

[WLG]

WIENER LINGUISTISCHE GAZETTE

Language Development in Typical Bilingualism and Bilingual Developmental Language Disorder

A Longitudinal Twin Case Study

Lisa-Marie Müller

Sonderdruck aus: *Wiener Linguistische Gazette* 96 (2024): 77–118

Universität Wien · Institut für Sprachwissenschaft · 2024

Eigentümer, Herausgeber und Verleger:

Universität Wien, Institut für Sprachwissenschaft
Sensengasse 3a
1090 Wien
Österreich

Redaktion: Florian Grosser, Jonas Hassemer & Carina Lozo
Redaktioneller Beirat: Markus Pöchtrager & Stefan Schumacher
Kontakt: wlg@univie.ac.at
Homepage: <http://www.wlg.univie.ac.at>

ISSN: 2224-1876
NBN: BI,078,1063

Die *Wiener Linguistische Gazette* erscheint in loser Folge im Open-Access-Format.
Alle Ausgaben ab Nr. 72 (2005) sind online verfügbar.



Dieses Werk unterliegt der Creative-Commons-Lizenz CC BY-NC-ND 4.0
(Namensnennung – Nicht kommerziell – Keine Bearbeitungen)

Language Development in Typical Bilingualism and Bilingual Developmental Language Disorder

A Longitudinal Twin Case Study

Lisa-Maria Müller*

Wiener Linguistische Gazette (WLG)

Institut für Sprachwissenschaft

Universität Wien

Ausgabe 96 (2024): 77–118

Abstract

Dieser Artikel präsentiert die Ergebnisse einer longitudinalen Zwillingsfallstudie, in welcher der Spracherwerb eines typisch entwickelten Englisch-Polnisch zweisprachigen Buben mit dem seines Zwillingsbruders, welcher mit einer spezifischen Sprachstörung diagnostiziert wurde, verglichen wird. Der Erwerb von Lexis, Morphosyntax und narrativer Kompetenz der beiden Buben wurde über einen Zeitraum von vier Jahren, zwischen 4;8 und 8;6 Jahren erhoben. Ein Vergleich mit monolingualen Normen zeigt, dass ihre Sprachentwicklung nicht linear verlief und sogar kurzfristiger rückläufiger Entwicklung unterlag. Ergebnisse deuten außerdem darauf hin, dass die Zuverlässigkeit verschiedener Spracherhebungstests zur Unterscheidung typisch entwickelter Zweisprachigkeit und beeinträchtigter zweisprachiger Entwicklung altersabhängig sein könnte, weshalb mehr longitudinale Studien benötigt werden.

Schlagwörter: Zweisprachigkeit, Sprachentwicklung, Developmental Language Disorder, Zwillingsfallstudie

* Lisa-Maria Müller, Chartered College of Teaching, London, Great Britain, lmuller@chartered.college

1 Introduction

Neither bilingualism nor developmental language disorder (DLD) are homogenous phenomena and surrounded by theoretical and terminological controversies. DLD has recently undergone a name change from SLI (Specific Language Impairment) after extensive debate in the field (Reilly et al. 2014a, 2014b; Bishop 2017) as it became clear that the language impairment might not be as ‘specific’ as previously assumed and affects more than children’s grammatical development. While a diagnosis of SLI traditionally required a mismatch between children’s linguistic and cognitive abilities (Tomblin 1996) this is no longer required as it has been found that such a mismatch does not accurately describe the majority of the clinical population (see Reilly et al. 2014a, 2014b). This also implies that the language profiles of children with DLD are more varied than previously assumed.

Bilingualism is also inherently heterogenous and influenced by factors such as the age at which children acquired their languages, if they acquired them simultaneously or sequentially, whether one of their languages is a minority language as well as its status in society, the language combination or simply the amount of exposure children receive in each of their languages. Each of these factors can potentially differ from one bilingual to another, which makes it inherently difficult to draw conclusions about bilingualism more widely – an issue that has, for example, become particularly evident in the recent replication crisis (Bak 2016; Paap 2013) around cognitive changes in bilingualism.

Twin case studies provide a unique opportunity to investigate differences in language abilities independently of environmental differences. Twins are of the same age, they have comparable exposure to their two languages, the same socio-economic status (SES) when they grow up in the same household with the same parents of the same educational background and they are genetically very similar. The twins in this study only differ in one crucial aspect; one of them was previously diagnosed with DLD, whereas the other is typically developing. This study responds to calls for a more holistic description of bilingual DLD

by investigating the children's lexical, morphosyntactic and narrative development over the course of four years.

2 Bilingualism and Developmental Language Disorder

Similar to research on monolingual DLD, morphosyntax has been a central focus in research on bilingual DLD. Morphosyntactic differences in typically and atypically developing bilinguals have been studied in a wide range of language combinations, such as English-Spanish (Eng & O'Connor 2000; Gutiérrez-Clellen et al. 2008; Jacobson & Schwartz 2002), Swedish-Arabic (Håkansson et al. 2003), Turkish-Dutch (Blom et al. 2013), Turkish-German (Clahsen et al. 2014; Rothweiler et al. 2010, 2012), English/Russian-Hebrew (Armon-Lotem 2010, 2014), Dutch-Moroccan/Arabic/Berber (Verhoeven et al. 2011, 2012) or Dutch-Turkish (Blom et al. 2013). Globally, these studies revealed that typically developing bilingual children and their peers with DLD differed in the quantity and quality of certain errors. They report higher overall accuracy for the typically developing bilingual groups (Bi-TD) as well as higher verb variety but more overgeneralizations for all investigated languages apart from German. The bilingual groups with developmental language disorder (Bi-DLD) tend to show lower performance in both languages, while at least one language is well developed in the Bi-TD groups. However, both groups produce bare verb stems. Additionally, more overlaps in the error patterns of later language learners and Bi-DLD groups were found than between Bi-DLD and Bi-TD groups. Hebrew data of the sequential bilinguals in this study further revealed an influence of subjects' L₁ on the error patterns of Bi-TD children but substitution for less complex forms for the Bi-DLD group (Armon-Lotem 2014).

Studies comparing the narrative abilities of bilingual children with and without DLD tend to find differences in the micro – but not the macrostructure of their stories (Altman et al. 2016; Iluz-Cohen & Walters 2012).

Only recently, group studies with a more integrated description of bilingual language abilities have emerged. Verhoeven et al. (2011), for example, investigated the Dutch phonological, morphosyntactic, lexical and textual development of Bi-TD and Bi-DLD with a variety of first languages. They found Bi-TD to outperform Bi-DLD on articulation, sentence repetition and auditory discrimination at all ages and lexicon and grammar to be particularly affected by language impairment.

Verhoeven et al. (2012) investigated the Turkish and Dutch phonological, morphosyntactic, lexical and textual development of 6–7, 8–9 and 10–11 year old Turkish-Dutch bilinguals with DLD, comparing it to the results of the Bi-TD group from their previous study. All tasks showed significant effects for age group. Additionally, a language effect was observed with both groups scoring better on Turkish auditory discrimination, nonword repetition, receptive vocabulary, word definition, sentence imitation and story comprehension.

Gillam et al. (2013) investigated English vocabulary, syntax, narrative and sensitivity of EpiSLI criteria¹ (Tomblin et al. 1996) in English-Spanish five to six year old Bi-TD and Bi-DLD. According to Tomblin et al. (1996) a cut-off score of -1.25 SD on two or more composite scores is required for a diagnosis of monolingual DLD. The application of a similar cut-off score to the bilingual population in Gillam et al.'s (2013) study led to the overdiagnosis of more than half the typically developing bilingual children as language impaired. In addition, neither of the seven subtests yielded acceptable levels of specificity and sensitivity to distinguish accurately between typical and atypical bilingual language development. The authors thus took a multivariate approach whereby five composite scores of comprehension, expression, vocabulary, grammar and narrative interact to predict impairment.

Kohnert et al. (2009) and Kohnert (2009) therefore suggest that a combination of language measurements rather than the performance on a single language measure may help to better characterize (bilingual) DLD and typical development. This is precisely what this study sets out

1 Epidemiological *SLI* criteria are five composite scores representing comprehension and production in vocabulary, grammar, and narration.

to do – a longitudinal, more holistic description of language skills in one typically and one atypically developing bilingual.

3 Methodology

The aim of this study was to give a holistic overview of the twins' language development by assessing their lexical, morphosyntactic and narrative development in order to answer the following research questions:

- How do these language skills develop over time in one typically developing bilingual English-Polish child and his twin brother with developmental language disorder?
- How does the development of their two languages relate to each other and does this relationship differ in the two children?
- Which assessments were more accurate in distinguishing between the child with typical bilingual development and the one with bilingual developmental language disorder at different ages?

3.1 Case Study Research

This section will outline the motives behind the design of this study. The first motive is rooted in the inherent heterogeneity of bilingualism and DLD, which makes generalisations problematic. Furthermore, the study of the individual allows for a more holistic picture of their overall language ability and can thus provide crucial insights for the study of language. Hymes even goes as far as to assign “foundational status” to individual language profiles and to see them as a “vantage point from which to consider questions of method and theory in the study of language in general” (1979: 35).

Case studies can also have “a high degree of completeness, depth of analysis, and readability” as emphasised by Duff (2008: 43) because qualitative analyses of language data are complemented by detailed case histories, which provide essential background knowledge to accurately

interpret language data. Such an in-depth description of background factors is more complicated in group studies. Concentrating on one or few cases further makes longitudinal studies and holistic description of language development more feasible.

As a case study approach is by definition exploratory and data-driven (Duff 2008: 44), it can generate theories or hypotheses which can in turn be verified or falsified on a larger cohort. One such example from neurology is Diamond et al.'s (1985) post-mortem dissection of Einstein's brain, which revealed extensive growth of dendritic spines on neurons. This finding was later replicated in a large number of cross-sectional brain studies (in Duff 2000). Similarly, some of the most crucial insights about the organisation of cognitive systems stem from Carramazza's (1986) case studies.

In linguistics, some of the most influential observations about monolingual and bilingual language development that still stand today, also initially stemmed from case studies (e.g. Leopold 1939, 1940, 1947, 1970; Stern & Stern 1970). These meticulous, longitudinal studies of child language development were able to address previous misconceptions (e.g. that children generally pronounce the word 'mama' first) and provide in-depth descriptions of the stages involved in monolingual and bilingual child language development. Leopold's own passionate appeal for more case studies investigating bilingual language development is thus not surprising and shows how necessary they were and still are today to complement and inform quantitative studies.

