

Uniwersytet Warszawski
Wydział Psychologii

Karolina Mieszkowska

Internal State Lexicon of bilingual
and monolingual pre- and early school children

Terminy wyrażające stany wewnętrzne u dwujęzycznych
i jednojęzycznych dzieci w wieku przedszkolnym i wczesnoszkolnym

Praca doktorska pod kierunkiem
dr hab. Ewy Haman (promotor)
dr hab. Agnieszki Otwinowskiej-
Kasztelanic (promotor pomocniczy)

Warszawa, 2018

Acknowledgments

I would like to thank the children and their parents who participated in the study, and the invincible research assistants and volunteers who were engaged in the testing, transcription, and coding of the data. I gratefully acknowledge the contribution of my supervisors, dr hab. Ewa Haman and dr hab. Agnieszka Otwinowska-Kasztelanic whose commitment and input in this work is inestimable and greatly appreciated. I would like to thank dr hab. Marta Bialecka-Pikul for her influence and knowledgeable advice in this study at many stages of its development. Last but not least, I would like to thank the Principal Investigators of Bi-SLI-Poland and WLRB, dr hab. Ewa Haman, dr hab. Zofia Wodniecka and dr hab. Agnieszka Otwinowska-Kasztelanic, for allowing me to be part of the two projects and re-use the data gathered there.

Podziękowania

Pragnę podziękować dzieciom i rodzicom, którzy wzięli udział w badaniu, oraz niezłomnym badaczom i wolontariuszom, którzy byli zaangażowani w badania, spisywanie i kodowanie danych. Podziękowania należą się również moim promotorkom, dr hab. Ewie Haman i dr hab. Agnieszce Otwinowskiej-Kasztelanic za ich nieocenione zaangażowanie i wkład w poniższą pracę. Chciałabym podziękować także dr hab. Marcie Bialeckiej-Pikul za jej wpływ i wnikliwe rady oferowane na wielu etapach realizowania tego badania. Na koniec pragnę podziękować kierownikom projektów Bi-SLI-Poland i WLRB, dr hab. Ewie Haman, dr hab. Zofii Wodnieckiej i dr hab. Agnieszce Otwinowskiej-Kasztelanic za to że mogłam być częścią tych dwóch projektów i wykorzystać zebrane w nich dane.

Abstract

Our lexicon includes terms that demonstrate our understanding of the fact that people's actions are governed by internal states such as beliefs, desires, and emotions. These words are internal state terms (IST) and their examples include “think”, “want”, “notice”, “surprised”, “scared”. Altogether, they constitute the Internal State Lexicon (ISL). The ISL might be placed at the intersection of language and theory of mind (ToM) development: IST need to be learned as any other words in the lexicon, and their use is an indicator of children's mentalizing abilities. The present thesis set out to investigate the use of IST in Polish-English bilingual children at pre- and early school age (4.5-7 years old). Bilingual upbringing may uniquely shape the use of IST. On one hand, bilingual children hear less of each language and have to exert constant control over the two languages. This may translate to lower language performance (e.g. Haman, et al., 2017). On the other hand, bilinguals often outperform monolinguals in their cognitive ability, including theory of mind (Farhadian, et al., 2010; Goetz, 2003; Kovács, 2009).

The primary aims of the present analysis were to investigate whether language status (bilingual vs. monolingual) influences the use of IST, and whether IST develop comparably in both languages of the bilingual child. The participants included 75 Polish-English bilingual children aged 4.5-7 years old and living in the UK, and matched 75 Polish monolingual children living in Poland. The internal state terms were elicited via child-made narratives based on a set of pictures. Three subclasses of IST were coded: emotional, mental, and perceptual terms. Children were also asked to retell the story immediately after listening to a model story, and to answer comprehension questions about the story protagonists' internal states. This was done to explore the effect of modelling on the IST production and to compare the IST production in story-telling and a relatively more interactive context of explicit conversation about internal states. Additionally, the children's vocabulary and grammar knowledge was assessed in their respective languages, and their theory of mind performance was measured with a test of reflection on thinking.

The results showed that while bilinguals exhibited relatively poorer language abilities than monolinguals, the children did not differ in the amount of internal state terms produced when telling a story. However, bilinguals outperformed monolingual peers on the theory of mind task. Thus, these two differences might have evened each other out, leading to no overall difference in the use of IST between the groups. Also, bilinguals used IST similarly in their two languages. It was also found that giving children a model story and explicitly asking them about the internal states of story protagonists sensitized them to their knowledge, desires, and beliefs which resulted in more internal state references in the retellings and answers to the questions than in the narratives told by children on the basis of pictures alone.

Keywords: Internal State Lexicon, internal state terms, mental state terms, theory of mind, narratives, bilingual children, child language development

Streszczenie

Nasz słownik zawiera pojęcia, które wskazują na to, że rozumiemy iż ludzkim zachowaniem kierują stany wewnętrzne takie jak przekonania, pragnienia i wiedza. Są to terminy wyrażające stany wewnętrzne (ang. *internal state terms*, IST) i należą do nich np. „myśleć”, „chcieć”, „zauważyć”, „zaskoczony”, „przestraszony”. Razem terminy te tworzą leksykon terminów wewnętrznych (ang. *Internal State Lexicon*, ISL). ISL można umiejscowić na styku języka i teorii umysłu: IST muszą zostać przyswojone jak każde inne słowo, a ich użycie jest jednym z wyznaczników zdolności do mentalizacji. Niniejsza praca doktorska bada użycie IST u polsko-angielskich dzieci dwujęzycznych w wieku przed- i wczesnoszkolnym (4,5 – 7 lat). Dwujęzyczność może w sposób unikalny wpływać na użycie IST. Z jednej strony, dzieci dwujęzyczne mają mniej kontaktu z każdym ze swoich języków, w porównaniu z jednojęzycznymi rówieśnikami i muszą monitorować aktywację obu języków. To może skutkować niższymi umiejętnościami językowymi w porównaniu z dziećmi jednojęzycznymi (np. Haman i in., 2017). Z drugiej strony, dzieci dwujęzyczne prześcigają swoich jednojęzycznych rówieśników w zdolnościach poznawczych, w tym teorii umysłu (Farhadian i in., 2010; Goetz, 2003; Kovács, 2009).

Głównym celem pracy było zbadanie czy status językowy (jedno- lub dwujęzyczność) wpływa na użycie IST i czy dzieci dwujęzyczne używają IST w podobny sposób w obu swoich językach. Uczestnikami badania było 75 polsko-angielskich dzieci dwujęzycznych w wieku 4,5 – 7 lat mieszkających w Wielkiej Brytanii i grupa dobranych 75 polskich dzieci jednojęzycznych mieszkających w Polsce. Badano użycie terminów wyrażających stany wewnętrzne w dziecięcych opowiadaniach tworzonych na podstawie zestawu obrazków. Kodowano trzy rodzaje terminów: emocjonalne, mentalne i percepcyjne. Dzieci proszone były również o ponowne opowiedzenie historyjki od razu po wysłuchaniu wersji modelowej i o odpowiedzenie na pytania dotyczące rozumienia historyjki, które skupiały się na stanach wewnętrznych postaci. Tym sposobem badano efekt modelowania na użycie IST i porównano produkcję IST podczas opowiadania historyjki i w stosunkowo bardziej interakcyjnym kontekście rozmowy o stanach wewnętrznych. Dodatkowo mierzono zasób słownictwa dzieci, ich zdolności rozumienia struktur gramatycznych w obu językach, oraz refleksję nad myśleniem.

Wyniki wskazały, że choć dzieci dwujęzyczne osiągają niższe wyniki w testach językowych od jednojęzycznych rówieśników, to obie grupy nie różnią się ilością IST użytych podczas opowiadania historyjki. Jednakże dzieci dwujęzyczne osiągnęły wyższe – niż dzieci jednojęzyczne – wyniki w teście teorii umysłu. Te dwie różnice w rozwoju dzieci mogły się zniwelować, prowadząc do braku różnic między grupami w ilości użytych IST. Ponadto, dzieci dwujęzyczne używały IST podobnie w obu swoich językach. Wyniki pokazały również, że prezentowanie dziecku historyjki modelowej i pytanie o stany wewnętrzne bohaterów uczuła dzieci na wiedzę, przekonania i pragnienia postaci, co prowadzi do zwiększenia – względem opowiedzianych historyjek – użycia IST w ich ponownie opowiedzianych historyjkach oraz w odpowiedziach na pytania dot. rozumienia historyjki.

Słowa kluczowe: słownik terminów wewnętrznych, terminy mentalne, stany mentalne, teoria umysłu, narracje, dzieci dwujęzyczne, rozwój językowy dzieci

List of contents

Acknowledgments	2
Abstract	3
Streszczenie	4
List of contents	5
Introduction	7
Key terms used in the present thesis	11
List of acronyms used in the present thesis	12
Chapter 1: Internal state terms – methodological issues	13
1.1. Internal state terms – most relevant characteristics	13
1.1.1. Internal state terms describe abstract processes.....	14
1.1.2. The meanings of internal state terms are often polysemous.....	14
1.1.3. Internal state terms as a small portion of the overall lexicon	15
1.2. Classification of internal state terms	17
1.2.1. The subclasses	17
1.2.2. The specific items	21
1.3. How has ISL been investigated.....	24
1.3.1. Interactional context of testing the ISL production	24
1.3.2. Non-interactive settings of testing the ISL production.....	27
1.3.3. Testing ISL comprehension.....	31
Chapter 2: Internal state terms – a developmental perspective	33
2.1. The building blocks: role of internal state terms in child development.....	33
2.2. Laying the foundations of internal state lexicon: development of thoughts about self and others	34
2.3. From the ground up: the role of interaction in the construction of social understanding.....	37
2.4. The development of internal state language in children	42
2.5. The Internal State Lexicon in relation to theory of mind performance and language	50
Chapter 3: Bilingual children	55
3.1. The bilingual speaker: theory of mind and the bilingual language acquisition	55
3.1.1. Theory of Mind development in bilingual children.....	57
3.1.2. Vocabulary skills in bilingual children.....	63
3.1.3. Narrative skills in bilingual children	66
3.2. Internal State Lexicon in bilingual children.....	68

Chapter 4: The current study – analysis of ISL in bilingual and monolingual children’s narratives	71
4.1. Pending questions about ISL in bilingual children	71
4.2. Methodological presumptions of the present analysis	75
4.3. The methods	77
4.3.1. The aims and the research questions of the current study	77
4.3.2. The hypotheses	78
4.3.3. The source study	82
4.3.4. Participants	83
4.3.5. General procedure and tools	86
Chapter 5: The results	96
5.1. Introductory analyses of the bilingual sample	97
5.1.1. Bilingual and monolingual comparison	97
5.1.2. Bilingual comparison in Polish and English	99
5.2. Main analyses related to ISL	101
5.2.1. RQ1: Do bilingual and monolingual children differ in the amount of internal state terms used when telling a story in Polish?	103
5.2.2. RQ2: Do bilingual children differ in the amount of internal state terms used when telling the stories in Polish and English?	108
5.2.3. RQ3: What are the best predictors of the use of internal state terms in the narratives told in Polish and English?	114
5.2.4. RQ4: Do children differ in the amount of internal state terms when telling a story, as compared to when they retell a story after listening to a model story told by an adult?	132
5.2.5. RQ5: Do children use more internal state terms when explicitly engaged in a conversation about the mental states of story protagonists, as compared to the amount of internal state terms in their spontaneous telling?	138
5.3. The summary of the results	146
Chapter 6: The discussion	150
Study limitations	162
Further directions	164
General conclusions and practical implications	167
References	170

Introduction

Internal state terms, such as “think”, “want”, “notice”, “surprised”, “scared”, indicate our ability to perceive the psychological disposition of self and others and they shape our cognitive and social development. They constitute what will be termed as the Internal State Lexicon (ISL). The present thesis investigates the use of internal state terms in fictional narratives told by pre- and early-school children aged 4.5 – 7 years old. The focus of the thesis will be on Polish-English bilingual children and their use of the ISL in story-telling. There is an evident lack of research on the ISL in bilinguals, even though their general linguistic and cognitive development is comprehensively examined. Here, the specific aims related to the bilingual population are to investigate:

(1) whether language status (bilingual vs. monolingual) may influence the use of the Internal State Lexicon (ISL); (2) whether the ISL develops at a comparable level in both languages of the bilingual child; (3) what are the predictors of the ISL production in the narratives in Polish and English. The bilinguals’ performance will be explored across their two languages and compared to that of carefully matched Polish monolinguals peers.

