studied, which was found to affect SRT-Dutch scores, unlike Havron et al.'s (2019) findings. Our results showed that a smaller age gap had a positive effect on the target children's language development, as measured by performance on a sentence repetition task.

Havron et al. (2019) expected to find an influence of age difference between monolingual siblings on the younger sibling's language development. More specifically, they hypothesised that a larger age gap would have a less detrimental effect for the younger sibling's linguistic skills, since there is less competition for parental resources (less care and supervision required at a higher age), and the older sibling's social and linguistic skills are likely to be more developed, suggesting that better linguistic input could be provided. However, they did not find such an effect in their study. A possible explanation for this finding could be that an older sibling with a smaller age gap is more likely to be a role model for the younger sibling, and input and feedback might be more attuned to the younger sibling's language learning process (Her et al., 2021).

Moreover, Havron et al. (2019) focussed on monolingual children, whereas the current study looked at bilingual children, meaning that more factors might be at play here. The different language(s) spoken in a home setting might alter the role of the older sibling's language input and feedback, as less input in the societal language from the parents could increase the importance of input from siblings. Additionally, sex of the sibling and overall exposure to the societal language (quantity and quality) might have impacted the results.

The effect of age difference was also found to be moderated by the percentage of Dutch input provided by the sibling. This suggests that the positive influence of a higher amount of input provided in Dutch by the sibling on the target child's Dutch language proficiency differed by age difference in the sense that a target child with a larger age gap with their sibling benefited more from a higher percentage of Dutch input compared to children with a smaller age difference with their sibling.

In conclusion, these findings show that having a sibling has a negative effect on the target children's Dutch language proficiency, but the effect of having an older

sibling was less detrimental for siblings with a smaller age difference, and this positive effect of age difference was larger for target children who received more Dutch input from their sibling. The older sibling might be functioning more as a role model when the age difference is smaller and input and feedback might be better aligned with the target child's language development process, both positively impacting the target child's societal language proficiency.

It should be mentioned that the exact age of the siblings was not known, as the age was only indicated in whole years and not months. This means that a sibling of, for instance, 7 years old could have had an age between 7;0 and 7;11, leading to an 11-month gap of uncertainty in the older sibling's age. It is unclear to what extent this might have had an effect on our findings. Therefore, this should be kept in mind when interpreting the results.

What was not considered in this study is the quality and the absolute amount of the sibling's input, the sibling's exact age, the sibling's sex, the language spoken at home by the parents (and the quantity and quality of any parental input provided in Dutch), and other measures of Dutch language proficiency besides sentence repetition abilities, as this was beyond the scope of the current study. Future research should address these shortcomings, allowing for a more comprehensive overview of how the presence of an older sibling in a bilingual home setting can affect the younger sibling's Dutch language development and what factors are at play here. Additionally, further studies could investigate the effect of language distance and its interaction with other variables explored in this research (e.g., sibling) to determine whether language distance moderates the effect of sibling on language proficiency in the sense that children who acquire two more distant languages have a higher need for input from a sibling than is the case with a less distant language pair. Finally, it would be interesting to see whether similar results are found when other societal languages than Dutch are studied.

In conclusion, we found that having a sibling resulted in lower Dutch proficiency scores, but that a smaller age difference between siblings did improve scores. However, future research should aim at increasing the degree to which other factors that might have influenced these results are considered, as well as showing whether results can be reproduced for different societal languages.

Acknowledgements

We would like to thank dr. Sharon Unsworth for giving us the opportunity to use the data she and her colleagues collected and for helping us throughout the process of

writing this paper during the course Child Bilingualism at the Faculty of Arts at Radboud University. Moreover, we want to thank Gerrit-Jan Kootstra, Chantal van Dijk, Figen Karaca, Elly Koutamanis, Jasmijn Bosch, Elise van Wonderen, and Joyce van Zwet, who were also involved in the data collection process. We would also like to thank the editors and two anonymous reviewers for their insightful comments on previous versions of this report.

References

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01

Bridges, K. J., & Hoff, E. (2012). Older sibling influences on the language environment and language development of toddlers in bilingual homes. *Applied Psycholinguistics, 35*(2), 225–241. https://doi.org/10.1017/s0142716412000379

Duncan, T. S., & Paradis, J. (2020). Home language environment and children's second language acquisition: The special status of input from older siblings. *Journal of Child Language*, *47*(5), 982-1005. https://doi.org/10.1017/S0305000919000977

Fox, J. (2003). Effect displays in R for generalised linear models. *Journal of statistical software*, *8*, 1-27. https://doi.org/10.18637/jss.v008.i15

Havron, N., Ramus, F., Heude, B., Forhan, A., Cristia, A., & Peyre, H. (2019). The effect of older siblings on language development as a function of age difference and sex. *Psychological Science*, *30*(9), 1333–1343. https://doi.org/10.1177/0956797619861436

Her, Y. C., Batur, Z. Z., Vergauwen, J., & Mortelmans, D. (2021). Siblings' similarities and dissimilarities: A review of theoretical perspectives and empirical insights. *Journal of Family Theory & Review, 13*(4), 447-462. https://doi.org/10.1111/jftr.12436

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of statistical software*, *82*, 1-26. https://doi.org/10.18637/jss.v082.i13

Marinis, T., & Armon-Lotem, S. (2015). Sentence repetition. *Assessing multilingual children: Disentangling bilingualism from language impairment*, 95-124.

R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org

Tsinivits, D., & Unsworth, S. (2020). The impact of older siblings on the language environment and language development of bilingual toddlers. *Applied Psycholinguistics*, *42*(2), 325-344. https://doi.org/10.1017/S0142716420000570

Unsworth, S., Kootstra, G. J., Van Dijk, C., Karaca, F., Koutamanis, E., Van Bosch, J., Van Wonderen, E., & Van Zwet, J. (2022). 2In1: Bilingualism in children [Dataset]. In *2In1 research project*.

Wellen, C. J. (1985). Effects of older siblings on the language young children hear and produce. *Journal of Speech and Hearing Disorders*, *50*(1), 84–99. https://doi.org/10.1044/jshd.5001.84

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. https://ggplot2.tidyverse.org