As generalisability of case study results to the broader population is often questioned, it needs to be stressed that the aim of case studies is not to generalise its findings to the whole population. Case studies should be regarded as food for thought or an invitation for further research. Therefore, rather than seeking a generalisable interpretation of a phenomenon, case study research seeks to complement and challenge common interpretations and existing (quantitative) research.

3.2 Participants

Participants initially took part in a larger research study as part of the cross-European COST Action ISo8o4: Language Impairment in a Multilingual Society, for which the author worked as a research assistant. Upon discovery of their unique profile, permission was sought from the parents to study them further as part of this case study. The case history presented here was elicited using two parental questionnaires and a follow-up interview with the children's mother, conducted in Polish by the author of this paper. The first questionnaire is the Polish version of a questionnaire used by the COST ACTION ISo8o4, which is based on Paradis (2007) and investigates children's (bilingual) language development as well as their current linguistic status. The second questionnaire is based on Stromswold (2006) and inquires about a number of pre-, peri-, and postnatal factors influencing language development.

The participants are a male dizygotic English-Polish bilingual twin pair living in the UK. The younger twin (henceforth 'LIB' – Language Impaired Bilingual) was diagnosed with DLD aged 4;6 (i.e.: prior to this study) in both languages by two qualified speech and language therapists and does not meet any of the exclusionary criteria for DLD. This diagnosis was re-confirmed by the school's speech and language therapist in November 2015, when he was assessed in his then dominant language English. At the time, his performance corresponded to the 5th percentile on the CELF-4 (Clinical Evaluation of Language Fundamentals; Semel et al. 2003). The older twin brother (henceforth 'TDB'; Typically Developing Bilingual) is typically developing.

Furthermore, the twins' mother voiced concerns about his language development in the parental questionnaire, the follow-up interview as well as subsequent conversations. She indicated a substantially later age of onset for LIB and mentioned that he generally spoke less than his brother. According to her, LIB's errors were grammatical (syntactic and morphological) and stylistic in nature. He received support from teaching assistants in both his English and his Polish (Saturday) schools in addition to speech and language therapy.

The information from the questionnaire on pre-, peri-, and postnatal factors influencing language development shows that the twins were not born prematurely but LIB was born second at only 1.8 kg, as opposed to his brother, whose birth weight amounted to 3.1 kg. Subsequently, LIB spent eight days in the NICU (Neonatal Intensive Care Unit) while his brother was released from the hospital after only five days. According to Stromswold (2006) low birth weight and prolonged time spent in the NICU are risk factors for language impairment.

The two boys are early sequential bilinguals who were born in the UK to Polish parents. They mainly speak Polish at home and English in school and during most of their social interactions. Both boys uttered their first words in Polish, their main form of interaction until the age of three when they entered pre-school, but LIB did so substantially later than TDB. While English is the children's language of education, Polish has remained their family language despite the fact that both parents are balanced bilinguals and use both languages on a daily basis. Books were read to them in Polish when they were younger and the boys still prefer to watch TV in Polish. Now that they can choose their own reading material, they tend to read in English and TDB prefers stories and novels, while LIB shows a preference for short, factual texts.

The twins also regularly interact in Polish with two aunts and two uncles and with their grandparents when they come to visit or they go to see them in Poland during the summer or winter holidays. They also attend Polish school for four hours a week on Saturdays. There they learn about Polish culture and to read and write in their family language. In a regular week, the children speak English more frequently than Polish as they use the majority language at school and during their extra-curricular activities. The twins use both their languages when interacting with each other, usually adapting to their environment.

The parents' views about bilingualism are very positive. They state a broader worldview, better cognitive development and biculturalism as advantages of bilingualism. They add, however, that the latter might also result in a lack of identity. The twins' parents have both completed tertiary education, so the twins' socio-economic background can be classified as mid-high.

Finally, when asked to characterise her children, the mother stated that TDB used to be more dominant and more controlling before the boys started school and went to different classes (a choice which is common practice for twins in the school they attend). The fact that TDB was more dominant linguistically also led him to regularly take over responsibilities and to talk for his brother (cf. Leonard 1998: 175). He is reported to do so less nowadays but this is still reflected by his rather extrovert nature. Quite contrarily, his younger brother is rather introvert, more creative and sensible and capable of playing or drawing quietly on his own for a longer period of time, according to the mother's description of her two children during the interview.

3.2 Materials

Participants were tested aged 4;8, 5;6, 6;8, 7;9 and 8;6 years in a longitudinal case study (Duff 2008). Their phonological, lexical, morphological, syntactic and narrative development was tested, using a combination of standardised assessments and some of the LITMUS tests developed by the *COST Action ISo804*. One series of testing took place as part of the bi-sli PL project², whereas the remaining data collection was part of a follow-up study, which was carried out with the permission and help of the bi-sli PL team and employed the tests used and partly developed or adapted by this group of researchers in addition to additional measurements, part of which are reported elsewhere

2 The project was carried out at the Faculty of Psychology, University of Warsaw, Poland in collaboration with Institute of Psychology, Jagiellonian University, Poland. The project was supported by the Polish Ministry of Science and Higher Education/National Science Centre (Decision 809/N-COST/2010/0). Data collection, data coding and maintenance were also partly supported by Foundation for Polish Science subsidy to Zofia Wodniecka and the Polish Ministry of Science and Higher Education grant (Decision 0094/NPRH3/H12/82/2014) *Phonological and Morpho-syntactic Features of Language and Discourse of Polish Children Raised Bilingually in Migrant Communities in Great Britain*, carried out at the Faculty of Modern Languages, University of Warsaw, Poland. The project was linked to the European COST Action ISo804.

(Müller 2018). Details of the tests are given below. The children were tested separately, in a quiet room by the author of this paper who is proficient in both languages. The two languages were assessed on different days with three weeks between testing sessions in each language.

3.3.1 Productive Lexis Tests

The Second Edition of the *EVT-2: Expressive Vocabulary Test* (Williams 2007) and the *ZNO: Zadanie Nazywania Obrazków* (Haman & Smoczyńska 2010) were administered to assess the twins' productive vocabulary. In these tests, participants have to name nouns, verbs and adjectives based on colourful stimuli. Both tests have been standardised on monolingual cohorts.

3.3.2 Receptive Lexis Tests

Receptive Vocabulary was assessed using the third edition of the *BPVS-3: British Picture Vocabulary Scale* (Dunn et al. 2009) and the *OTS-R: Polish Obrazkowy Test Słownictwa* (Haman & Fronczyk 2012). In both tests, participants have to choose the right of four images based on auditory stimuli.

3.3.3 Morphology and Syntax

The *SASIT: School Age Sentence Imitation Task* (Marinis et al. 2011) and its Polish adaptation (Banasik et al. 2011) were used to assess the twin's knowledge of morphology and syntax in their two languages. In these tasks, participants hear recorded sentences through headphones and have to repeat them. Their production is scored for syntactic and morphological accuracy.

Their understanding of grammar was tested using *the TROG-2: Test of Reception of Grammar* (Bishop 2003) and its Polish translation *TROG-*

PL (Smoczyńska 2005). In this task, participants have to choose the right out of four images based on a pre-recorded sentence they hear.

3.3.4 Narrative Task

The Multilingual Assessment Instrument for Narratives (MAIN) (Gagarina et al. 2012; Kiebzak-Mandera et al. 2012; Mieszkowska et al. 2020) was used to assess their narrative abilities in both languages. Stories are elicited based on six colourful picture stimuli. This test has been widely used with multilingual populations.

4 Results

4.1 Receptive Vocabulary

LIB's and TDB's receptive vocabulary scores in percentiles are presented in Figure 1. Norms for the OTS-R, the Polish receptive vocabulary task, are only available until the age of 6;11. Therefore the scores at the last two data collection points were compared to these younger monolingual norms.

Figure 1 shows a persistent gap between TDB and LIB's receptive English vocabulary scores and continuous fluctuation in both children's receptive lexis scores for both languages, in relation to age norms. TDB and LIB's BPVS-3 scores decreased by 27 percentile points between the ages of 4;8 and 5;6, when LIB's score remained stable, while TDB's receptive vocabulary score decreased further (-13) until the age of 6;8. Between the ages of 6;8 and 7;9 LIB's and TDB's BPVS-3 scores increased in relation to age norms but more so for TDB (+26) than for LIB (+14). Finally, TDB's and LIB's receptive English lexical development showed a reverse trend, with TDB's score gaining another fourteen percentile points in relation to monolingual English norms, reaching the

82nd percentile. LIB's score, on the other hand, decreased drastically from the 28th to the 8th percentile between the ages of 7;9 and 8;6.

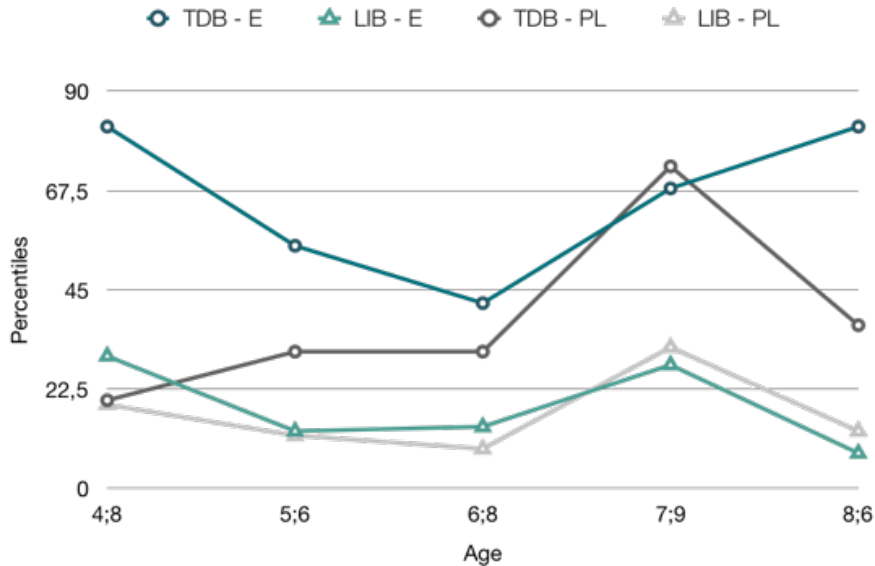


Fig. 1: Receptive Vocabulary Scores in Percentiles

Their development of Polish receptive vocabulary also fluctuated. While TDB and LIB scored roughly within the same percentile (20th and 19th) aged 4;8, TDB scored within the 31st percentile aged 5;6 and 6;8. In contrast, LIB's performance dropped to the 12th percentile aged 5;6 and the 9th percentile aged 6;8. When compared to younger (6;11), norms aged 7;9, TDB performed slightly above average (73rd %ile) and LIB within the mean (32nd %ile). Finally, aged 8;6 the twins' performance on the Polish receptive vocabulary task had decreased and they scored in the 37th and 13th percentile of norms based on the performance of 6;11-year-old monolingual Polish-speaking children.