The present thesis investigates the use of internal state terms in child-made narratives based on a set of pictures. Though the terms are associated both with the lexical and theory of mind development, the ISL is neither the content of typical lexical tests nor the focus of false-belief tasks typically used to measure theory of mind. Rather, the internal state terms are often investigated in spontaneous speech or a relatively more structured context, such as story-telling. Two further aims of the present thesis focus on the nature of the used tool, a narrative. These two aims investigate whether: (4) the ISL production in a narrative context can be momentarily improved by presenting a child with a model story told by an adult; (5) whether the child’s ISL use when telling a story differs from the ISL use when specifically asked about story protagonists internal states.

The thesis is organized as follows. The first chapter explores the qualities of internal state terms: their characteristics (e.g. abstractness, polysemy), their typical classifications into subclasses and the specific terms included in those subclasses. The abstract and polysemous nature of internal state terms has important consequences for their acquisition. Chapter 2 presents an overview of the acquisition of the ISL and links this to the developing lexicon and theory of mind in children. Since the focus of the present thesis is on the use of the ISL in bilingual children, Chapter 3 discusses the differences between bilingual and monolingual language acquisition and explains how these differences influence the bilingual use of internal state terms. Subsequently, the research questions and the method of analysis are presented in Chapter 4. The results of the statistical analyses performed to answer the research questions are reported in Chapter 5. Chapter 6 provides a discussion of the results, together with practical implications for parents and practitioners working on language development in bilingual children.

The present thesis is a secondary analysis of existing data. The data used here was gathered largely in the Bi-SLI-Poland project “*Cognitive and language development of Polish bilingual children at the school entrance age - risks and opportunities*” (2010 – 2015, PIs: dr hab. Ewa Haman and dr hab. Zofia Wodniecka). The project was carried out at the Faculty of Psychology, University of Warsaw in collaboration with the Institute of Psychology, Jagiellonian University. The project was supported by the Polish Ministry of Science and Higher Education / National Science Centre (Decision 809/N-COST/2010/0) and in part by the Foundation for Polish Science subsidy to dr hab. Zofia Wodniecka. Data collection, data coding and maintenance were also partly supported by 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*” (2014 – 2016, PI: dr hab. Agnieszka Otwinowska-Kasztelanic), henceforth WLRB, carried out at the Faculty of Modern Languages, University of Warsaw. The projects were also linked to the

European COST Actions IS0804 “*Language Impairment in a Multilingual Society: Linguistic Patterns and the Road to Assessment*” (2009 – 2013) and IS1306 “*New Speakers in a Multilingual Europe - Opportunities and Challenges*” (2013 – 2017).

Both studies, i.e. the Bi-SLI-Poland and WLRB, were large-scale projects that gathered scientists from the fields of linguistics, psycholinguistics, and developmental psychology. The author of the present thesis was actively involved in both projects at many levels of study continuation: as a project manager and a secretary for both projects the author was largely involved in the recruitment of participants into the study, planning and management of data collection, and the administrative side of the projects. Then, as a starting PhD student, the author participated in the data transcription and coding, data cleaning and in conducting the preliminary analyses. Finally, the author was also engaged (together with the research teams from UW and UJ) in the final data analyses, data dissemination (via academic publications, conference presentations, and meetings with practitioners), and in writing the final reports for the funding bodies. The author is still continuing her involvement in the process of analyzing the data and preparing publications going far beyond the initial goals and time frame of the projects. This thesis is one of the examples of this endeavor.

Part of the data analysed here was also used in a published article by Otwinowska, Mieszkowska, Białecka-Pikul, Opacki, Haman (2018) of which the author of the present thesis is one of the two first authors. Specifically, data from 131 of the 150 children (87%) analysed here were also analysed in the Otwinowska et al. (2018) article. However, the two papers differ in the subject of their analysis. The article examined the differences between bilingual and monolingual stories told spontaneously and stories retold after a model story. It investigated whether retelling might improve bilingual and monolingual stories to the same extent. The analyses published there focused on two areas of narrative quality: the macrostructure (i.e. general coherence of a story) and microstructure (e.g. type-token ratio). Finally, the article

included the use of internal state terms as one of the measures of narrative macrostructure and considered only the overall number of internal state terms, without distinguishing between the different subclasses. Thus, the two analyses might be considered discrete.

Key terms used in the present thesis

Short definitions of key terms used in the present thesis are presented below. Those key terms are often diversely defined and interpreted, hence the following compendium was prepared to explain how these terms are understood in the present thesis.

Internal State Lexicon (ISL) – the inventory of internal state terms known by an individual; the ISL is the focus of Chapter 1.

Internal state terms (IST) – terms that refer to knowledge, belief, desires, emotions and perception states.

Narratives – in the present thesis, narratives are fictional stories told by children and based on a set of pictures; for information on narrative abilities in bilingual children, see Subsection 3.1.3; for a detailed description of the tool used to elicit narratives, see Subsection 4.3.5.

Narrative macrostructure – in general, the ability to tell a coherent story. More specifically, macrostructure is related to the concept of story structure which includes referring to the setting (time and place), and the episode structure. The episode structure is composed of an initiating event that triggers the protagonist's response, internal responses (the internal state of the protagonist as a response to the initiating event), the goal of the protagonist, the attempt to reach the goal, and the outcome of the undertaken actions (Gagarina, 2016; Stein & Glenn, 1975). For more details, see Subsection 3.1.3.

Bilingual – in the present thesis, the bilingual participants are children of immigrants who acquire two languages simultaneously and from early age. They acquire one language at home (here: Polish, home language, spoken by at least one of the parents) and a different language is spoken by the community around them (here: English, majority language).

Monolingual – in the present thesis, the monolingual participants are children who acquire one language from birth. Specifically, these are Polish children of Polish parents living in Poland. Thus, the language that they use for home and outside of home communication is the same.

List of acronyms used in the present thesis

BPVS-3	British Picture Vocabulary Scale – 3 rd edition (Dunn, Dunn, & Styles, 2009), a picture recognition test used to measure receptive vocabulary size in English;
EVT-2	Expressive Vocabulary Test – 2 nd edition (Williams, 2007), a picture naming test used to measure expressive vocabulary size in English;
ISL	Internal State Lexicon;
IST	Internal state terms;
L1	The first language (in the order of acquisition) of a bilingual child;
L2	The second language (in the order of acquisition) of a bilingual child;
OTSR	Obrazkowy Test Słownikowy (Haman & Fronczyk, 2012), a picture recognition test used to measure receptive vocabulary size in Polish;
ToM	Theory of mind;
TNW	Total number of words;
TRM	Test of Reflection on Thinking (Białęcka-Pikul, 2012), a test used to measure theory of mind development (more specifically: the development of reflection on thinking);
TROG	Test of Reception of Grammar – (in English: TROG-2, Bishop, 2003; in Polish: TROG, translation by Smoczyńska, 2008, unpublished), a picture recognition test used to measure receptive grammar skills;
ZNO	Zadanie Nazywania Obrazków (Haman & Smoczyńska, 2010, unpublished), a picture naming task used to measure expressive vocabulary size in Polish;
CDI	MacArthur-Bates Communicative Developmental Inventory (Fenson et al., 1993, 2007).

Chapter 1: Internal state terms – methodological issues

The present chapter will focus on the important methodological issues related to the testing of the Internal State Lexicon (ISL). It will start by presenting the characteristics of internal state terms that bear consequences on both the acquisition and investigation of the ISL (Section 1.1.). Next, it will present the most and the least commonly investigated subclasses and prototypical terms, based on the author's review of studies that investigated the ISL (Section 1.2.). The next section will discuss the tools and procedures used to investigate the ISL in child language comprehension and production in either interactional or non-interactional contexts (Section 1.3.).

1.1. Internal state terms – most relevant characteristics

Our lexicon includes many terms that demonstrate our understanding of the fact that people's actions are governed by internal states such as beliefs, desires, and emotions. These words are referred to as internal state terms, or mental state terms (see e.g. Nielsen & Dissanayake, 2000; Symons, 2004). The specific terms known by an individual, i.e. the inventory of the terms, is the Internal State Lexicon (Frank & Hall, 1991). It may include different subclasses, from those referring to strictly *mental* processes (e.g. “know”, “imagination”, “surprised”, see e.g. Bretherton, McNew, & Beeghly-Smith, 1981; Shatz, Wellman, & Silber, 1983) and *desires* (e.g. “want”, “dream”, “envy”, see e.g. Jenkins, Turrell, Kogushi, Lollis, & Ross, 2003; Ruffman, Slade, & Crowe, 2002) to those referring to *emotions* (e.g. “love”, “fear”, “angry”, see e.g. Bretherton et al., 1981; Dunn, Bretherton, & Munn, 1987). Other subclasses include states referring to (*moral*) *obligation* (e.g. Miranda, Baixauli, & Colomer, 2013), *perception* (e.g. Hall & Nagy, 1979), and *physiological needs* (e.g. Rumpf, Kamp-Becker, Becker, & Kauschke, 2012).

Let us first discuss those characteristics that are most relevant to the acquisition and testing of internal state terms in children. These are of three kinds. First, internal state terms describe abstract processes. Second, they are often polysemous in nature. Third, they constitute a small share in the child's language input and production.

1.1.1. Internal state terms describe abstract processes

First, internal state terms by definition refer to psychological experiences and processes which are generally abstract, complex, and not clearly visible in behaviour (Dunn & Brophy, 2005; Hall & Nagy, 1979). The internal state terms describe “[...] things that people experience, rather than what they are, in the long-term sense of the word” (Hall & Nagy, 1979:13, original emphasis). Some of these experiences may have behavioural manifestations, e.g. desire may be manifested by reaching for an object to request it. When the child reaches out for an object, the mother's common response is to confirm the child's request by directly referring to the internal state (e.g. “You want the ball, yes?”) and to hand the object (Slaughter, Peterson, & Carpenter, 2009). This might help the child to map the term “want” onto this behaviour. Nonetheless, most internal state terms are characterized by low imageability and sensory richness, i.e. they are less likely to bring to mind a sensory image, as compared to concrete nouns and verbs. To add to the complexity, some of the internal state terms describe a single process that may be named in more than one way (e.g. “teach” vs. “learn”, depending on the direction of the action, Gleitman, 1990). That is one of the reasons why internal state terms pose a conceptual challenge to young children (Thompson, 2006).

1.1.2. The meanings of internal state terms are often polysemous

Many of the internal state terms are polysemous in nature, i.e. have several meanings that signify different internal processes. As an example, “I think that she's angry” refers to an opinion, while “I'm thinking about buying a car” implies a mental process of consideration (for more examples see Naigles, 2000; Shatz, Wellman, & Silber, 1983). Here, both uses of the

word “think” carry an internal, or mentalistic, meaning. However, not all uses of an internal state term may be considered to refer to the internal state. Hall and Nagy (1979:86) discuss the two meanings of the perceptual term “see”: that of “perceiving visually” and that of “visiting” (e.g. “He went to see his grandmother”). According to Hall and Nagy, only one of the two meanings – the former – can be regarded as an instance of an internal state term. Similarly, many researchers (e.g. Jenkins, Turrell, Kogushi, Lollis, & Ross, 2003; Nielsen & Dissanayake, 2000; Shatz et al., 1983) point to the fact that some internal state terms may appear in speech (of children and adults) as highly conventionalized conversational devices, i.e. without a clear mentalistic meaning: e.g. “You know, there might be some cake left” or “The ball, I mean, the car”. Hall and Nagy (1979) call such use of internal state terms as “pragmatic usage” and note that these might or might not be mentalistic, depending on the sentential context. They identify seven categories of such pragmatic uses: *conversational devices* (e.g. “You’re not serving the children, you know”), *indirect requests and suggestions* (e.g. “Do you want to take out the garbage, please?”), *rhetorical questions* (e.g. “Do you know what happened to Harry?”), *exam questions* (e.g. “Do you know who discovered America?”), *hedges* (e.g. “He’ll go on holidays, I suppose”), *opinion questions* (e.g. “Don’t you think we should decide what we’re going to do?”), and *attentional devices* (e.g. “Look what I did!”). They suggest that each of these examples need to be carefully considered when deciding about actual meaning and function of a given instance in child/child-directed speech.