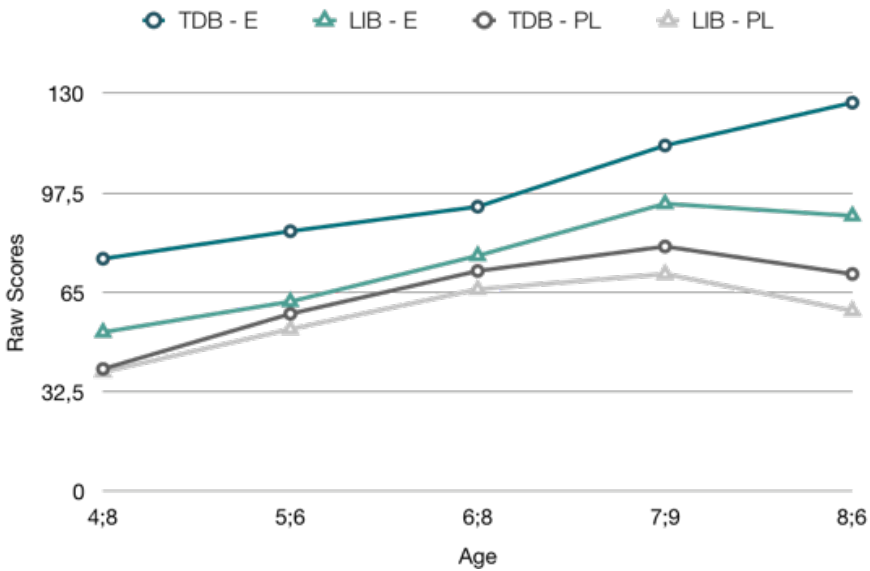


Fig. 2: Raw receptive vocabulary scores.

Raw scores (figure 2) confirmed that the children's Polish receptive vocabulary also decreased in absolute terms between 7;9 and 8;6 (-9 words for TDB and -19 for LIB). In the same time frame, TDB's English receptive vocabulary increased by 14 words on the BPVS-3 while LIB's English score decreased by 4 points.

This graph further indicates that both children knew more words in English than in Polish at all ages, but the gap between TDB's English and Polish receptive vocabulary scores was consistently larger than between LIB's two scores.

Figure 2 also shows that the gap between TDB and LIB's Polish receptive vocabulary is substantially smaller than between their English receptive vocabulary scores. Starting out with a comparable raw score of 40 and 39 on the OTS-R aged 4;8, the gap gradually increased to twelve points aged 8;6. In the BPVS-3, however, the gap between TDB and LIB's score aged 4;8 amounted to 24 points, which decreased marginally to twenty-three aged 5;6 and further to sixteen points aged 6;8. At 7;9 the gap had widened to nineteen points and finally to 37 points aged 8;6.

4.2 Productive Lexis

Figure 3 illustrates their longitudinal development of productive lexis in percentiles. Norms for the Polish productive vocabulary task are only available until the age of 6;11. Therefore, the scores at the last two data collection points were compared to these younger monolingual norms.

Both children perform better on the EVT-2 than the ZNO, when compared to age norms (or younger norms for the Polish test aged 7;9 and 8;6). It is also apparent that TDB clearly outperformed LIB, but both performed within or slightly above the mean.

Figure 3 also indicates a slightly differentiated development profile in relation to their monolingual peers. While LIB's score increased gradually (34th – 39th – 45th – 50th – 58th %ile) and in a linear fashion over the five time points, TDB's performance fluctuated more when compared to monolingual norms.

Their Polish productive lexis showed a different development. LIB performed consistently in the 1st percentile of monolingual Polish norms, even when compared to younger peers (6;11) aged 7;9 and 8;6. TDB, on the other hand, performed three percentile points above LIB aged 4;8, where he stagnated until 5;6, before his score drastically increased to the 14th percentile aged 6;8. When compared to younger age norms (6;11), TDB performed in the 39th percentile aged 7;9 but aged 8;6 his performance had dropped to the 31st percentile of younger monolingual peers

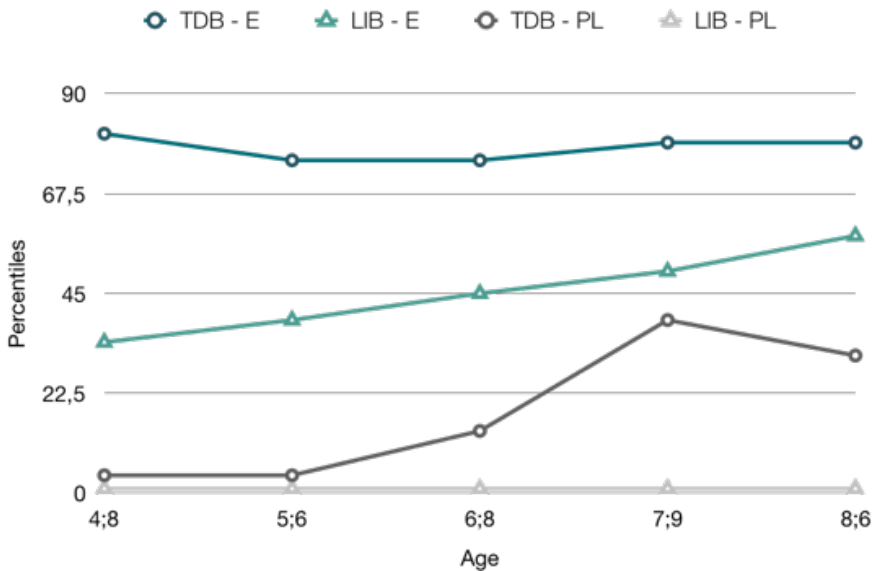


Fig. 3: Productive Lexis in percentiles.

However, the raw scores presented in Figure 4 indicate that TDB's productive vocabulary did not decline as drastically as the percentile score may suggest. Indeed, he only knew one fewer word aged 8;6 than aged 7;9, which resulted in a score of minus eight percentile points.

The graph further shows differences in the gap between TDB's and LIB's English and Polish productive lexis scores. On the EVT-2, the gap between TDB and LIB decreased gradually from 18 (4;8) to 11 (8;6). In contrast, the difference between their scores on the ZNO increased until the age of 6;8 (9 – 16 – 20) before it decreased again aged 7,9 (15) and 8;6 (13). In other words, the gap between TDB's and LIB's score decreased gradually on the EVT-2 but fluctuated on the ZNO.

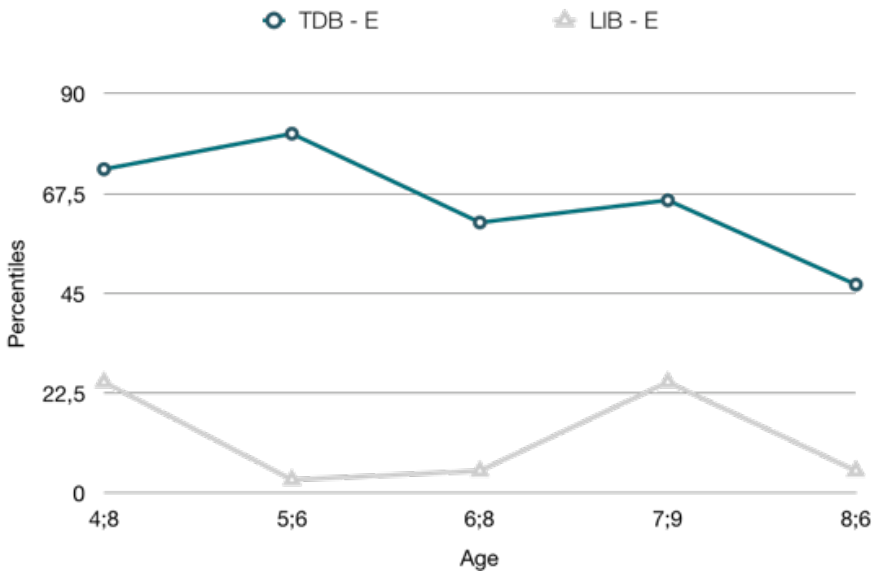


Fig. 4: Raw productive lexis scores.

4.3 Receptive Morphology and Syntax

Figure 5 illustrates TDB's and LIB's longitudinal receptive morphosyntactic development in relation to English age norms. Polish norms are not available because the Polish version has not been standardised.

The graph in Figure 5 indicates fluctuation in TDB's and LIB's morphosyntactic development in relation to English age norms. TDB performed on the 75th percentile aged 4;8, the 81st percentile aged 5;6, the 61st percentile aged 6;8, then improved slightly in relation to monolingual norms aged 7;9, when he scored in the 66th percentile and finally his performance in relation to age norms declined again and he scored in the 47th percentile aged 8;6.

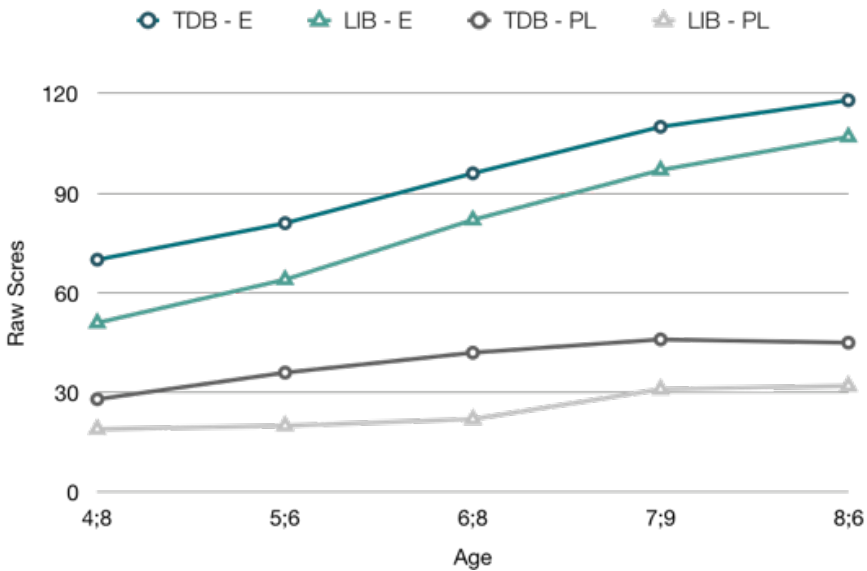


Fig. 5: TROG-2 and TROG-PL results.

LIB's performance showed a reverse trend between the ages of 4;8 and 5;6, when he scored in the 25th and 3rd percentile respectively. In contrast to TDB, LIB's performance stayed relatively stable between the ages of 5;6 and 6;8, when he scored in the 5th percentile. In parallel to TDB's performance, LIB's percentile score also improved between the ages of 6;8 and 7;9 but LIB's score showed a steeper increase, from the 5th to the 21st percentile. Aged 8;6 LIB's score plummeted again and he scored in the 5th percentile.

To compare TDB's and LIB's receptive grammatical knowledge in both languages over time, their TROG-2 scores as 'correct sets'³ are provided in Figure 6.

3 To ensure better reliability, TROG contains four items per grammatical structure. A set is scored as correct if a child repeats all four sentences in one category correctly.

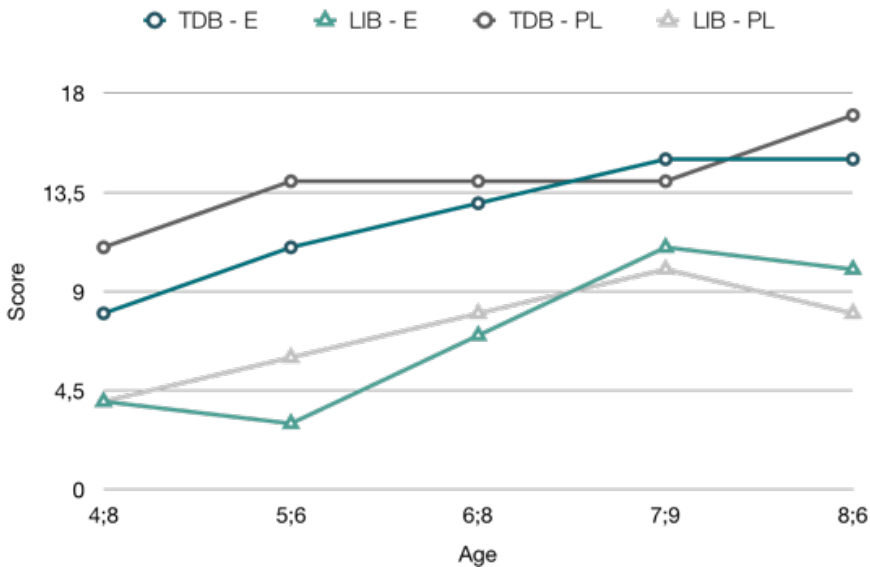


Fig. 6: TROG-2 scores in sets correct.

The gap in ‘correct sets’ between TDB’s and LIB’s TROG-2 score amounted to four sets aged 4;8 and doubled to eight sets aged 5;6, before it decreased to a difference of six sets aged 6;8 and even further to four sets aged 7;9, before increasing again slightly to five aged 8;6.

In Polish, the difference in their performance was nearly twice as large as in English aged 4;8 and amounted to seven sets. At 5;6 years 6;8 and 7;9, the gap between TDB’s and LIB’s performance was identical in their two languages (eight, six and four sets respectively). Aged 8;6, the gap between TDB’s and LIB’s performance on the Polish adaptation of the TROG test increased again substantially to nine sets.

TDB and LIB differed substantially in regard to their language dominance aged 4;8. While TDB clearly scored higher on the Polish than the English version of the TROG, LIB scored equally in both languages. Between the ages of 5;6 and 7;9 TDB and LIB displayed a parallel course of development. They both performed better in Polish than in English until the age of 6;8 when their English performance surpassed their Polish score. Aged 8;6, however, TDB’s Polish score further improved to 17 correct sets and his English performance stagnated at 15 sets while

LIB's performance decreased by one set in English and two sets in Polish. This resulted in a higher English score for LIB aged 8;6 (10 vs. 8 sets) but a higher Polish score for TDB (17 vs. 15 sets).

4.4 Productive morphology and syntax

Figure 7 shows that TDB outperformed LIB on both the English and the Polish version of the sentence repetition task at all ages. Aged 4;8, LIB was unable to complete the activity, resulting in discontinuation of the task. TDB, on the other hand, completed the task and repeated six English and eleven Polish sentences correctly. Aged 5;6, LIB repeated one English and three Polish sentences accurately, in contrast to TDB's 15 correct English and 23 correct Polish sentences. At 6;8, TDB accurately repeated 26 Polish and 29 English sentences, as opposed to seven English and five Polish sentences by LIB. Aged 7;9 TDB scored 43 in English and 40 in Polish, while LIB scored 16 in English and seven in Polish. Finally, aged 8;6, TDB and LIB both improved in English, so that TDB scored 45 and LIB 17, but only TDB could further improve his Polish score to 47. LIB, on the other hand, only repeated six Polish sentences accurately aged 8;6.

In both languages, the gap between TDB's and LIB's scores increased steadily between the ages of 5;6 and 8;6 but more so in Polish than in English (14 vs. 21 sentences).

While the gap between TDB's English and Polish performance decreased over time, from eight aged 4;8 to two aged 6;8, LIB's morphosyntactic development displayed an opposite trajectory. Aged 5;6 and 6;8 the difference between LIB's scores in English and Polish amounted to two sentences, while by the age of 8;6 this gap had increased to eleven. Moreover, TDB's two languages developed in parallel but LIB's English morphosyntactic knowledge displayed a steeper developmental trajectory than his Polish.

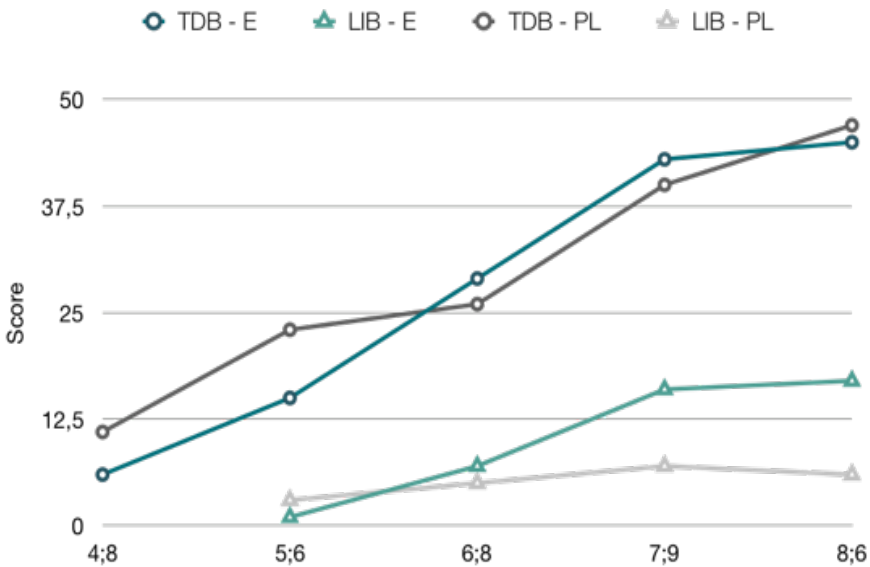


Fig. 7: Results from sentence repetition tasks.

This difference becomes particularly apparent in TDB's higher Polish score aged 4;8 and 5;6 but a higher English score aged 6;8 and 7;9 and finally, aged 8;6, a marginally higher Polish score again. In contrast, LIB only scored higher on the Polish sentence repetition task aged 5;6 and from 6;8 onwards, he scored increasingly higher on the English than the Polish version of the test.

4.5 Narration

4.5.1 Story Length

Figure 8 shows that TDB told longer stories in English at all ages but 7;9. At this age, TDB commented that “*you can’t really put a lot of detail into this story*”, suggesting that he may not have been sufficiently challenged. This would be in line with him telling a shorter story at the age of 8;6. In Polish, TDB and LIB told a story of similar length (66 vs. 69 words) at the age of 4;8 but TDB told longer stories than LIB at all subsequent data collection points.

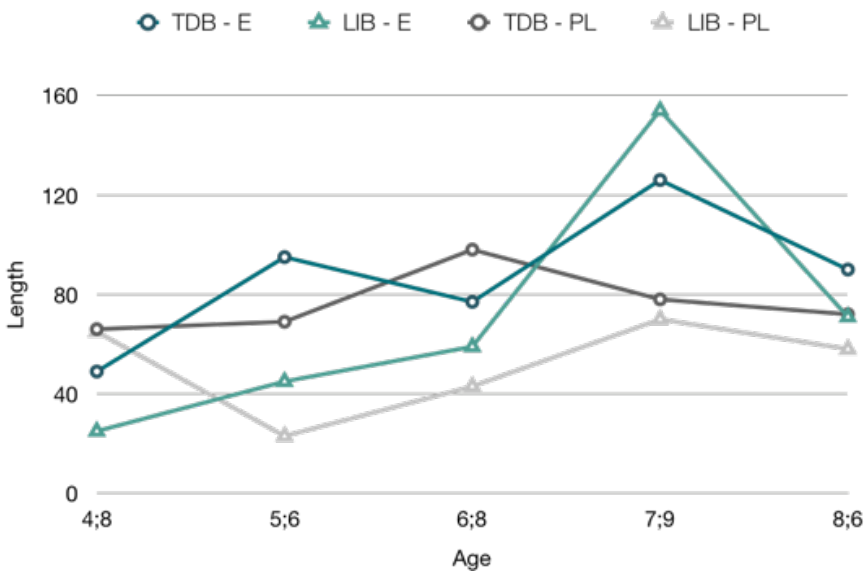


Fig. 8: Story Length (in words).

While LIB consistently told longer stories in English from the age of 5;6 onwards, the relationship between TDB’s two languages fluctuated. TDB’s Polish and English stories differed in length by an average of 20 words at all ages, apart from the age of 7;9, when his English story was substantially longer than his Polish narrative (difference of 38 words). The gap between LIB’s two languages was also most substantial at this

age, when his English narrative was more than twice the length of his Polish story (158 vs. 70 words). Aged 8;6 both children told stories of comparable length in both their languages.

4.5.2 Mean Length of Utterance (MLU)

Mean length of utterance in morphemes was calculated using CLAN (Child Language Analysis programme) and results are presented in Figure 9. TDB consistently outperformed LIB in both languages and at all time points, apart from the first Polish session aged 4;8.