1.1.3. Internal state terms as a small portion of the overall lexicon

Research on lexical development in children, specifically in the early childhood, has focused on the child’s acquisition of words that refer to easily imaginable objects (e.g. “apple”) and actions (e.g. “write”) that can be presented to children on picture boards to find and name (Clark, 1993, 2016, 2017). Accordingly, these have a large share in the lexical tests designed to measure the size of children’s lexicon. Comparatively little research has investigated children’s

acquisition of terms that refer to internal states that are not only less imagineable but also less common both in adult or child spontaneous speech. When it comes to adult speech towards children, Adrián, Clemente, and Villanueva (2007) found that English speaking mothers' use of cognitive terms when narrating a wordless picture story constituted up to 1.2% of overall mothers' talk (see also Slaughter, Peterson, & Mackintosh, 2007 for similar findings). As for the amount of internal state terms produced by children, the author of the present thesis searched for internal state terms in the Word Birth Browser (an interactive application that includes data of language produced by a single child from birth till the age of two, see Roy, Frank, DeCamp, Miller, & Roy, 2015). The part of the corpus available online in the Word Birth Browser includes 679 different lexemes produced by the child till the age of 2. Out of 679 different lexemes found in the corpus, 35 were internal state terms. This implies that internal state terms constituted approximately 5% of all the lexemes used by this particular child till the age of 2. Similarly, Bartsch and Wellman (1995) calculated the use of cognitive terms in spontaneous speech of English-speaking toddlers. They found that at the age of 2.5 to 3 years old, 5% of the children's utterances contained cognitive terms. This tendency seems to be language universal. Tardif and Wellman (2000) found a similar proportion in a slightly younger sample of Chinese toddlers: at the age of 27 months (2;3), 5% of all their utterances contained a cognitive term. Similarly, Pascual, Aguado, Sotillo, and Masdeu (2008) studied longitudinally Spanish-speaking children (tested from the age of 3 to 5 years) and found that only 3% of their utterances contained cognitive terms. Thus, internal state terms seems to constitute a relatively small portion of both the child's input, and the child's overall lexicon, possibly across contexts and languages, at least at the age below 5 years.

Having a small share in the child's overall lexicon, the ISL is also of interest to few. Many studies on the ISL are done by researchers studying Autism Spectrum Disorders (ASD) with the aim to compare the use of internal state terms in ASD populations and typically

developing controls (Capps, Losh, & Thurber, 2000; Losh & Capps, 2003) or children with developmental or language impairments (Capps et al., 2000; Miranda et al., 2013; Norbury & Bishop, 2003; Rumpf et al., 2012; Tager-Flusberg, 1992). The studies that have focused on the production of internal state terms in a narrative context have shown that children with ASD may use fewer cognitive terms than typically developing peers (Rumpf et al., 2012), or age- and language-matched children with Down syndrome (Tager-Flusberg, 1992). However, children with ASD and developmental delays seem to refer to emotional states with the same frequency as typically developing controls, though they tend to simply label the protagonists' emotions, without identifying causes for internal states (Capps et al., 2000). Thus, studies on the atypical populations have highlighted the importance of investigating various subclasses of internal state terms.

1.2. Classification of internal state terms

Studies of the Internal State Lexicon differ in regard to the choice of the investigated subclasses. Some may examine only one subclass, e.g. cognitive/mental or emotion terms. Others may include such rarely-studied subclasses as linguistic (e.g. “say”, “call”, “warn”) or behavioural/trait terms (e.g. “brave”, “funny”), see Dyer-Seymour, Shatz, Wellman, & Saito, (2004). Another matter is the choice of the individual terms under investigation: some studies focus on the cognitive verbs only, others will include a range of word categories and even phrases (e.g. verb or noun phrases). Yet others will exclude the conversational mentions of internal state terms. We will now discuss the different approaches to the classification of the Internal State Lexicon.

1.2.1. The subclasses

The most common subclasses of internal state terms are the cognitive/mental terms, desire/volition terms, and emotional terms. A review, carried out by the author, aimed to bring together a list of internal state terms coded in different studies of child-adult interaction (with

typically developing children). The author found 46 studies published online. The search was done in the Google Scholar and included keywords such as: mental state language, mental state talk, internal state language, metacognitive terms, mental terms, mental verbs, emotion state language, feeling states. The search yielded 46 studies that were found to be relevant. The studies were published between 1979 and 2017. Out of the 46 studies, 24 appended or included a whole or substantial list of words coded as internal state terms. However, 4 of the 24 studies were excluded from the review list because they used a CDI questionnaire with terms coded originally by Bretherton, McNew, Beeghly-Smith (1981) which was already included in the review list. Another 4 studies were excluded from the review list because they investigated the use of several terms only, e.g. “want”, “think” and “know”, or “remember”, “know”, and “guess”. One study was excluded because the list was composed of internal state terms used by an ASD population, not of interest to the present study. Finally, 3 studies were excluded because they investigated children’s comprehension of terms (through forced-choice format), and thus the terms included in the lists were not derived from children’s speech. Appendix A contains details on all the 46 studies, together with reasons for including/excluding it from the final review list.

The review list (see Appendix A) included 12 studies that appended at least a substantial list of terms coded as the ISL, investigated typically developing child populations and derived their terms from children’s speech. Eleven of the studies were conducted in English, and one study investigated the ISL in Spanish, but provided English translations of the terms, which enabled their inclusion in the list (Pascual et al., 2008). Out of the 12 studies that provided a list of words, 9 studies coded internal state terms in spoken child-adult interaction (Bretherton et al., 1981; Dunn, Bretherton, & Munn, 1987; Jenkins et al., 2003; LaBounty, Wellman, Olson, Lagattuta, & Liu, 2008; Nielsen & Dissanayake, 2000; Pascual et al., 2008; Ruffman, Slade, & Crowe, 2002; Shatz et al., 1983a; Zevenbergen, Haman, & Zevenbergen, in review).

Additionally, 2 studies (Dyer, Shatz, & Wellman, 2000; Dyer-Seymour et al., 2004) listed examples of terms used in children's books (English and Japanese, altogether 130 books for 3-, 4-, and 5- year olds), and one study (Hall & Nagy, 1979) was a theoretical review providing an extensive list of words that, according to the authors, could be considered as ISL. The review attached in the Appendix A includes also a column with terms coded as the ISL in the present thesis.

Cognitive terms are the focus of the majority of studies that investigate the ISL. The review showed the cognitive/mental terms as the most commonly studied subclass of the ISL – they were investigated by all 12 studies (though Dunn and colleagues, 1987 refer to those as “states of consciousness”). The cognitive terms included in the review list generally refer to the mind, imagination, and metacognition. The most commonly studied examples (i.e. those appearing in at least seven out of the twelve studies) included the following words: “mean”, “dream”, “forget”, “understand”, “guess”, “know”, “remember”, “think”, “wonder”, “pretend” (for the full list of terms, see Appendix A). The second most common subclass of the ISL were emotional terms, coded in 9 out of 12 studies. The most commonly studied examples (i.e. appearing in at least 6 out of the 9 studies) were: “feel”, “happy”, “love”, “surprised/surprising”, “afraid”, “angry”, “like”, “sad”, “scared/scary”, “upset” (for the full list of terms, see Appendix A). Desire/volition subclass came third, coded in 7 out of 12 studies. The most commonly studied examples (i.e. appearing in at least 4 out of 7 studies) were: “hope”, “wish”, “want”, “would like to”, “need” (for the full list of terms, see Appendix A). Other less common subclasses referred to: morality, obligation and moral evaluation (included in 3 studies) with terms such as “have to”, “must”, “horrible”; physiological terms (included in 3 studies) with key terms such as: “hungry”, “sleepy”, “thirsty”, “tired”; and perceptual terms (included in 2 studies) with key terms such as: “hear”, “look”, “notice”, “see”. Of interest are the terms referring to perception. They are not that widely studied, but the very act of seeing seems to

guide children's early knowledge attribution. In fact, from early on, children attribute the knowledge or lack thereof based on whether a person saw or did not see something happen. For example, Surian, Caldi, and Sperber (2007) found that 13-month old infants expected the agent to correctly infer the location of the object only when the agent had seen the object being placed, and were surprised (as evinced by longer looking times) when the agent correctly "guessed" the location without having seen where the object was placed. Indeed, most of the false-belief tasks used to measure theory of mind in children (i.e. the ability to attribute mental states to oneself and others and to predict people's behaviour on the basis of their mental states, Astington & Jenkins, 1995) are based on the protagonists either having seen something done or not, e.g. observing a change of location, seeing the real contents of a box, perceiving another figure in a picture turned upside down (e.g. Bartsch & Wellman, 1989; Gopnik & Rosati, 2001; Lewis & Osborne, 1990; Wimmer & Perner, 1983). Consequently, children may use both cognitive and perceptual terms to refer to the other's state of knowledge.

However, physiological or perceptual terms seem to include many ambiguous cases. Hall and Nagy (1979) discuss a few such cases, e.g. the two meanings of "see" (discussed above) or "ache" and "hurt" which describe sensory, rather than internal (or psychological) experiences. Words such as "tired", "awake", "cold", "hungry" or "thirsty" are even more disputable. Hall and Nagy (1979) talk about such cases as "not clearly categorizable" and advocate for careful considerations when such words are to be coded, but they do – on the whole – include such items in the broad category of internal state terms (1979:12). Since then, the items have entered the coding schemas in several studies (e.g. "hungry"/ "thirsty": (Bretherton et al., 1981; Dunn et al., 1987; Gagarina et al., 2012; Kay-Raining Bird, Cleave, Curia, & Dunleavy, 2008; Miranda et al., 2013; Pinto, Tarchi, & Bigozzi, 2016; Ruffman et al., 2002; Rumpf et al., 2012); "cold"/ "hot": (Bretherton & Beeghly, 1982; Bretherton et al., 1981; J. Dunn et al., 1987; Kay-Raining Bird et al., 2008; Pinto et al., 2016) .

Due to their polysemous nature, a clear-cut classification of some terms into appropriate subclasses may be problematic. Even in the review list described here, it happened that different researchers coded specific terms as belonging to different subclasses (see Appendix A). For example, the term “hope” was considered as a cognitive term in 4 studies (Hall & Nagy, 1979; Nielsen & Dissanayake, 2000; Shatz et al., 1983; Zevenbergen et al., in review), and as a desire/volition term in 5 studies (e.g. Bretherton et al., 1981; Dyer et al., 2000; Dyer-Seymour et al., 2004; LaBounty et al., 2008; Ruffman et al., 2002). Sometimes the sentential context in which the term is used might determine its category. A good example is the classification of the term “feel”, which, according to Hall and Nagy (1979:21), can be categorized as an emotion term, a cognitive term, or a perceptual term. For example, in “how would you feel if they said that to you?” “feel” is an emotion term. In the sentence “I don’t feel he can handle the job” “feel” is a cognitive term. Finally, in a sentence “It felt warm and soft”, “feel” is a perceptual term. The many uses of a word, and consequently, distinctive meanings, might be the main reason why researchers sometimes classify the terms differently. In fact, out of 441 items in the review list (see Appendix A), 48 items were categorized differently in at least 2 out of 12 studies, and 9 items were categorized differently in 3 out of 12 studies. This means that altogether 13% of the items were categorized into different subclasses across studies. Thus, a divergent categorization of items is not that uncommon and should be a matter of researchers’ careful examination.

1.2.2. The specific items

Most studies of internal state terms focus on verbs, and most verbs refer to cognitive processes e.g. “know”, “think”, “guess”, “pretend”. In fact, the review list (see Appendix A) allows us to sketch some patterns, for example, most cognitive terms are verbs (58%), while most emotion terms are adjectives (46%). However, these are just some trends, and a specific syntactical category of a word (e.g. verb) cannot definitely determine its subclass (e.g.

cognitive). As Hall and Nagy (1979) note: “There are some syntactic categories that will be typical of internal state words, but there are no syntactic criteria that will determine whether or not a word belongs in this class” (p. 8). What Hall and Nagy imply is that no syntactical clues should be followed to categorize a specific word (e.g. a noun) as a particular subclass, e.g. a cognitive, emotional, or a perceptual term.