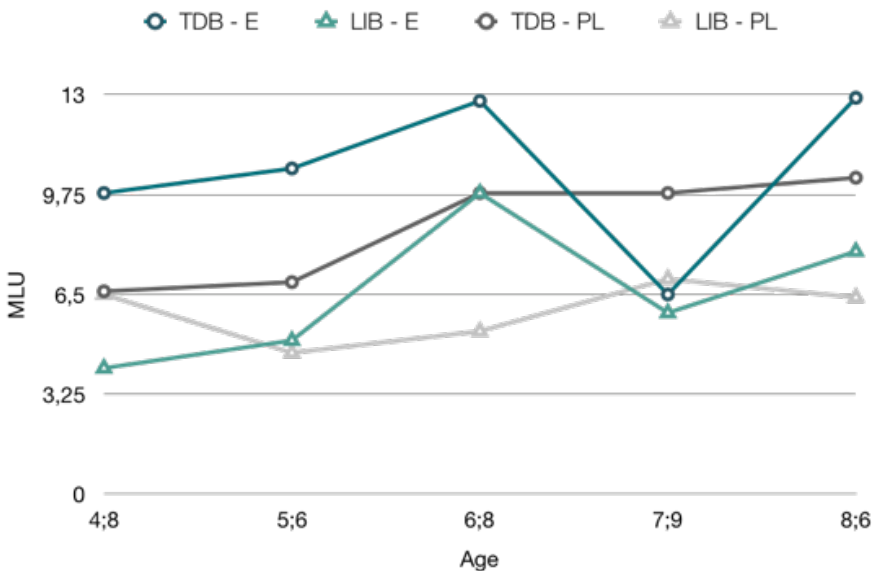


Fig. 9: MLU in Narratives.

TDB's and LIB's MLU development followed a parallel course. Their English MLU increased steadily until the age of 6;8 before it plummeted to a comparable mean length for both children (TDB: 6.5, LIB: 5.9) at the age of 7;9. Their MLU had increased again considerably at the age of 8;6 and amounted to 12.9 for TDB and 7.9 for LIB. This parallel pattern of development could also be observed for their Polish MLU between the

ages of 5;6 and 8;6. Their MLU increased steadily until the age of 7;9, where it stagnated.

Regarding the relationship between TDB's and LIB's two languages, slightly divergent patterns emerged for the two children. TDB's English MLU was higher at all ages apart from 7;9. In contrast, LIB had an MLU of 6.5 in Polish but 4.1 in English aged 4;8. Aged 5;6 his MLU in both languages was comparable (E: 5, PL: 4.6). At the age of 6;8, LIB's English MLU was nearly twice as high as his Polish MLU (E: 9.8, PL: 5.3). By 7;9, LIB's English MLU had decreased to 5.9, while his Polish MLU had increased to 7.0. This pattern had reversed by the age of 8;6 (E: 12.9, PL: 10.3). In contrast to story length, the relationship between English and Polish MLU was more stable in TDB than in LIB. Overall, MLU differentiated well between the two children at all ages and in both languages, apart from 4;8 in Polish and 7;9 in English.

4.5.3 Type-Token Ratio (TTR)

Vocabulary complexity was assessed by calculating the type-token ratio. The TTR assesses how many distinct words participants use in their narratives and thus provides a more precise measure of their vocabulary range than story length alone. It was calculated using CLAN (Child Language Analysis programme).

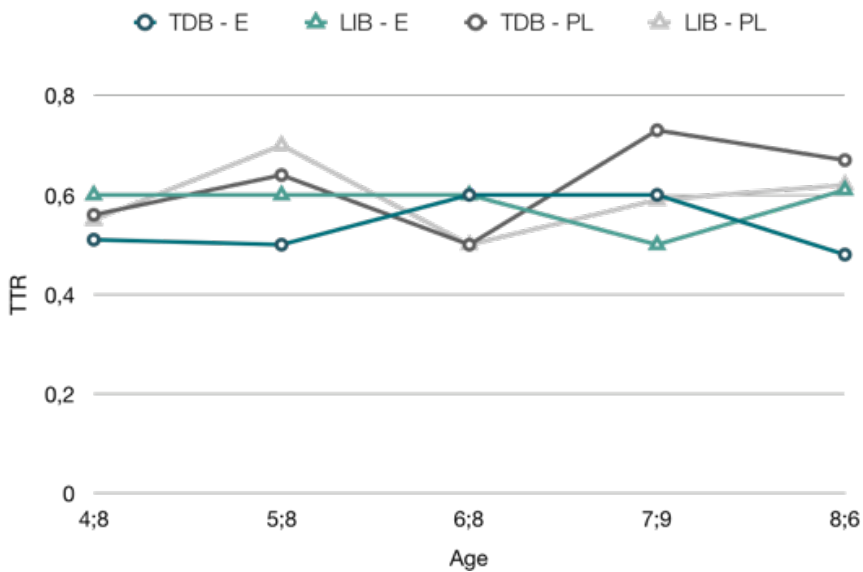


Fig. 10: TTR in narratives

The analysis presented in Figure 10 shows that the difference between TDB and LIB never exceeded more than 0.1 at any time and for either language. Therefore, the following observations have to be treated with caution and it needs to be noted that differences in their TTR were minimal at any given time point.

It is nevertheless noteworthy that LIB outperformed TDB on the English narration task aged 4;8, 5;6 and 8;6. Aged 6;8 they had a similar TTR (0.6) while aged 7;9 TDB marginally outperformed LIB in English (0.6 vs. 0.5). In contrast, TDB and LIB's Polish TTR were comparable at all ages apart from 7;9 (0.6 vs. 0.7).

An investigation of the relationships between their two languages shows that aged 4;8 their TTR in both languages was comparable while aged 5;6 the TTR in their Polish narratives was higher than in their English stories. This pattern reversed for both children by the age of 6;8 but aged 7;9 their Polish TTR was higher again. At the last data collection point, aged 8;6, the TTR in LIB's English and Polish stories was similar (0.6), while TDB's Polish story had a higher TTR than his English narrative (0.67 vs. 0.48).

An analysis of TTR change over time revealed that both children's Polish TTR fluctuated more than the TTR in their English narratives. This was calculated by analysing the difference in TTR between any one given time point and the next. This calculation did not reveal a bigger difference than 0.1 for either child in English and for LIB the TTR even stagnated between the ages of 4;8 and 6;8. In contrast, LIB's Polish TTR decreased from 0.7 to 0.5 between the ages of 5;6 and 6;8 while TDB's Polish TTR increased by 0.2 between the ages of 6;8 and 7;9. Based on these results, TTR on the MAIN narrative task did not appear to differentiate well between TDB and LIB in either of their languages.

4.5.4 Story Structure

Story structure was assessed using the scoring sheets provided in the MAIN manual. The number of story elements (i.e. internal state terms as initiating event, goal, attempt, outcome, internal state term as reaction) was counted for both children and each story in both languages. Results are presented in Figure 11.

The analysis of story structure revealed that TDB outperformed LIB in Polish macrostructure at all ages and in English story structure at every age apart from 6;8. Moreover, the macrostructural development diverged between the two children. LIB displayed a linear development in that the number of macrostructural elements he mentioned during his narratives increased steadily in his two languages until the age of 7;9, before it increased more in Polish but decreased slightly (from 7 to 6) in English.

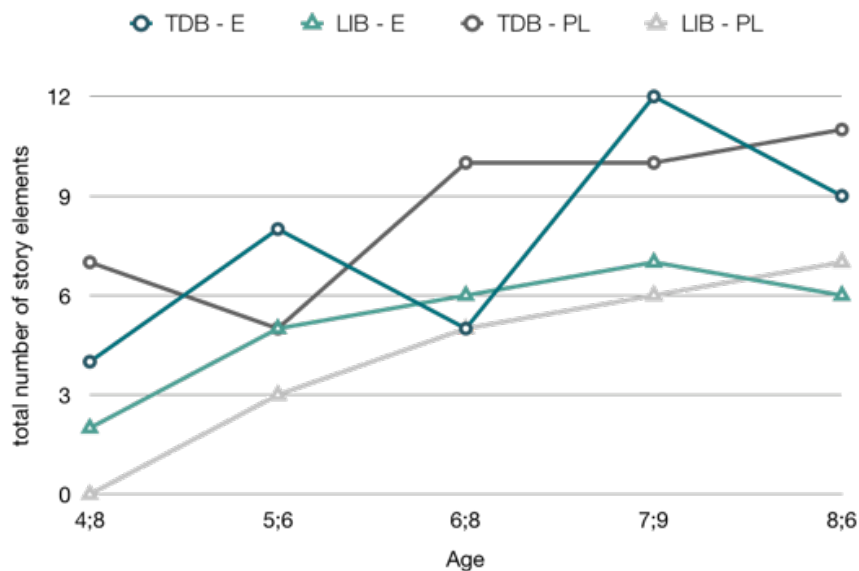


Fig. 11: Story Structure.

In contrast, TDB's macrostructural development fluctuated substantially. In English, the number of story elements he mentioned in his narratives increased from four to eight between the ages of 4;8 and 5;6, before it decreased to five aged 6;8 and rose to twelve aged 7;9. Finally, aged 8;6 TDB mentioned nine of the 17 elements in his English narrative. In Polish, TDB displayed a mirrored pattern. Whenever the number of story elements increased in English, it decreased in Polish (4 – 8 – 5 – 12 – 9 vs. 7 – 5 – 10 – 10 – 11), apart from the age of 7;9 when it stagnated at 10.

Overall, LIB's use of macrostructural elements developed in a linear manner and in parallel in his two languages, while TDB displayed more fluctuation and his use of story elements in English and Polish appears to have developed in reaction to one another (i.e. when one was increasing, the other was decreasing).

4.5.5 Story Complexity

Story complexity was calculated by analysing how many AO (attempt-outcome), single G (goal), GA (goal-attempt), GO (goal-outcome) or full GAO (goal-outcome-attempt) sequences the children used in their narratives. Results are presented in Figure 12.

TDB outperformed LIB with regard to story complexity in English aged 4;8, 5;6 and 6;8 and at all ages but 5;6 in Polish. Additionally, LIB displayed a more linear development in both languages as the number of sequences he described in his stories increased in both languages until the age of 7;9 and then rose further to three sequences in Polish but decreased back to one sequence in English.

The level of story complexity in English and Polish differed between the two participants. LIB displayed a parallel development in English and Polish that was merely staggered for Polish. In contrast, the story complexity in TDB's English narratives fluctuated between one and two sequences, but was fairly consistent at three sequences for Polish at all ages but 5;6.

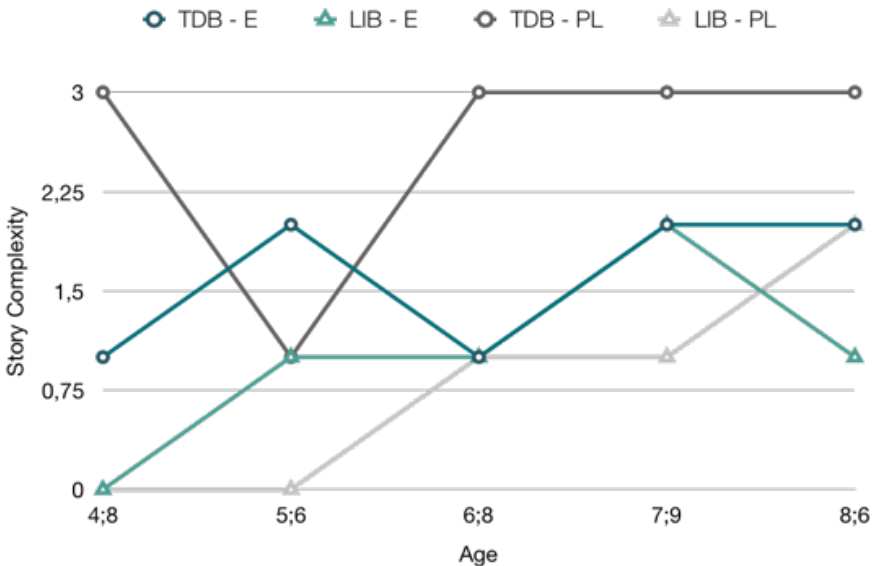


Fig. 12: Story Complexity.

The analysis of number of full GAO sequences revealed that TDB used two such sequences, one in English aged 7;9 and one in Polish aged 8;6. LIB, on the other hand, did not use a single complete GAO sequence in any of his narratives.

4.5.6 Internal State Terms (IST)

IST are terms to describe a character's mental state and can give an indication of a child's Theory of Mind development. The total number of internal state terms in tokens was calculated. Results are provided in Figure 13.

TDB used more IST in English at all ages but 4;8. In Polish, LIB outperformed TDB in his use of IST aged 4;8 and 5;6 but from 6;8 onward this pattern reversed. In contrast to story structure and complexity, both children displayed considerable levels of fluctuation in their use of IST in both languages. This suggests that the use of IST may distinguish better between typically developing English-Polish bilingual children and their peers with DLD at later ages.

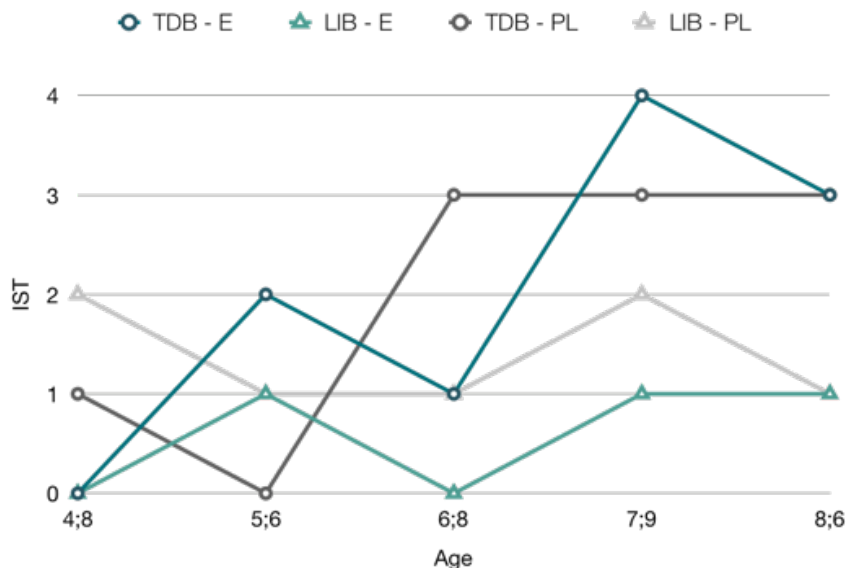


Fig. 13: Internal State Terms.

5 Discussion and Conclusion

In sum, tests that assessed more integrated language skills such as grammar recognition, sentence repetition, story structure and complexity generally distinguished between TDB and LIB, even though interesting age differences will be discussed in more detail below. The fact that vocabulary alone was generally found to be less reliable in this sample may have to do with the children's relatively high socio-economic background, which has been shown to be associated with higher vocabulary scores (Hoff 2006, 2013). By taking a longitudinal approach, the study was also able to show interesting fluctuations in the children's language development, suggesting that it may be qualitatively different from monolingual language development. Details are discussed below. These differences will need to be investigated further in larger longitudinal studies to determine if they are specific to this case or apply to bilingual language development more widely.

5.1 Receptive vocabulary

The analysis of TDB's and LIB's receptive lexical development revealed that TDB outperformed LIB at all ages and in both languages apart from aged 4;8 in Polish. However, the receptive vocabulary tests did not appear particularly reliable in the diagnosis of bilingual DLD as LIB scored within the average or low average range in both languages and at all ages, with the exception of his English score aged 8;6 and his Polish score aged 6;8. This is in line with Verhoeven et al. (2011), who did not find a significant difference between their Bi-TD and their Bi-DLD group in Dutch receptive lexis either. Verhoeven et al.'s (2012) finding of a significant difference between Bi-TD and Bi-DLD in Turkish but not Dutch expressive and receptive lexis could not be confirmed in this study. On the contrary, the gap between TDB's and LIB's productive and receptive vocabularies was larger for English than Polish, where they even overlapped at the age of 4;8 for receptive and approached each other until the age of 5;6 for expressive vocabulary.

The receptive vocabulary tests thus only distinguished well between TDB and LIB at older ages. As the norms of the Polish tests end at the age of 6;11, LIB's and TDB's scores could only be compared to younger monolingual Polish norms at the last two data collection points. It therefore remains unclear whether LIB's performance actually decreased further in relation to age norms.

As vocabulary development is highly correlated with exposure, it is possible that TDB benefitted more from additional exposure than LIB. Therefore, the data from this study suggests that receptive vocabulary tests should be treated with caution, particularly when used with younger bilingual children.

The analysis of TDB's and LIB's BPVS-3 scores in percentiles revealed an interesting pattern with regard to bilingual development. The analysis showed that TDB's and LIB's percentile scores fluctuated for both languages. This suggests that their receptive vocabulary did not evolve at the same pace as their monolingual peers'. This finding was substantiated by the fact that TDB's and LIB's English and Polish vocabulary increased in absolute terms between the ages of 4;8 and 7;9 while it fluctuated in relation to age norms. Note that TDB's Polish and LIB's English and Polish absolute scores had decreased between the age of 7;9 and 8;6, which highlights the fluctuating nature of bilingual language development.

The receptive vocabulary in both languages of both children evolved in a largely comparable manner. This indicates that even though LIB's score in both languages was consistently lower, both children were equally influenced by their bilingual status. This suggests that bilingual language development may be qualitatively different from monolingual development and that monolingual vocabulary growth curves do not necessarily apply to bilingual children.

5.2 Productive vocabulary

The fluctuation in relation to monolingual norms was more evident for the receptive than the productive vocabulary scores. Here TDB performed consistently between the 75th and the 80th percentile in English. LIB, on the other hand, showed slightly more fluctuation (34th-58th) but it needs to be noted that the last two data collection points were compared to younger norms. This finding suggests that TDB's and LIB's productive English vocabulary evolved largely in parallel to monolingual peers and both performed within or above the monolingual average range.

For Polish, LIB's performance was consistently far below monolingual age norms and he performed within the 1st percentile at all data collection points. TDB also performed in the low average and borderline range at all ages where monolingual norms were available. This suggests that the growth of their Polish productive vocabulary was influenced more substantially by their bilingual status than their English vocabulary, which is likely due to English being the majority language.

Overall, the investigation of vocabulary development highlighted that productive and receptive vocabulary tests were less reliable in the distinction between TDB and LIB. Furthermore, both children scored higher on the English than the Polish tests, which reflects the minority-majority status of their two languages. Additionally, bilingual-specific developmental patterns were revealed. Finally, the overall results of the vocabulary tests emphasise the role of exposure for vocabulary development, which was highlighted by the minority-majority language difference.

5.3 Receptive morphology and syntax

TDB outperformed LIB on the TROG in both languages and at all ages. It needs to be noted, however, that the gap between their scores was more substantial at older ages (from 5;6 onward) with the exception of 7;9, when both children performed within the average range again. This suggests that their performance on the English TROG test was linked to

exposure and highlights once more the importance of input for bilingual development. This corresponds to Jacobson (2011) and Jacobson and Walden (2013) who found profiles of English-Spanish bilingual children with and without language impairment to overlap more in lower than in higher grades.

Both children performed better on the Polish task at younger ages with the exception of LIB aged 4;8, when he performed similarly in both languages. This suggests that the increased structured input they received in the majority language in school may have improved their English receptive grammar, so that by the age of 7;9 they performed better on the English than the Polish task.

Finally, an observation regarding the stability of TDB's and LIB's grammatical systems could be made as TDB's receptive grammatical understanding increased steadily over the course of the four years, while LIB's understanding of grammatical structures decreased between the ages of 4;8 and 5;6 in English and for both languages between the ages of 7;9 and 8;6. This implies that TDB's grammatical system may be more robust.

5.4 Productive morphology and syntax

TDB outperformed LIB at all ages and in both languages on the sentence repetition task. This provides further evidence for sentence repetition tasks as a reliable tool to distinguish between typical language development and DLD (Archibald & Joanisse 2009; Armon-Lotem & Meir 2016; Riches 2012; Stokes et al. 2006; Thordardottir & Brandeker 2013). Furthermore, the two children displayed a slightly diverging relationship between their two languages. The gap between TDB's English and Polish performance decreased consistently over time, while the gap between LIB's English and Polish performance widened substantially between 6;8 and 8;6. Once more this difference cannot be traced back to different exposure to the two languages as both children heard a similar amount of both languages. However, English speech and

language therapy and regular formal instruction in the majority language may explain LIB's higher score in English.