The review list (see Appendix A) showed that verbs constituted 178 out of 441 (40%) of internal state words studied. Additionally, there were altogether 34 (8%) words that might be either nouns or verbs in English (e.g. “wish”). The 57 nouns constituted 13% of all words (441) in the list. The list also contained 156 adjectives (35%) and 14 adverbs (3%). The list included some terms that could not be unequivocally assigned to a single syntactic category, e.g. “hug”, was used both as a verb and a noun, “pretend” was found to be used both as a verb and as an adjective. Moreover, there were idioms and phrases, e.g. “pay attention”, or “slip one’s mind”. These were relatively infrequent compared to more common words and they constituted 3% of the sample. Still, as noted by Hall and Nagy (1979:8) such phrases should not be underestimated and left out of the classification as they are a significant part of lexical resources of a language.

The review list included over 7 categories altogether: cognitive terms, desire/volition terms, emotion terms, moral/obligation/evaluation terms, perceptual terms, physiological terms, and traits (sometimes referred to as behavioural terms). The most common subclasses of ISL included:

- *emotion terms*, altogether 214 terms: 29% were verbs, 13% were nouns, 46% were adjectives; 3% were adverbs. Additionally, 9% of words could be coded either as nouns or verbs (not specified by the authors). The most common terms were: “feel”, “happy”, “love”, “surprised/surprising”, “afraid”, “angry”, “like”, “sad”, “scared/scary”, “upset”;
- *cognitive terms*, altogether 132 terms: 58% were verbs, 16% were nouns, 16% were adjectives, 3% were adverbs. Additionally, 6% of words could be coded either as

- nouns or verbs (not specified by the authors). The most common terms were: “mean”, “dream”, “forget”, “understand”, “guess”, “know”, “remember”, “think”, “wonder”, “pretend”;
- *desire/volition/modality terms*, altogether 27 terms: 59% were verbs, 11% were nouns, 11% were adjectives; 7% were adverbs. Additionally, 11% of words could be coded either as nouns or verbs (not specified by the authors). The most common terms were: “hope”, “wish”, “want”, “would like to”, “need”;
 - *perceptual terms*, altogether 38 terms: 14 verbs, 3 nouns, 18 adjectives. Additionally, 3 words could be coded either as nouns or verbs (not specified by the authors). The most common terms were: “hear”, “look”, “notice”, “see”;
 - *moral/obligation/evaluation terms*, altogether 15 terms: 6 verbs, 2 nouns, 6 adjectives, 1 adverb. The most common terms: “have to”, “must”, “horrible”.

In short, internal state terms are classified into different subclasses: the most commonly studied subclasses include cognitive/mental terms (referring to the mind, imagination, and metacognition), desire/volition terms (referring to desires but also needs), and emotional terms (referring to both positive and negative emotions). Fewer studies investigate terms that refer to morality, obligation and moral evaluation, or physiological and perceptual terms. However, internal state terms are often polysemous and hence, clear-cut and unanimous classification into subclasses may be problematic. In fact, it may happen that a single term is categorized by researchers as belonging to different subclasses, based on the sentential context and the context-specific meaning. Last, many of the internal state terms are verbs and nouns, while comparatively fewer terms are adjectives and adverbs. Thus, internal state terms form an open and very heterogeneous group of words. Being also polysemous and abstract, these words provide a methodological challenge to researchers who want to study ISL. Let us turn now to the different methods used to investigate ISL in children.

1.3. How has ISL been investigated

The development of the ISL cannot be accurately tested neither by the tools used to measure vocabulary size in children, nor by the classic false-belief tasks, since neither of them focus on eliciting internal state terms. As cooperative social interaction is the main source of the ISL acquisition, it is also the most instinctive context of testing the ISL production and comprehension. The production of the ISL can be investigated in interactional settings (mostly naturalistic, e.g., spontaneous speech, free-play interaction, in shared book-reading or other kinds of everyday home interaction) or comparatively non-interactional settings (often also structured, e.g. tasks that require the child to tell a story based on a specific book, tasks that elicit picture description or person/protagonist description). The present section gives an overview of various procedures used to measure the ISL production in the two contexts (interactional vs. non-interactional) and also briefly presents some of the ways to measure the ISL comprehension in children.

1.3.1. Interactional context of testing the ISL production

Initially, interactional and naturalistic settings of the ISL investigation, such as collecting speech samples or coding live unstructured interactions, were the primary source of knowledge about the development of the ISL. For example, Hall and Nagy's (1979) report on the theoretical issues in the investigation of the ISL was based on 300 hours of recorded conversations of individuals from diverse social and ethnic backgrounds. Shatz, Wellman and Silber (1983) were the first to draw a timeline of early appearance of cognitive/mental terms. In a case-study, they investigated a single child longitudinally from the age of 2;4 to 4;0. Similarly, Bretherton and colleagues (1981) gathered a corpus of children's spontaneous use of the ISL with mothers, which served as a basis for the Internal State Language Questionnaire (ISLQ, Bretherton & Beeghly, 1982; Bretherton et al., 1981), a 78-word checklist for assessing internal state terms production in 30-month old toddlers.

The interactional contexts most commonly employed in research include spontaneous speech in children, or in child-adult and child-peer interactions at home, shared book-reading between the child and the parent or the child and the teacher, and child-made stories. We will now look closely at each of those contexts.

Naturalistic observations of everyday home interactions, e.g. between the child and his/her parent, or between siblings, yield results that are most representative of children's real-life environment. The proportion of internal state talk might be low in such contexts (approximately up to 5%, see Bartsch & Wellman, 1995; Slaughter et al., 2007; Tardif & Wellman, 2000), but for this reason the observation time is usually relatively longer than in other settings (e.g. book-reading). However, there is a large variation in the time of observation: in some studies, the data is gathered in a few sessions each lasting an hour or longer, and the samples are of c.a. 30-50 participants (Brown & Dunn, 1992; Dunn, Brown, & Beardsall, 1991; Jenkins et al., 2003). Notably, in a study by Jenkins et al. (2003), each of the 37 families was recorded altogether 12 times (divided between two time points), which amounted to the total of 18 hours of recordings from each family. In some cases, especially in smaller samples, researchers leave the recording device in the house for several days. For instance, Kay-Raining Bird et al. (2008) did a case-study of internal state talk directed to an autistic child and recorded home interaction for three days (from Saturday to Monday). Importantly, during the recordings, the family members are explicitly encouraged to carry out with their usual daily routine, and the experimenters (if present) do not participate in the family affairs, and respond as little as possible to the comments of family members. In a study that received much publicity, Deb Roy collected audio and video recordings from an immediate linguistic environment of a single, typically developing child. The data was gathered in different rooms in the child's home, from the child's birth to the age of 3 years old, adding up to more than 200,000 hours of recordings (Roy et al., 2015). Such naturalistic, interactional settings yield large amounts of data, which

significantly lengthens the transcription and coding processes. In a now classic longitudinal study by Hart and Risley (1995), which did not focus on the ISL but on the child-directed speech in general, the recorded casual interactions in 42 families amounted to more than 1 300 hours in total – consequently, it took 6 years of data transcription, coding and analysis before the first results were attained (Hart & Risley, 2003).

As many significant factors related to the naturalistic settings remain uncontrolled, and such interactions may vary widely from family to family, many researchers turn to more structured contexts and use a prompt to elicit internal state talk. To maintain the naturalistic character of the observation, the setting is often at home, but within the recorded session there is a short appointed time when a purposefully communicative task is introduced. For example, this may be a memory game, i.e. finding matching sets of cards (e.g. Howard, Mayeux, & Naigles, 2008). Other times, the researcher brings a box of selected toys and dressing-up materials (e.g. Hughes & Dunn, 1997). In such contexts, the interaction is still considered largely naturalistic, but aided by prompts to ensure there is some mentalistic exchange involved.

A much more common procedure is engaging the child and an adult in a shared book-reading (e.g. Adrián et al., 2007; Martucci, 2016), and story-telling (Adrian, Clemente, Villanueva, & Rieffe, 2005; Clarke-Stewart & Beck, 1999; Slaughter et al., 2007; Symons, Peterson, Slaughter, Roche, & Doyle, 2005). The book-reading or narrative context may be particularly rich in internal state talk. In fact, Sabbagh and Callanan (1998) showed that parents and their five-year-olds children who were jointly reading illustrated wordless picture books used more internal state terms than comparable parent–child pairs in the Bartsch and Wellman's (1995) sample of everyday conversations. Narratives may be a significant source of internal state talk for a few reasons. First, they involve numerous protagonists, each with their own knowledge and goals. This differing knowledge offers multiple perspectives on the same event and paints the landscape of consciousness, i.e. how a particular action is interpreted by the

protagonists (Bokus, 2000, 2004, 2013; Bruner, 1986). The landscape of consciousness is one of two dimensions of a narrative first described by Bruner (1986) and investigated in detail by Bokus (2000, 2004, 2013). The second dimension of a narrative is the landscape of action, a representation of the actual actions of the story protagonists. Most of the spontaneous stories of pre- and early school children already include the double landscape of narration (Bokus, 1998; Ligęza, 1998). In order to relate the landscape of consciousness, the story-teller must depict the protagonists as mental agents and make use of internal state terms to describe their knowledge, goals and beliefs (Bruner, 1986). Second, the activity of book-reading and story-telling is a relatively common, or at least familiar form of family interaction. This joint construction of a story offers an opportunity to describe and label internal states, ask questions, paraphrase and elaborate on the action. Moreover, if parents repeatedly encourage the child to label, describe or elaborate on a protagonist's internal states (e.g. through questions, prompts), it may become natural for the child to ask and answer such questions (Symons et al., 2005). Thus, parental questions and prompts may serve as a scaffolding that improves the child's understanding of a story and their own storytelling abilities. In fact, parental questions and encouragements were found to be positively related to the number of preschool children's contributions to co-constructed narratives (Zevenbergen, Holmes, Haman, Whiteford, & Thielges, 2016) and the children's event recall and verbal complexity of their own (retold) stories (Clarke-Stewart & Beck, 1999).

1.3.2. Non-interactional settings of testing the ISL production

Some researchers have signaled that investigating the ISL production in a naturalistic, interactional context may in fact imprecisely assess the cognitive competence. For example, Shatz and colleagues (1983) have pointed out that children younger than 4 years old may be yet unable to express mental reference due to still developing language skills. Meins, Fernyhough, Johnson, and Lidstone (2006) have found that many of the children's uses of internal state terms

during social interaction are likely to be mere responses to the interlocutor's focus on the mind or emotions. Thus, some researchers choose to employ non-interactive tasks, where the child's production is unaided and unobstructed by the experimenter. Such tasks include telling a story (spontaneous or based on a specific book or props), or describing a friend or a picture.

Using child-made narratives based on a specific book/pictures is a common technique to elicit not only internal state terms, but to study the mentalizing ability in general (Bialecka-Pikul, 2012; Charman & Shmueli-Goetz, 1998; Symons et al., 2005). In a typical procedure, a child is given a wordless picture book (often a commercially available children's storybook) or a set of pictures that constitute a story. The topics of the books/pictures vary, e.g. from child-relevant problems such as a first day of school, to a fictional story about adventures of a pet (e.g. a dog, or a frog). Regardless of the topic, the plot usually involves mistaken identity, or deception (e.g. Meins et al., 2006) and shows the protagonists' emotional reactions to events (e.g. Symons, 2004). The child may be asked to look at the book/pictures by themselves, and not to show any pictures to the experimenter. This is done in order to ensure the minimum intervention into the child's narrative (e.g. the procedure for MAIN in Gagarina et al., 2012). Another reason is to ensure that the child is motivated to tell a story, as Bokus (1978) showed that children tell more complex stories when their listeners cannot see (and therefore do not know) the story. Usually, the experimenter is allowed to encourage the child to tell the story, especially if the child seems shy at the beginning. However, while the child is telling the story, the experimenter does not offer any prompts and does not ask any questions. Oftentimes, the child's story is followed by scripted questions from the experimenter to check the child's comprehension of the plot.

The narrative task in a non-interactive context still shares some of the characteristics of the naturalistic setting: (1) it includes the two dimensions of a narrative (landscape of action and landscape of consciousness); (2) it is based on a familiar and well-established practice often

found in children's playtime routines, i.e. joint book-reading and story-telling; (3) it offers an opportunity to describe and label internal states. Importantly, the story-books and pictures may encourage the child to refer to the internal states of the protagonists, but it is also possible to tell a story or describe the pictures without any such references (Symons et al., 2005). Thus, such tasks should reveal variability of the responses, with some stories including many internal state terms to others including none.

Narratives are also used to train children's narrative abilities and their understanding of the internal states of the story protagonists (Lewis, Freeman, Hagestadt, & Douglas, 1994). The training is based on the idea that a model story told by an adult may draw children's attention to the internal states of the characters and encourage them to use more internal state terms. In a series of experiments by Lewis and collaborators (1994), 3-year old children who failed a standard false-belief task were familiarized with the events that comprised the task by listening to a story version of the task. The children were then asked to retell the story back to the experimenter. The results of the experiments showed that when given the opportunity to link story events into a coherent narrative, children exhibited an accurate understanding of the internal states of story protagonists. Taking a step further, a different analysis on a set of partially overlapping data, carried out by Otwinowska, Mieszkowska, Bialecka-Pikul, Opacki, and Haman (2018), established that when retelling a story after an adult model, both monolingual and bilingual children told more coherent stories and included more references to internal states. The children also showed an increased understanding of internal states of the story protagonists.

In some cases, the children are asked to compose their own stories, not aided by adults or props, or even suggested topics. Children choose their own story protagonists, subjects and plots. A strong advocate of this technique is Angeliki Nicolopoulou, who suggests that only the spontaneous narratives reflect children's actual narrative abilities, including the capacity for

viewing protagonists as mental agents (e.g. Nicolopoulou, 2002; Nicolopoulou, Cates, de Sá, & Ilgaz, 2014; Nicolopoulou & Richner, 2007).

To elicit internal state terms, some researchers use a series of questions (in a form of an interview), instead of a narration. An example of this is a “describe-a-friend” task (Grazzani & Ornaghi, 2012; Meins et al., 2006), in which children are asked scripted questions about a friend of theirs (Do you have a best friend? What is your best friend’s name? Can you describe [friend] for me? What do you like about [friend]? What sort of person is [friend]? Is there anything else you’d like to tell me about [friend]?). Such interviews may elicit both non-mentalistic comments (e.g. “he lives round the block to me”, or “she has freckles”, Meins et al., 2006:187), and comments that include internal states. Meins distinguished between mentalistic comments (referring to mental life and intellect, e.g. “he’s a clever person”, p. 187) and behavioral comments (referring to social interactions and traits, e.g. “he plays with me”, “he’s friendly”, p. 187).

The methods of eliciting IST described above are characterized by different features. Collecting spontaneous speech samples or coding live unstructured interactions are most representative of children’s real-life environment, but require large amounts of data in order to be truly indicative of the child’s mentalizing abilities. The non-interactive tasks like telling a story (spontaneous or based on a specific book or props) or describing a friend provide a small sample of the child’s speech, but the stimuli is meant to tap into mentalizing abilities and the child’s production is unobstructed. It is also important to note that although the non-interactive tasks use different means to direct the child’s attention to inner motives of other people, most of them tend to imitate spontaneous speech or talk. In the present thesis a narrative based on a set of pictures is considered as an optimal way of eliciting internal state terms. Most importantly, such a narrative is based on a familiar practice of story-telling, and offers an opportunity to describe and label internal states but does not create a situation where the child

can mimic the interlocutor's focus on the mind or emotions. Moreover, in the present study, children were asked to tell a story without showing the pictures to the experimenter. Such a practice ensures no intervention from the experimenter and provides a close to real-life motivation for the child to tell an unfamiliar story to the interlocutor.

1.3.3. Testing ISL comprehension

Just like the acquisition of general vocabulary, the acquisition of the full meanings of internal state terms is not an all-or-nothing process. First, the child uses a word with a limited understanding of it, but as their exposure to different uses of a word accumulates, they gradually achieve a better understanding of the word, as stated by Nelson's (1998) "use before meaning" hypothesis. Some researchers investigate the ISL acquisition in children by checking their understanding of small discrepancies between the full meanings of internal state terms. This is usually done by measuring children's comprehension of internal state terms. For example, Moore, Bryant, and Furrow (1989) designed a task in which children aged 3 to 8 years old were asked to determine the location of a hidden object. The object was placed in one of the two possible locations, and the only cues to the exact location were two conflicting statements involving internal state terms "know", "think", and "guess". Thus, two puppets would each give their statement as to the location of the object, e.g. "I guess it's in the blue" vs. "I know it's in the red box" and subsequently, the child would be asked to say where they think the object is (forced-choice format).

This paradigm inspired other researchers to create forced-choice tasks that would include more internal state terms. Astington and Pelletier (2004) designed Metacognitive Vocabulary Test (METVOC) for early school children to target the degrees of certainty (e.g. the difference between "know" and "guess"), or the variation in knowledge (e.g. "to remember" which implies prior knowledge, and "guess" which implies absence of knowledge). The specific terms tested by METVOC are 12 verbs: "know", "guess", "remember", "forget", "wonder",

“figure out”, “explain”, “understand”, “learn”, “teach”, “predict”, “deny”. In METVOC, the experimenter reads the children a story that consists of 12 episodes, each illustrated by pictures, e.g. “Dad comes into the room and says: <<Time for bed. If it's sunny tomorrow, we'll go to the park>>. In the morning John gets out of the bed and looks out the window. He sees the rain pouring down. <<Oh no Look at that! We won't be going to the park today>>”. At the end of each episode, children are required to select one of two cognitive verbs to describe the character’s state of mind, e.g. “does John know it’s raining or does John remember it’s raining?”. Each verb was used twice in the questions, once as the correct choice and once as the incorrect choice. METVOC is already adapted into French (Astington & Pelletier, 2004), Italian (Antonietti, Liverta-Sempio, Marchetti, & Astington, 2006) and Polish (Mieszkowska, Haman, Białecka-Pikul, & Otwinowska, 2016).

To recapitulate, the naturalistic observations of everyday family interactions were the primary method of investigating production of the ISL. However, as many uncontrolled factors vary from family to family, many researchers decide to employ more structured tasks. These commonly include shared book-reading to elicit multiple perspectives on the same event, as viewed by different story characters. Still, during social interaction children may simply mimic their interlocutor’s focus on the mind or emotions, which is why some researchers opt for non-interactive tasks, where children are asked to tell a story or describe a friend or a picture and are unaided by the experimenter. Others focus on the comprehension of internal state terms, assuming that full acquisition of the ISL involves the understanding of small discrepancies between the full meanings of internal state terms. This is done with the use of the forced-choice tasks in which children select one of two words to describe the character’s state of mind. Such tasks allow for a conscientious investigation of degrees of certainty (e.g. the difference between “know” and “guess”), or the variation in knowledge (e.g. “to remember” and “guess”).

Chapter 2: Internal state terms – a developmental perspective

The present chapter investigates the acquisition of internal state terms from a developmental perspective. A metaphor of a construction site is employed throughout the chapter, with the Internal State Lexicon seen as something that is being developed and constructed. Section 2.1. focuses on the building blocks – internal state terms – and discusses their role in the child’s social and cognitive development. Section 2.2. describes the foundations of the whole construction: the early awareness of thoughts about self and others. Section 2.3 describes what joins the building blocks, i.e. how social interaction helps children to construct the social world and, consequently, helps build up the ISL. The next section (2.4.) gives an overview of the process, showing when the particular types of internal state terms appear in the child’s language production and how is that related to the child’s environment (e.g. social interaction). Last section (2.5.) links the development of the ISL to theory of mind and language development in children.

2.1. The building blocks: role of internal state terms in child development

The acquisition of the ISL is linked to cognitive and social development. It is proposed that an involvement in internal state conversation causes children to acquire the internal state concepts and labels, and ultimately contributes to the construction of theory of mind (ToM), (Astington & Jenkins, 1995; Bialecka-Pikul, 2004; Bialecka-Pikul, 2012). Such evidence comes from longitudinal studies of parent-child interaction, which found that parental talk about internal states predicts children’s theory of mind development at later ages. For instance, the frequency of spontaneous family discourse about emotions (i.e. mother to child and child to mother) at children’s age of 3 years old was shown to predict the children’s ability to identify emotions at 6 years old (Dunn et al., 1991). Symons, Fossum, and Collins (2006) studied free play interaction between mothers and children aged 2 and found that the number of mothers’

references to desire states was significantly related to the children's theory of mind performance (on false-belief tasks) at 5 years old. They also found that this relation was independent of other confounding variables, including maternal sensitivity (understood as the ability to perceive, identify and respond accurately to signals from the child), socio-economic status, and child's language ability (as measured by Mean Length of Utterance, MLU). Thus, it seems that the exposure to internal states precedes and supports the development of theory of mind.

Importantly, the opposite direction does not seem to hold, i.e. it is not the case that children's earlier theory of mind development predicts the amount of maternal references to internal state terms. This has been shown by Ruffman, Slade and Crowe (2002) who investigated the two possibilities: whether maternal internal state talk enhances ToM development in children or whether mothers talk more about internal states because of their children's ahead-of-time social understanding. They collected data on the children's ToM development and the mothers' use of the ISL at three timepoints over one year (individual children differed in age at each time point, but the average age at T1 was 3 years old, at T2: 3,5 years old, and at T3: 4 years old). They found significant positive correlations between early maternal internal state talk and the children's later ToM after partialing out the children's early ToM. However, the authors did not find any significant correlations between the early ToM understanding in children and later maternal internal state talk (after partialing out earlier maternal internal state talk). Thus, Ruffman and colleagues suggest there is a unidirectional relation between the mother's internal state utterances at early timepoints and the child's later ToM (for similar results in 4-7 year old children see Adrián, Clemente, & Villanueva, 2007).

2.2. Laying the foundations of internal state lexicon: development of thoughts about self and others

First internal state terms are shown to appear relatively late in the language production of a typically developing child, i.e. by their second birthday (Fenson et al., 1994). It appears

that much has to be done before children begin to explicitly name and talk about the internal states of self and others. One of the important milestones is the development of joint attention with another person. Around the 9-10 months of age, infants move from primarily dyadic interactions (i.e. with a care-giver), to triadic interactions, i.e., between the child and another person about “an object of knowledge” – a referent in the outside world (e.g. a toy) (Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998; Chapman, 1991; Tomasello, 1995). This is the time when eye gaze becomes an important cue to the communicative intentions of others. Tracking another person’s eye gaze allows the child to identify which entities are relevant and should remain the focus of his/her attention. Concurrently, by the age of 9-10 months, infants pay more attention to objects of knowledge that are the focus of joint attention in comparison to those that are not (Brooks & Meltzoff, 2005). Establishing joint attention is necessary for linking the linguistic input (i.e. utterances, words that the child hears) with the objects and actions in the world. In fact, the amount of time infants spend in joint attention with their mothers is a strong predictor of the child’s comprehension of language and production of first words and gestures, as measured by the MacArthur-Bates Communicative Development Inventory, CDI (Brooks & Meltzoff, 2005; Carpenter et al., 1998).

Another important milestone preceding the development of the Internal State Lexicon is the growing awareness of others’ knowledge and beliefs. Studies show that at the age of 18, 15, or even 13 months, infants are already capable of tracking others’ knowledge and internal states and have non-verbal expectations of their behaviour, as evidenced by looking-time paradigms (e.g. Meltzoff, 1999; Onishi & Baillargeon, 2005; Surian, Caldi, & Sperber, 2007). In those experiments, infants watched an agent (an adult or an animal) reach for a toy in one of two locations. The agent’s choice of the location was based on either their knowledge (or lack thereof) or a belief (true or false) about the object’s whereabouts. Using a violation-of-expectations paradigm, Onishi and Baillargeon (2005) found that 15-month old infants

developed predictions about the agent's choice: the infants expected the agent to reach where he/she believed the toy to be (regardless whether the belief was true or false) and were surprised, i.e. looked longer, when the agent acted contrary to the belief he/she should have. Going a step further, Surian, Caldi, and Sperber (2007) found that 13-month old infants expected the agent to correctly infer the location of the object only when the agent had seen the object being placed, and looked longer when the agent correctly "guessed" the location without prior knowledge.