The finding that LIB's grammatical production was not balanced across his two languages is also important to bear in mind with regard to bilingual language impairment. While bilingual children with language impairment tend to show a low performance in both languages, typically developing children will perform well in at least one of their two languages. This is not to say, however, that a mismatch between the two languages is exclusive to bilingual children with typical development. It seems to also occur in bilingual DLD but the important point to note is that their overall performance in both languages is substantially lower than that of their typically developing peers.

5.5 Narration

Story length was more reliable in the distinction between TDB and LIB before the age of 6;8. This finding provides further evidence for a declining correlation between age and storytelling abilities in typically developing bilingual children, from the age of seven onward (Bohnacker 2016; Gagarina 2016; Mavis et al. 2016; Roch et al. 2016).

Previous studies only occasionally found story length to differ substantially between children with and without language impairment (Altman et al. 2016; Cleave et al. 2010; Fey et al. 2004). Therefore, story length does not appear to be a very reliable indicator of language impairment in bilingual children, particularly at older ages.

In contrast, MLU in morphemes was found to be reliable in the distinction between TDB and LIB in both languages and across ages. This finding corresponds to LIB's difficulties with morphology and is also in line with Altman et al. (2016).

In contrast, the differences between TDB's and LIB's TTR scores were only marginal. This stands in contrast with Iluz-Cohen and Walters (2012), Altman et al. (2016) and Tsimpli et al. (2016), who all found vocabulary diversity to differ between Bi-TD and Bi-DLD. This divergence between the results obtained in this study and previous

research could be due to LIB's relatively high (English) productive vocabulary scores, which in turn could be influenced by the twins' relatively high socio-economic background (Hoff 2006, 2013). Support for his hypothesis is provided by the finding that the difference between TDB's and LIB's TTR was larger for the English than the Polish stories, which corresponds to LIB's substantially lower Polish vocabulary score.

Therefore, if TTR is to be considered in the distinction of typical and atypical bilingual language development, productive lexis scores need to be taken into account. The TTR of bilingual children with language impairment but with a productive vocabulary score within the normal range, might not differ significantly from their typically developing bilingual peers.

The story structure analysis revealed that TDB outperformed LIB at all ages in Polish and at every age apart from 6;8 in English. This suggests that story structure (i.e.: the number of story elements) can be indicative of bilingual DLD. This corresponds to Squires et al.'s (2014) findings but stands once more in contrast to Iluz-Cohen and Walters (2012), Altman et al. (2016) and Tsimpli et al. (2016), who did not find macrostructure to differentiate between the two bilingual groups. This study therefore adds to the controversial findings regarding the distinctive potential of macrostructural elements in the assessment of bilingual children. Due to this controversy, story structure alone probably cannot be recommended for the diagnosis of DLD in bilingual children.

It was shown that LIB's development of story structure was linear and increased steadily between time point one and four in English and Polish. In contrast, TDB's use of story structure elements fluctuated in both languages and whenever he used more elements in one language, he used fewer in the other. Judging from these findings, it seems like TDB's languages stand in a more dynamic relationship to each other.

The observations from the story structure analysis were further confirmed in an assessment of TDB's and LIB's story complexity. The children differed substantially in the number of sequences (i.e.: attempt-goal; single goal; goal-attempt, goal-outcome; full goal-attempt-outcome) they described. TDB outperformed LIB and the difference was

particularly apparent for Polish and in their use of complete GAO sequences.

This finding implies that LIB has more difficulty arranging his narratives coherently. The fact that LIB did not use a single whole GAO sequence could also imply difficulty in his working memory, as GAO sequences require to keep the goal in mind while describing the attempt and finally the outcome. This hypothesis would also be supported by the fact that LIB prefers to read shorter factual texts, while TDB shows a preference for narratives and novels. The reading of novels requires a longer attention span and better working memory as the reader needs to keep a lot of information in mind while processing subsequent sentences and chapters. Reading difficulties have been linked to DLD (Baird et al. 2011) and this paper provides further reason to investigate this link in more depth in the bilingual population.

Finally, TDB's and LIB's use of internal state terms showed that TDB used more internal state terms in both languages from the age of 5;6 onward for English and from 6;8 onward for Polish. This finding is in line with Squires et al. (2014). TDB's preference for reading stories and novels, where internal state terms are more common than in factual publications, might also influence this result. This hypothesis seems to be supported by the fact that the difference between TDB and LIB became particularly visible at older ages, when they would have become more fluent readers and start to influence their choice of reading material.

5.7 Conclusion

Overall, this study hence shows that bilingual language development does not necessarily map on to monolingual growth curves, which needs to be taken into account during assessments. Moreover, some language measures only showed differences between TDB and LIB at certain ages, which highlights a further level of complexity in bilingual language development and emphasises the need for more holistic and longitudinal approach to the study of typical and atypical bilingual

development. Finally, this study highlighted that measures of vocabulary size and complexity may be less reliable in distinguishing between typical bilingual development and bilingual DLD in children with a higher socio-economic background, adding a further level of complexity to the assessment of bilingual children.

Acknowledgements

I would like to thank emer. o. Univ.-Prof. Mag. Dr. Wolfgang Dressler, Professor Shula Chiat and Univ.-Prof. Mag. Dr. Ute Smit for their supervision and mentorship during this project and beyond, Prof. Dr. hab. Ewa Haman and Dr. hab. Zofia Wodniecka-Chlipalska for their support throughout, the twins for their willingness to participate over the years, and their parents for allowing them to do so.

References

- Altman, Carmit, Sharon Armon-Lotem, Sveta Fichman & Joel Walters. 2016. Mental state terms in the narratives of English – Hebrew bilingual preschool children with and without specific language impairment. *Applied Psycholinguistics* 37. 165–193.
- Archibald, Lisa M. D. & Marc F. Joanisse. 2009. On the sensitivity and specificity of nonword repetition and sentence recall to language and memory impairments in children. *Journal of Speech, Language, and Hearing Research* 52(4). 899–914.
- Armon-Lotem, Sharon. 2010. Instructive bilingualism: Can bilingual children with specific language impairment rely on one language in learning a second one? *Applied Psycholinguistics* 31(2).
<http://doi.org/10.1017/S0142716409990385>.
- Armon-Lotem, Sharon. 2014. Between L2 and SLI: inflections and prepositions in the Hebrew of bilingual children with TLD and monolingual children with SLI. *Journal of Child Language* 41(1). 3–33.
<http://doi.org/10.1017/S0305000912000487>.
- Armon-Lotem, Sharon & Natalie Meir. 2016. Diagnostic accuracy of repetition tasks for the identification of specific language impairment (SLI) in

- bilingual children: Evidence from Russian and Hebrew. *International Journal of Language and Communication Disorders* 1–17.
<http://doi.org/10.1111/1460-6984.12242>.
- Baird, Gillian, Vicky Slonims, Emily Simonoff & Katharina Dworzynski. 2011. Impairment in non-word repetition: A marker for language impairment or reading impairment? *Developmental Medicine and Child Neurology* 53(8). 711–716. <http://doi.org/10.1111/j.1469-8749.2011.03936.x>.
- Bak, Thomas. 2016. Cooking pasta in La Paz: Bilingualism, bias and the replication crisis. *Linguistic Approaches to Bilingualism* 6(5). 699–717.
- Banasik, Natalia, Ewa Haman & Magdalena Smoczyńska. 2011. Zadanie Powtarzania Zdań.
- Bishop, Dorothy VM. 2003. Test for reception of grammar: Version 2: TROG-2. Pearson Assessment. London: Pearson.
- Bishop, Dorothy V. M. 2017. Why is it so hard to reach agreement on terminology? The case of developmental language disorder (DLD). *International Journal of Language & Communication Disorders* 52(6). 671–680. <https://doi.org/10.1111/1460-6984.12335>.
- Blom, Elma, Jan De Jong, Antje Orgassa, Anne Baker & Fred Weerman. 2013. Verb inflection in monolingual Dutch and sequential bilingual Turkish-Dutch children with and without SLI. *International Journal of Language and Communication Disorders* 48(4). 382–393. <http://doi.org/10.1111/1460-6984.12013>.
- Bohnacker, Ute. 2016. Tell me a story in English or Swedish: Narrative production and comprehension in bilingual preschoolers and first graders. *Applied Psycholinguistics* 37. 19–48.
- Caramazza, Alfonso. 1986. On drawing inferences about the structure of normal cognitive systems from the analysis of patterns of impaired performance: The case for single-patients studies. *Brain and Cognition* 5. 41–66.
- Clahsen, Harald, Monika Rothweiler, Franziska Sterner & Solveig Chilla. 2014. Linguistic markers of specific language impairment in bilingual children: The case of verb morphology. *Clinical Linguistics & Phonetics* 49. 1–13. <http://doi.org/10.3109/02699206.2014.886726>.
- Cleave, Patricia L, Luigi E Girolametto, Xi Chen & Carla J Johnson. 2010. Narrative abilities in monolingual and dual language learning children with specific language impairment. *Journal of Communication Disorders* 43(6). 511–522. <http://doi.org/10.1016/j.jcomdis.2010.05.005>.

- Duff, Patricia. 2008. *Case Study Research in Applied Linguistics*. New York: Taylor & Francis Group.
- Dunn, Loyd M., Douglas M. Dunn, Ben Styles & Julie Sewell. 2009. *British Picture Vocabulary Scale: Third Edition*. GL Assessment. Brentford: GL Assessment.
- Eng, Nancy & Barabara O'Connor. 2000. Acquisition of Definite Article + Noun Agreement of Spanish-English Bilingual Children with Specific Language Impairment. *Communication Disorders Quarterly* 21(2). 114–124. <http://doi.org/10.1177/152574010002100205>.
- Fey, Marc E., Hugh W. Catts, Kerry Proctor-Williams & J. Bruce Tomblin B. 2004. Oral and Written Story Composition Impairment. *Journal of Speech Language and Hearing Research* 47, 1301–1318. <http://doi.org/1092-4388/04/4706-1301>.
- Gagarina, Natalia. 2016. Narratives of Russian – German preschool and primary school bilingual: Rasskaz and Erzählung. *Applied Psycholinguistics* 37. 91–122.
- Gagarina, Natalia, Daleen Klop, Saria Kunnari, Kula Tantele, Taina Välimaa, Ingrida Balciuniene, Ute Bohnacker & Joel Walters. 2012. ZASPiL Nr. 56 – December 2012 MAIN: Multilingual Assessment Instrument for Narratives.
- Gillam, Ronald B., Elizabeth D. Pena., Lisa M. Bedore, Thomas M. Bohman & Anita Mendez-Perez. 2013. Identification of SLI in Bilingual Children: 1. Assessment in English. *Journal of Speech, Language, and Hearing Research* 56(December). 1813–1823. [http://doi.org/10.1044/1092-4388\(2013/12-0056\)](http://doi.org/10.1044/1092-4388(2013/12-0056)).
- Gutiérrez-Clellen, Vera F., Gabriela Simon-Cerejido & Christine Wagner. 2008. Bilingual children with language impairment: A comparison with monolinguals and second language learners. *Applied Psycholinguistics* 29(1). 3–19. <http://doi.org/10.1017/S0142716408080016>.
- Håkansson, Gisela, Eva-Kristina Salameh & Ulrika Nettelblatt. 2003. Measuring language development in bilingual children: Swedish-Arabic children with and without language impairment. *Linguistics* 41(2). 255–288. <http://doi.org/10.1515/ling.2003.009>.
- Haman, Ewa & Krzysztof Fronczyk. 2012. *Obrazkowy Test Słownikowy - Rozumienie (OTSR)*. Gdańsk: Pracownia Testów Psychologicznych i Pedagogicznych.
- Haman, Ewa & Magdalena Smoczynska. 2010. *Zadanie Nazywanie Obrazków*. Faculty of Psychology. University of Warsaw. Unpublished material.