With the criticism of such implicit tests based on looking-times, there have been attempts to involve a child in a more active behavioural response, i.e. helping the experimenter (Buttelmann, Carpenter, & Tomasello, 2009; Buttelmann, Over, Carpenter, & Tomasello, 2014). In the latter experiment with the use unexpected-contents task, 18 month old infants were sitting with the experimenter and were shown a series of boxes. The boxes were covered in pictures of children's building blocks and contained blocks. However, the last box in a series (target box) turned out to contain a spoon (unexpected content). What differed between true-belief and false-belief conditions was whether the experimenter was present with the child when the contents of the last box, the target box (i.e. spoon) were discovered or not. In the true-belief condition, both the child and the experimenter saw that the last box contained a spoon, not blocks. In the false-belief condition, the experimenter was absent from the room exactly when the child discovered that the last box contained the unexpected spoon. Then, in both conditions, the experimenter would try to reach for the target box, unsuccessfully. The child was invited to help the experimenter by giving him what he wanted: the child could choose between a block (correct in false-belief condition) or a spoon (correct in true-belief condition). Thus, in order to help the experimenter appropriately, infants had to predict what the experimenter thought was in the box, based on his or her belief about its' contents. In either condition (true- or false-belief) 67% of the 18-month old infants correctly chose the item that the experimenter should expect to find in the target box.

These results seem to contradict the Piagetian theory of egocentrism, which views children before the age of 6 years as unable to share another's perspective: "The child being ignorant of his own ego takes his own point of view as absolute and fails to establish between himself and the external world of things that reciprocity which alone would ensure objectivity" (Piaget, 1923:197). However, it is important to note that the experiments reported here have purposefully reduced the difficulty of the task by tracking infants' behaviour in entirely non-verbal or mostly non-verbal situations. As such, these experiments point to an early and gradually developing awareness of other's knowledge and beliefs, what some call the "precursors of theory of mind" (Baron-Cohen, 1991; Tomasello, 1995). Nonetheless, in infancy, this awareness remains implicit in behaviour. The act of verbally communicating the children's expectations about others comes indeed later.

To sum up, before children begin to refer to internal states of self and others, two important milestones need to be reached. First, children need to engage in joint attention which enables them to link the linguistic input with the objects and actions in the world. It serves as the basis of early language learning, including the acquisition of the ISL. The second important milestone is the development of awareness of others' knowledge and beliefs, the precursor of theory of mind. It will direct the children's attention towards the mind and shape their early interactions with care takers. Hence, these two constitute the early foundation of the ISL use in the later linguistic interactions of children.

2.3. From the ground up: the role of interaction in the construction of social understanding

The early developing awareness of others knowledge and beliefs is facilitated through interaction and communication. In this interaction, children may experience the social and psychological world and construct an understanding of internal states (Carpendale & Lewis, 2004; Carpendale & Lewis, 2006). Slaughter, Peterson, and Carpenter (2009) investigated early

child-adult interactions in pre-verbal infants, when the communication on the part of the children is largely gestural. They observed 9-12 months infants for their production of communicative gestures in free-play with their mothers. The authors distinguished between the children's imperative gestures (e.g. reaching towards an object to request it) and declarative gestures (e.g. holding up an object to show it). They found that the child's tendency to use imperative gestures positively impacted the mother's frequency of labelling their children's internal states (e.g. "Want Mommy to get that?", "Did you figure that out?", "You like the ball, huh?"). The relationship remained significant even after controlling for the mother's general verbosity (i.e. total number of words produced). Thus, at the child's pre-verbal stage, the mother (in response to her child's communicative gestures) provides mapping between the internal experiences of the child and the specific terms for referring to those experiences. However, it is important to note that this mapping, i.e. the mother's tendency to explicitly name her infant's internal states, is not driven solely by the child's own interest in the mental world (as communicated by the gestures). Mothers first need to be sensitive to their children's communicative efforts. The mothers' attunement to their infants' internal states is also referred to as "mind-mindedness" (see e.g. Meins, 2013; Meins & Fernyhough, 1999).

Mind-mindedness is composed of two factors (which Elizabeth Meins calls indices): the caregivers' tendency to comment appropriately on their infants' assumed thoughts and feelings, and to misread their infants' internal states. An example of the first is commenting that an infant is surprised if he/she widens his/her eyes and stares in response to an event (appropriate mind-related comments). An example of the latter is stating that the infant is surprised in the absence of any overt clue to such an emotion (non-attuned mind-related comments) (Meins, 2013). Mind-mindedness influences the child's development of theory of mind and the child's use of internal state language. Meins, Fernyhough, Arnott, Leekam, and de Rosnay (2013) investigated 206 infant-mother pairs: they measured the mothers' use of internal state terms

during a 20-minute free play, the children’s use of internal state terms (as reported by mothers in a CDI questionnaire), and their theory of mind. To the latter aim they used Wellman and Liu’s ToM scale (Wellman & Liu, 2004) which measures children’s ability to recognize a belief different from their own, their understanding that knowledge depends on a previous access to crucial information, their recognition that another person will predict the contents of a container on the basis of its appearance, and the children’s ability to predict the protagonist’s behaviour on the basis of his/her false belief. As a result of the study, the authors constructed a model presenting the influence of mind-mindedness indices on the child’s ToM and ISL. The model is summarized in Figure 1. The mothers’ appropriate mind-related comments on their 8-months olds’ internal states were directly related to the infants’ theory of mind performance at the age of 51 months, but were not related to the children’s use of internal state terms. On the other hand, the non-attuned mind-related comments were negatively correlated with the Internal State Lexicon at 26 months and were unrelated to ToM. The authors suggest that mothers who misread their infants’ internal states in the first year of life might impede the children’s acquisition of the Internal State Lexicon. In conclusion, mother’s appropriate use of internal state terms supports ToM development (Dunn et al., 1991; Ruffman et al., 2002; Symons et al., 2006), and while mothers’ inappropriate use of internal state terms does not directly affect ToM development, it might impede the ISL acquisition in children.

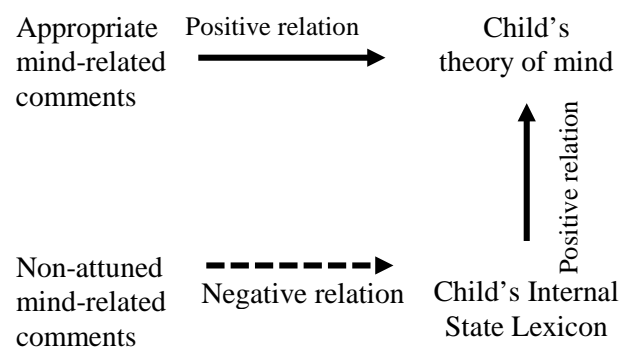


Figure 1. Model presenting the influence of mind-mindedness on the development of Theory of Mind and Internal State Lexicon in infants. Model based on Meins et al. (2013).

As the child grows, their experience in the social world becomes more diversified and richer through entering into new relationships, e.g. with peers and the surrounding community. Increasing opportunities to engage in social interactions allow children to further develop an understanding of the mind. Jeremy Carpendale and Charlie Lewis (2004; 2006) stress the role of such a relational, action-based perspective on the child's construction of an understanding of the mind. They suggest that concepts about the internal states or the mind are not merely passed on from adults, nor are they entirely built by the child. Rather, the child gradually constructs those concepts precisely through involvement in social activities. In practice, this view presupposes that the child learns to understand and use internal state terms through encountering them in appropriate contexts: "the circumstances surrounding the use of psychological words become the criteria for their use" (Carpendale & Lewis, 2004:88).

This interactional perspective on the development of social knowledge is shared by Katherine Nelson (Nelson, 1998, 2007; Nelson et al., 2003). According to her, children show an *experiential* nature and construct their own subjective view of the world (*personal mind*), which later on, through immersion in social interactions, becomes a *socially shared mind*. Nelson proposed the idea of the child "entering into the community of minds" (Nelson, 1998, 2007; Nelson et al., 2003) as a metaphor of theory of mind development. Her approach sees the child as entering into "new relationships of self and other understanding" within community: family, peers, and other people (Nelson et al., 2003:26). This entering is done through regular interactions, e.g. everyday problems that children face, and their actions. Thus, Nelson perceives children as not mere observers of social interactions, but actual actors, learning and making use of others' behaviour.

Bialecka-Pikul (2012) points to yet another aspect of the child's interaction that shapes theory of mind – the production of language. In her longitudinal study of over 100 Polish

children (tested at the age of 3.5 and 5.5 years old), Bialecka-Pikul investigated the impact of child-individual factors such as language production (children's description of pictures, eliciting verbs and prepositional phrases within grammatically complex sentences) and social factors, such as the frequency of child-induced activities (e.g. playing, doing puzzles, reading or listening to stories). She found that language production at the age of 3.5 was a strong predictor of ToM performance at the age of 5.5. Moreover, the impact of language production on ToM was significant even after controlling for the child's non-verbal intelligence and social factors. Thus, Bialecka-Pikul echoes Nelson's interactional perspective on ToM development and further suggests that interaction in fact enhances the quality of children's speech, which in turn shapes their theory of mind.

With reference to ISL development, the outcomes and perspectives discussed above imply that internal state terms and concepts are acquired through interaction and mentalistic conversation with others. This provides the child with opportunities to hear an internal state term in different contexts. These contexts may involve different semantic uses (meanings) of a given term (e.g. "I feel cold" vs. "I feel this is going to be fun"). They may also offer some syntactic clues that scaffold false belief representation (de Villiers, 2005; de Villiers, 2007). Specifically, internal state terms and communication terms are transitive verbs and take complements (e.g. "The man thought that it was a rock") and thus, complements may direct the child to guess that a novel verb is an internal or a communication verb. Levy and Nelson (1994) suggest that initially, children's use of internal state terms will be constrained to the previously heard uses. However, as children accumulate exposure to internal state talk, they build up inferential understanding of the terms. Thus, frequent exposure to internal state talk should provide children with more observational data which, in turn, should lead to an increased and a more flexible use of the terms.

2.4. The development of internal state language in children

The acquisition of internal state terms is based on interrelations between the internal state talk that the child hears (input), and the opportunities the child gets to use the internal state terms (output). Figure 2 gives an outline of these interrelations pictured throughout the child's development over the first 6 years. Let us now explain the nature of these interrelations between ISL input and output.

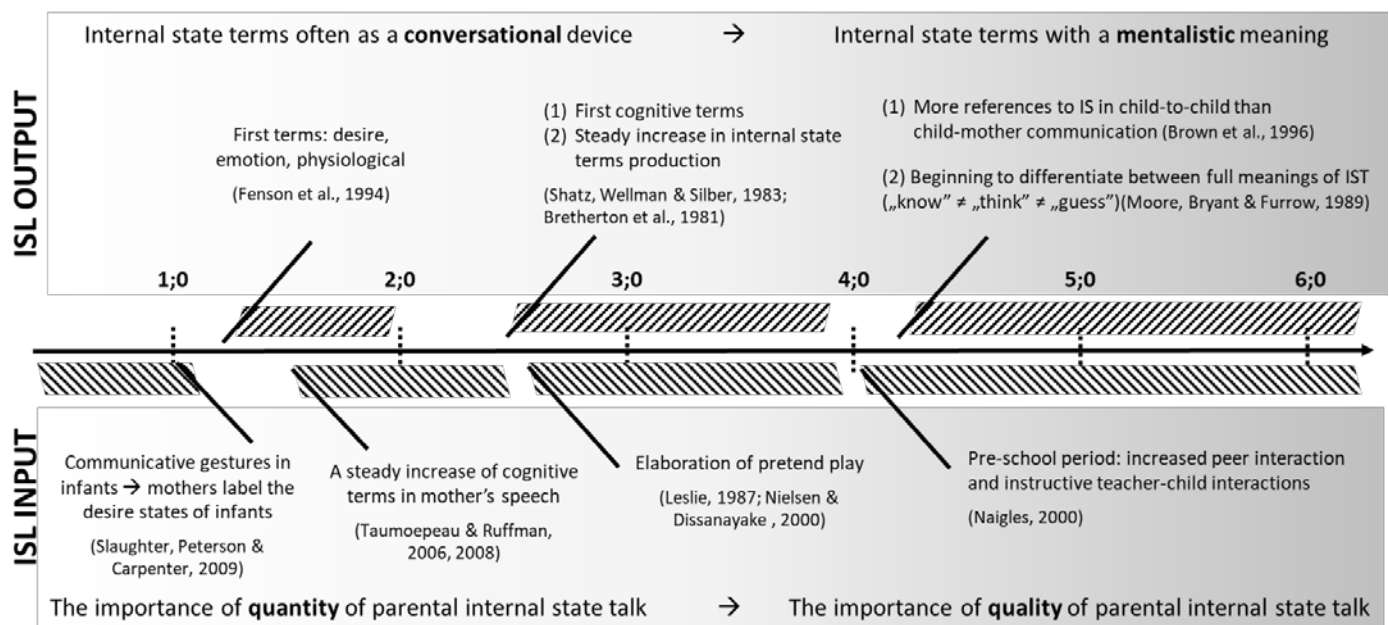


Figure 2. An outline of the associations between internal state language (ISL) input and output over the first 6 years of the child's development. The information included in the figure is based on studies reported in the present chapter, references are made to exemplary key studies.