- Hoff, Erika. 2006. How social contexts support and shape language development. *Developmental Review* 26. 55–88.
- Hoff, Erika. 2013. Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology* 49. 4–14.
- Hymes, Dell. 1979. Sapir, Competence, Voices. In C. J. Fillmore, D. Kempler & W. S.-Y. Wang (Eds.), *Individual Difference in Language Ability Language Behavior*, 33–45. New York: Academic Press.
- Iluz-Cohen, Peri & Joel Walters. 2012. Telling stories in two languages: Narratives of bilingual preschool children with typical and impaired language. *Bilingualism: Language and Cognition* 15(1). 58–74.
<http://doi.org/10.1017/S1366728911000538>.
- Jacobson, Peggy F. & Richard Schwartz. 2002. Morphology in incipient bilingual Spanish-speaking preschool children with specific language impairment. *Applied Psycholinguistics* 23(1). 23–41.
<http://doi.org/10.1017/S0142716402000024>.
- Jacobson, Peggy F. 2011. The effects of language impairment on the use of direct object pronouns and verb inflections in heritage Spanish speakers: A look at attrition, incomplete acquisition and maintenance. *Bilingualism: Language and Cognition* 15(1). 22–38.
<http://doi.org/10.1017/S1366728911000484>.
- Jacobson, Peggy F. & Patrick R. Walden. 2013. Lexical diversity and omission errors as predictors of language ability in the narratives of sequential Spanish-English bilinguals: a cross-language comparison. *American Journal of Speech- Language Pathology* 22(3). 554–65.
[http://doi.org/10.1044/1058-0360\(2013/11-0055\)](http://doi.org/10.1044/1058-0360(2013/11-0055)).
- Kiebzak-Mandera, Dorota, Agnieszka Otwinowska & Marta Bialecka-Pikula. 2012. *MAIN Multilingual Assessment Instrument for Narratives: Polish Version*. Faculty of Psychology. University of Warsaw. Unpublished material.
- Kohnert, Kathryn, Jennifer Windsor & Kerry Danahy Ebert. 2009. Primary or “specific” language impairment and children learning a second language. *Brain and Language* 109(2–3). 101–111.
<http://doi.org/10.1016/j.bandl.2008.01.009>.
- Kohnert, Kathryn. 2010. Bilingual children with primary language impairment: Issues, evidence and implications for clinical actions. *Journal of Communication Disorders* 43(6). 456–473.
<http://doi.org/10.1016/j.jcomdis.2010.02.002>.

- Leonard, Laurence B. 1998. *Children with Specific Language Impairment*. Massachusetts: MIT Press.
- Leopold, Werner. 1939. Speech Development of a Bilingual Child: A Linguist's Record, Volume I: Vocabulary Growth in the First Two Years. *Northwestern University Studies in the Humanities*, Volume 6. vii-185.
- Leopold, Werner. 1940. The Study of Child Language and Infant Bilingualism. *WORD* 4(1). 1-17.
- Leopold, Werner. 1947. *Speech Development of a Bilingual Child, Volume II: Sound-Learning in the first Two Years*. New York: AMS Press.
- Leopold, Werner. 1970. *Speech Development of a Bilingual Child, Volume III: Grammar and General Problems in the First Two Years*. New York, NY: AMS Press.
- Marinis, Theo, Shula Chiat, Sharon Armon-Lotem, Gibbons, D. & E. Gipps E. 2011. School Age Sentence Imitation Test (SASIT) – English. Reading: University of Reading.
- Mavis, Ilknur, Uge Müge & Natalia Gagarina. 2016. Macrostructure components in narrations of Turkish – German bilingual children. *Applied Psycholinguistics* 37, 69–89.
- Mieszkowska, Karolina, Agnieszka Otwinowska, Marta Białecka-Pikul, Dorota Kiebzak-Mandera, Marcin Opacki & Ewa Haman. 2020. Polish MAIN: How was it developed and how has it been used so far? *ZAS Papers in Linguistics* 64, 169–181. <https://doi.org/10.21248/zaspil.64.2020.571>.
- Müller, Lisa-Maria. 2018. Kommunikationsstrategien im typischen Bilingualismus und unter den Bedingungen von spezifischer Sprachstörung: Eine explorative Zwillingsfallstudie. *Sprache, Stimme und Gehör* 42(2). 82-86.
- Orgassa, Antje & Fred Weerman. 2008. Dutch gender in specific language impairment and second language acquisition. *Second Language Research* 24(3). 333–364. <http://doi.org/10.1177/0267658308090184>.
- Paap, Kenneth R. 2013. The role of componential analysis, categorical hypothesising, replicability and confirmation bias in testing for bilingual advantages in executive functioning. *Journal of Cognitive Psychology* 26(3). 242-255.
- Paradis, Johanne. 2007. Bilingual children with specific language impairment: Theoretical and applied issues. *Applied Psycholinguistics* 28(3). 551–564. <http://doi.org/10.1017/S0142716407070300>.
- Reilly, Sheena, Dorothy V. M. Bishop & Bruce Tomblin. 2014a. Terminological debate over language impairment in children: forward movement and

- sticking points. *International Journal of Language and Communication Disorders* 49. 452–462.
- Reilly, Sheena, Bruce Tomblin, James Law, Cristina McKean, Fiona K. Mensah, Angela Morgan, Sharon Goldfeld, Jan M. Nicholson & Melissa Wake. 2014b. Specific Language Impairment: a convenient label for whom? *International Journal of Language and Communication Disorders* 49. 416–451.
- Riches, Nick G. 2012. Sentence repetition in children with specific language impairment: An investigation of underlying mechanisms. *International Journal of Language and Communication Disorders* 47(5), 499–510. <http://doi.org/10.1111/j.1460-6984.2012.00158.x>.
- Roch, Maja, Elena Florit & Chiara Levorato. 2016. Narrative competence of Italian – English bilingual children between 5 and 7 years. *Applied Psycholinguistics* 37. 49–67.
- Rothweiler, Monika, Solveig Chilla & Babur Ezel. 2010. Specific language impairment in Turkish: evidence from case morphology in Turkish-German successive bilinguals. *Clinical Linguistics & Phonetics* 24(7). 540–555. <http://doi.org/10.3109/02699200903545328>.
- Rothweiler, Monika, Solveig Chilla & Harald Clahsen. 2012. Subject–verb agreement in Specific Language Impairment: A study of monolingual and bilingual German-speaking children. *Bilingualism: Language and Cognition* 15(1). 39–57. <http://doi.org/10.1017/S136672891100037X>.
- Semel, Eleanor, Elisabeth H. Wiig & Wayne A. Secord. 2003. *Clinical Evaluation of Language Fundamentals - Fourth Edition (CELF-4)*. San Antonio, TX: Pearson Assessment.
- Smoczynska, Magdalena. 2005. *Test for Reception of Grammar - Polish translation*. Krakow: Uniwersytet Jagiellonski.
- Stern, Clara & William Stern. 1907. *Die Kindersprache. Eine psychologische und sprachtheoretische Untersuchung*. Leipzig: Barth.
- Stokes, Stephanie F., Anita M-Y. Wong, Paul Fletcher & Laurence B. Leonard. 2006. Nonword repetition and sentence repetition as clinical markers of specific language impairment: the case of Cantonese. *Journal of Speech, Language, and Hearing Research* 49(2), 219–236. [http://doi.org/10.1044/1092-4388\(2006/019\)](http://doi.org/10.1044/1092-4388(2006/019)).
- Stromswold, Karin. 2006. Why aren't identical twins linguistically identical? Genetic, prenatal and postnatal factors. *Cognition* 101(2). 333–384. <http://doi.org/10.1016/j.cognition.2006.04.007>.

- Squires, Katie Ellen, Mirza J. Lugo-Neris, Elizabeth Peña, Lisa Bedore, Thomas M. Bohman & Ronald B. Gillam. 2014. Story retelling by bilingual children with language impairments and typically developing controls. *International Journal of Language & Communication Disorders* 49(1). 60–74.
<http://doi.org/10.1111/1460-6984.12044>.
- Thordardottir, Elin & Myrto Brandeker. 2013. The effect of bilingual exposure versus language impairment on nonword repetition and sentence imitation scores. *Journal of Communication Disorders* 46(1). 1–16.
<http://doi.org/10.1016/j.jcomdis.2012.08.002>.
- Tomblin, J. Bruce, Nancy L. Records & Xuyang Zhang. 1996. A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech and Hearing Research* 39(6). 1284–1294.
- Tsimpli, Ianthi M., Eleni Peristeri & Maria Andreou. 2016. Narrative production in monolingual and bilingual children with specific language impairment. *Applied Psycholinguistics* 37. 195–216.
- Verhoeven, Ludo, Judit Steenge & Hans van Balkom. 2011. Verb morphology as clinical marker of specific language impairment: Evidence from first and second language learners. *Research in Developmental Disabilities* 32(3). 1186–1193. <http://doi.org/10.1016/j.ridd.2011.01.001>.
- Verhoeven Ludo, Judit Steenge & Hans van Balkom. 2012. Linguistic transfer in bilingual children with specific language impairment. *International Journal of Language & Communication Disorders* 47(2). 176–83.
<http://doi.org/10.1111/j.1460-6984.2011.00092.x>.
- Williams, Kathleen T. 2007. *Expressive Vocabulary Test*, Second Edition (EVT-2). London: Pearson Assessment.