In response to infants' communicative gestures, mothers provide mapping between the internal experiences of their children and the specific terms for referring to those experiences (Slaughter et al., 2009). At this point of development, the mere quantity of internal state talk from the mother or another predominant caretaker is most predictive of the child's subsequent use of internal states. This has been established in a series of longitudinal studies by Taumoepeau and collaborators (Taumoepeau & Reese, 2013; Taumoepeau & Ruffman, 2006, 2008). In the latter two studies (Taumoepeau & Ruffman, 2006, 2008) mothers described

pictures to their 15 month old and 24 month old infants. The authors coded the mothers' descriptions for internal state terms (desire, emotion, cognitive terms, and terms to indicate modulation of assertion, e.g. "maybe", "must", "might"), and asked parents to report the child's internal state talk at 24 and 33 months. The results demonstrated that: (1) the quantity of the mothers' use of desire terms at 15 months predicted the children's (reported) use of desire terms at 24 months; (2) the mothers' use of cognitive terms increased significantly between 15 and 33 months, while the use of desire and emotion terms remained relatively stable across the timepoints; (3) the mothers' use of cognitive terms at 24 months was the best predictor of the children's internal state talk (all subclasses) at 33 months. The results of these studies highlight two important matters: that desire talk precedes cognitive talk in both mothers and children, and that in the first years of the child's life, the quantity of the mother's use of internal state terms predicts the child's early internal state talk.

Nonetheless, internal state terms are acquired late compared to common first words spoken by children. A study of typically developing American infants showed that by the age of 16 months, when children generally have started to name items present in their everyday life, e.g. "bottle", "banana", "eye", "juice", "kitty", "shoe", the majority of infants understood their first internal state term, a desire term "wanna" (as reported by parents), but only 5% were reported to produce the word (in a total sample of 1789 children from 8 to 30 months; data gathered through parental Communicative Development Questionnaire; Fenson et al., 1994). For the majority of English-speaking children (50% or more of the Fenson's et al. total sample) the desire term "wanna" appeared in the production by their second birthday, together with the emotion term "happy" and the physiological term "hungry". Likewise, Brown & Dunn (1991) and Hughes & Dunn (1998) reported that around the age of two, children began to use internal state terms referring to volition and perception to relate internal states of self and others. At this time, they also started to label their basic emotions, both positive (e.g. "happy"), and negative

(e.g. “sad”, “mad”). This timeline is also confirmed by American English data that has been gathered in Wordbank¹ (Frank, Braginsky, Yurovsky, & Marchman, 2017). Thus, some of the basic internal state terms from the subclasses of desire, emotion and physiological terms might be expected to appear in typically developing children’s production in their second year of life (1;0-1;11). Importantly, production of desire, emotion and physiological terms precedes production of cognitive terms. This pattern seems stable across languages, as found by a cross-linguistic study of Italian, German, English and French infants (Kristen et al., 2014). A similar effect of desire terms preceding cognitive terms is found in Polish, as observed in the CDI data gathered from 73 Polish parents of monolingual and 73 parents of Polish-English bilingual children (children aged 2;0 – 3;4, mean age: 2;5) (Miękisz, 2016; Miękisz et al., 2017). The Polish desire term “chcieć” (want) was produced by 68% of the monolingual two-year-olds and 47% of the bilinguals. As a comparison, the cognitive term “wiedzieć” (know) was produced by 40% of the Polish two-year-olds, and only 15% of the Polish-English bilinguals.

Words referring to cognitive experiences are reported to appear in the child’s speech much after their second birthday (i.e. in the third year of life, 2;0-2;11) and even later than concrete verbs and some low-frequency words. In a study of a single child through ages 2;4 to 4;0, Shatz, Wellman, and Silber (1983) tracked the child’s use of four cognitive terms: “remember”, “think”, “know”, and “dream” in spontaneous speech. All of the terms appeared (with a mentalistic meaning) in the child’s speech around his third birthday (“think” at 2;8, “remember” at 2;9, “know” at 2;10, and “dream” at 3;0). The Wordbank data confirms this: both “have to” and “think” are produced by the majority of American English children by 30 months of age. This is partly a reflection of a parallel increase of cognitive terms in the mothers’ speech to their children (Taumoepeau & Ruffman, 2008). However, between the ages of 2 and

¹ Wordbank (<http://wordbank.stanford.edu/>) is an open database of children’s vocabulary development that contains data gathered through CDI questionnaires from over 64,000 children (over 72,000 CDI administrations) across 25 languages.

4, children often may use mental terms as an idiomatic or conversational device, i.e. without indicating the actual mental states. An instance of a conversational use of mental terms would be: “I know what this is for” (Nielsen & Dissanayake, 2000, p. 616), “What do you think?” (Nielsen & Dissanayake, 2000, p. 616) or “X, I mean, Y”, e.g. “I saw a snake, I mean, a lizard!” (Brown, Donelan-McCall, & Dunn, 1996, p. 840; Nielsen & Dissanayake, 2000, p. 616; Shatz et al., 1983a, p. 308). This is often done to maintain the conversation, rather than to indicate the comprehension of an internal state. In the study by Shatz and collaborators (1983), 38% of the uses of the term “know” by the child across the ages of 2;4 to 4;0 occurred in the phrase “I don’t know”, which the authors interpreted as an “idiomatic negative expression” (p.308) with no direct reference to an internal state.

A rapid spurt in the production of ISL takes place during the third year of life (2;0-2;11), and this is also the time when children start to produce longer utterances (often starting with telegraphic speech). At this time, children are shown to incorporate internal state terms to talk about the past and future experiences, in questions, as to confirm their evaluation (“Is X mad at me?”, Bretherton et al., 1981, p. 356), and in negations, to deny a state (e.g. “me no hungry”, Bretherton et al., 1981, p. 354). This is also the time when elaborated pretend play emerges, i.e. children start to treat inanimate objects as if they are real, and create imaginary objects or persons with no tangible referents in the immediate environment (Nielsen & Dissanayake, 2000). Some researchers propose that the mentalistic use of internal state language appears when children start to engage in pretend play (e.g. Brown, Donelan-McCall, & Dunn, 1996; Hughes & Dunn, 1997; Leslie, 1987). Pretence emerges before the third birthday, and becomes consolidated into the child’s play in the fourth year of life (3;0-3;11). By the fourth birthday, children may involve imaginary characters and animals in their pretend play (Nielsen & Dissanayake, 2000).

Leslie (1987) linked the pre-school child's ability to engage in pretend play with the development of the Internal State Lexicon and the emergence of theory of mind. Specifically, he proposed that it is the ability to appreciate dual representations (employed in pretense), that enables children to relate meta-representational expressions to the language of thought, and hence construct the Internal State Lexicon. According to Hughes and Dunn (1997), what is crucial for ToM and internal state talk development is the *shared* pretense which, being a social activity, involves peer-to-peer cooperation in order to create and sustain imagined situations. Indeed, Hughes and Dunn (1997) and Brown and colleagues (1996) observed that when involved in shared pretend play, children often referred to internal states to get their peer's attention ("Do you know what this is?", Brown et al., 1996, p. 840), referred to their thoughts or memories (e.g. "Do you think Captain Hook could be policeman?", Hughes & Dunn, 1997, p. 1030) or to direct the joint activity ("Pretend we're pirates", Hughes & Dunn, 1997, p. 1030). Hughes and Dunn (1997), who observed peer dyads aged 3 and 4 years old, found more such references to internal states in the context of shared pretend play, than outside of pretense (i.e. remaining speaker turns between the children). Thus, shared pretend play may serve as a social context that facilitates the use (and the acquisition) of internal state terms.

Brown and colleagues (1996) highlighted yet another important matter, namely, the increasing role of peer-to-peer interaction in the development of the child's ISL. The authors examined internal state talk in 4 year-old children's unstructured conversations with adults (mothers) and compared it with their unstructured interactions with other children (older siblings and friends of the same age). The results revealed significantly more references to internal states in the children's conversations with siblings and friends than with mothers (despite the mothers' frequent use of internal state terms to the children). Still, frequent use of internal state terms by both partners was related to more cooperative interaction in both child-friend and child-sibling dyads. Hence, it may well be that around the age of 4, parental internal

state talk still serves to build up the child's Internal State Lexicon, but the terms are now increasingly used by the child in their social interactions with other children. In fact, Jenkins and colleagues (2003) found that four-year-old children with an older sibling, produced and heard more cognitive talk and less desire talk than children without an older sibling. This is also in line with evidence showing that children with older siblings score better on theory of mind tasks than children without older brothers or sisters (Ruffman, Perner, Naito, Parkin, & Clements, 1998) or that siblings can compensate for slower language development in false belief understanding (Jenkins & Astington, 1996).

The acquisition of the full meanings of particular internal state terms comes gradually. As Nelson's (1998) "use before meaning" hypothesis suggests, children first develop a partial understanding of a word, but through continuous exposure to new instances of the word, they eventually acquire all its possible meanings and uses. Research shows that this may hold true for the Internal State Lexicon as well. Around the ages 4 to 6 children start to appreciate the small discrepancies between the full meanings of some internal state terms. For example, they start to understand that different cognitive terms indicate different degrees of certainty (Johnson & Wellman, 1980; Moore et al., 1989). This was tested by Moore, Bryant, and Furrow (1989) in a forced-choice task already described in the previous chapter (see Section 1.3.3.). Children were asked to determine the location of a hidden object based on two puppets giving conflicting statements with the use of internal state terms "know", "think", and "guess". The results showed that by the age of 4, some of the children could differentiate "know" from "think", and "know" from "guess". They were also aware of the fact that "know" indicates a higher reliability than "think" and "guess". This understanding was evident in all children by the age of 5. However, the distinction between "think" and "guess" was not well understood even by the age of 8 years old (the oldest group studied in the experiment).

Naigles (2000) suggests that the full mastery of the internal state concepts (i.e. the differentiation between the concepts) may be enhanced through the pre-school environment, i.e. via interaction with peers and via purposefully instructive teacher-child interactions. However, the other suggestion is that the full mastery comes as a consequence of a shift from the role of quantity of adult internal state talk, to the importance of quality of internal state talk. Howard, Mayeux, and Naigles (2008) investigated the matter by comparing the effect of schooling and the effect of varied quality of maternal internal state talk. They audio-taped 30-minute interactions between the middle- and high-SES mothers and their children (attending pre-school on regular basis or not attending) in a naturalistic context (e.g. during cooking, mealtime, arts and crafts, play) and during one structured event, i.e. playing a memory game. The mother-child interactions were coded for the type of utterance containing a cognitive term, i.e. questions (both wh-questions, e.g. “What do you think?”, and Yes/no questions, e.g. “Do you know what this is?”), statements (e.g. “I know that song”), and directives (e.g. “Think about it”). The results showed that, with regards to the mothers’ input, questions improved the children’s ability to distinguish between the verbs, while statements were inhibitory. Moreover, questions were more often used by the mothers of non-pre-schoolers (compared to the mothers of pre-schoolers). With regards to the pre-school attendance (or lack thereof), the authors found that it did not significantly predict the child’s ability to differentiate between the mental verbs. The authors suggest that within the middle- and high-SES families, mothers who do not send their children to pre-school may deliberately attempt the type of conversational interactions that they assume pre-schoolers possibly receive from their teachers.

In a study by Slaughter, Peterson, and Mackintosh (2007), a distinction was made between the quantity and quality of mental state talk in mothers’ narratives of a picture-book to their pre-school children. This way, the authors wanted to investigate whether simple parental mentions of internal state terms are enough, or whether parents need to elaborate on the internal

states in order to improve their children's theory of mind understanding. They distinguished between simple statements, i.e. references to cognitive terms (e.g. "He remembers", "She doesn't realize") and clarifying statements, i.e. phrases or sentences that (a) explicitly state the contents of characters' minds (e.g. "He remembers that he has not done the bedroom yet"), (b) include explanations for sources of knowledge (e.g. "She didn't see them playing so she will not know who has messed up her dressing table"), or (c) note the discrepancies between different characters' mental states or between mental states and reality (e.g. "He puts all the make-up back so Mummy doesn't know what they've been up to"). The children were then given a false-belief task, i.e. change of location task (based on Baron-Cohen's Sally-Anne task, Baron-Cohen, Leslie, & Frith, 1985). The results of the study demonstrated that it is the mothers' clarifications of others' thoughts rather than the mere use of cognitive terms that are significantly associated with the children's false-belief understanding.

To recapitulate, internal state talk emerges in children's speech around the second year of life, with desire, emotion and physiological terms appearing first. Then, around the third birthday, children start to refer to mental/cognitive states. In many cases, the early use of mental/cognitive terms is of conversational or idiomatic nature. During the fourth year of life there is an increase in the frequency of internal state terms in children's speech. It is also at this time that children start to differentiate between the full meanings of internal state concepts (e.g. "know", "think", "guess"). The acquisition of the Internal State Lexicon is also a reflection of the developing social interaction: during infancy and the first two years of life, it is the mother or other mostly present caretaker who is the role model for the ISL acquisition: she will first label the internal experiences of her child and provide sentential context for figuring out the basic meaning of internal state terms. At the beginning, the quantity of maternal internal state talk matters the most for the child's subsequent ToM development and internal state language. However, as the child is progressing in their ISL production, around 4 years of age, the

importance shifts to the quality of internal state talk: explanations of the feelings of story characters in a book, describing casual connections, elaborating on different possible perspectives, turn-taking in conversations. It is also at that age that children usually start their formal education, which brings about, among others, an increase in peer-to-peer interaction. As a consequence, communication with other children (e.g. in shared pretend play) becomes a new prevalent context for the ISL use.

2.5. The Internal State Lexicon in relation to theory of mind performance and language

Naturally, researchers studying the ISL see the acquisition of the internal state terms as an indicator of a developing theory of mind. Thus, many have searched for a link between the performance on classic ToM tests, i.e. false-belief tasks, and the use of internal state terms. The previous sections have mentioned the link between the maternal use of the ISL and the child's later ToM performance (e.g. Ruffman et al., 2002). However, of more interest to the present thesis is the possible link between the child's own use of internal state terms and their performance on theory of mind tests. Nielsen and Dissanayake (2000) investigated this association in pre-school children. The children's use of 16 mental terms was coded during a free play with a parent: "know", "think", "forget", "mean", "remember", "guess", "pretend", "dream", "bet"/"reckon", "hope", "trick", "wonder", "wish", "figure", "believe", and "understand". Additionally, the children were administered three standard false-belief tasks. In the unexpected contents task following the procedure from Bartsch and Wellman (1989) the children were presented with a Band-Aid box that in fact did not contain band-aids; these were found in an unmarked box. Next, children were introduced to a Tigger puppet who fell over and was looking for band-aids. The children were asked where would Tigger look for band-aids. In order to pass the task, the children had to answer that Tigger will look for band-aids in the Band-Aid box (though they already knew the band-aids were not really there). The

second task was a standard unexpected transfer task (Wimmer & Perner, 1983), in which Winnie the Pooh's snack, originally placed in one location, was transferred (without him seeing) to another location. The children were asked where will Winnie the Pooh look for his snack, when he comes back. The correct answer would be to attribute Winnie the Pooh with a false belief that the snack is where he originally placed it. The last false-belief task was Lewis and Osborne's (1990) Smarties task in which children were shown a smarties tube. On opening, it turned out that the tube contained pencils. The children were then asked what they originally thought was in the tube (attributing a false belief to self) and what would Tigger think was in the tube (attributing a false belief to other). The results revealed a significant positive correlation between the performance on all of the false-belief tasks and the use of internal state terms with a mentalistic meaning ($r = .30$). Interestingly, significant but weak correlations were found also between the scores on the false-belief tasks and the use of internal states in idiomatic expressions, e.g. "I don't know" ($r = .17$) and in conversational use, e.g. "What do you think?" ($r = .14$).

Apart from ToM, vocabulary size could also relate to the use and comprehension of internal state terms. A study by Antonietti and collaborators (2006) points to a direct link between children's receptive vocabulary knowledge and their understanding of internal state terms. They tested Italian children aged from 4 to 8 years with a receptive vocabulary test and the Metacognitive Vocabulary Test (METVOC, Astington & Pelletier, 2004, see Section 1.3.3.) developed to measure children's acquisition of the full meanings of internal state terms (e.g. the difference between "know" and "guess"). The results revealed a significant positive correlation between the score on the METVOC and the score on receptive vocabulary test, $r = .62$.

Another aspect of language that may have consequences for the acquisition of ToM and ISL is syntax. Astington and Baird (2005), de Villiers (2007), Milligan, Astington, and Dack

(2007) suggest that knowledge of syntax may be involved in ToM development in two ways. First, syntactic clues may provide scaffolding for false belief representation (de Villiers, 2005; de Villiers, 2007), e.g. the fact that internal state terms are transitive and take complements (e.g. “He thought that X”, “She said that Y”) may direct the child to infer that if a novel verb is appears in this specific syntactic context, it may be an internal or a communication verb. Moreover, while the main clause may be true (e.g. “He thought that...”, “She said that...”), the embedded clause may be false (e.g. “...this was a rock”). Thus, complement structure not only implies an internal state but also provides the opportunity to observe contradictions between internal states and reality (de Villiers, 2000; Hale & Tager-Flusberg, 2003). Accordingly, passing a false-belief task requires understanding of complex grammatical structures. In their longitudinal study of three-year-olds, Astington and Jenkins (1999) found that syntactic knowledge, summed across the receptive and expressive subtests, predicted later performance on false-belief tasks but the reverse pattern did not hold. Lohmann and Tomasello (2003) confirmed this result with a training study of three year olds. The children were shown deceptive objects (e.g. an object that looked like a flower, but was really a pen). In the full-training condition, the experimenter talked about each of the objects using internal state terms (e.g. “think”, “know”, “say”) inserted in complement structures (e.g. “What do you think this is?”, “You thought it was a flower”). In the discourse-training condition, the deceptive nature of the objects was highlighted, but the experimenter did not use either internal state terms or complement structures. Instead, the experimenter stated what the object was (e.g. “It was a flower, now it is a pen”). There was also a no-language-training condition, in which the experimenter highlighted the deceptive nature of the object non-verbally. For example, the experimenter showed the child the object saying “Look!”, then revealed the true nature of the object and said “Oh!”. The children in each training-group took part in three 20-30 minute training sessions (over a two week period), and finally took a false-belief post-test to measure

their ToM improvement. The results showed that both the full-training and discourse-training groups equally improved their performance on false-belief tasks, while the no-language-training condition did not show any change. Thus, the study revealed that exposing the deceptive objects using either nouns or internal state terms embedded in a sentential context is a strong facilitator of false-belief understanding.

Some studies found no significant relation between the number of internal state terms used and theory of mind performance or vocabulary knowledge. For example, Meins and colleagues (2006) have investigated the use of internal state language in children aged from 7 to 9 in a describe-friend-task and when telling a story after a picture book (both non-interactive tasks). They also tested the children's theory of mind with the use of Happé's (1994) strange stories task that investigate understanding of complex states like misunderstanding and double bluffing. Finally, they examined the children's receptive vocabulary knowledge with a standardized picture recognition task. The results demonstrated significant cross-task stability in the children's use of the internal state terms while narrating a picture story or describing a friend. However, the use of internal state terms in either task was not significantly correlated neither with the performance on the ToM task or the receptive vocabulary test. Charman and Shmueli-Goetz (1998) also failed to establish a significant relationship between 7-year olds' theory of mind performance and their use of internal state terms when narrating a picture book. Meins and colleagues conclude that children's tendency to attend to and refer internal states when explaining people's behaviour generalizes across contexts (thus a high cross-task stability). However, children's individual differences in the tendency to use internal state terms might be independent of their cognitive development to reason about internal states: "having a ToM is somewhat different to spontaneously using one's ToM abilities to describe, explain and interpret the behaviors of others" (Meins et al., 2006:183). Thus, Meins suggests that children's use of internal state terms in a non-interactive

context might reflect their spontaneous tendency to use their theory of mind capacity, rather than the ToM capacity itself.

Thus, there is reason to believe that the acquisition of ISL is linked to the developing theory of mind and language, including vocabulary and grammar. Though research evidence yields mixed results, the potential link should be of great interest to those researchers who study bilingual language development. The upcoming chapter discusses bilingual language acquisition in relation to those two areas influential for ISL: language and theory of mind.

Chapter 3: Bilingual children

This chapter will present the rationale for studying the Internal State Lexicon in bilingual children. To this end, the chapter will discuss the differences between the bilingual and monolingual language acquisition and explain how these differences might influence the bilingual use of internal state terms. Section 3.1. will start by briefly defining a bilingual speaker. Next, it will discuss the bilingual language acquisition in three domains that seem to be most relevant to the study of the ISL: vocabulary knowledge, theory of mind development, and narrative skills. The comparisons will be made between the bilingual and monolingual developmental trajectories, and between the bilinguals' performance in two languages. Section 3.2. will focus on the available evidence on the acquisition of Internal State Lexicon in bilingual children: the comparisons between the amount of internal state terms produced by bilingual and monolingual children, and in the two languages of a bilingual. This will serve as a background for Section 4.1. of the next chapter, which will present the pending questions concerning the acquisition of ISL in bilingual children.

3.1. The bilingual speaker: theory of mind and the bilingual language acquisition

The most commonly employed definition of bilingualism is the use of two languages on regular, everyday basis (Grosjean, 1984, 2010). While such definition may seem broad, it is deliberately meant to be inclusive. Bilingualism is a heterogeneous construct that includes different types of bilingual speakers. More specifically, it accommodates those who have acquired another language in a context of formal education (e.g. bilingual school), and those for whom bilingualism was not a goal, but a side-effect of life circumstances (e.g. bilingualism due to immigration, language-mixed marriages / bicultural families, or in bilingual societies). Bilinguals may exhibit comparable proficiency in their two languages (balanced bilinguals), or be more fluent in one of the languages (unbalanced bilinguals). Some bilinguals have acquired

the second language (L2) in their childhood (early bilinguals), or in teenage years or later (late bilinguals). Those who have acquired two languages at the same time are referred to as simultaneous bilinguals, and those who have started acquiring the L2 after having considerably acquired their first language (L1²) are often referred to as sequential bilinguals. Finally, those who have learned the L2 and reinforced their L1 are called additive bilinguals, while those who have learned the L2 at the expense of the L1 (L1 attrition) are often referred to as subtractive bilinguals. A useful review of the topic is given in De Houwer (2009); or in Polish in Kurcz (2007) and in Wodniecka, Mieszkowska, Durlik, & Haman (2018). The present thesis focuses on bilingual children acquiring the L2 early, in the immigration context (children of at least one Polish parent, living in the UK), who may be considered simultaneous and additive bilinguals, i.e. are acquiring two languages at the same time, and are keeping both.

What is common to all the types of bilingual speakers is that “bilingual is not the sum of two complete or incomplete monolinguals; rather, he or she has a unique and specific linguistic configuration” (Grosjean, 1984, 1989, 2010). This has important consequences on the way researchers and practitioners should view bilingual language acquisition. First, a bilingual person does not have to attain an equal mastery of the two languages (Grosjean, 1984, 2010). In fact, bilinguals’ proficiency in a given language is largely dependent on the amount and quality of input they receive in that language, as well as on the opportunity to use the language (Haman et al., 2017; Hoff, 2017; Hoff & Core, 2013; Hoff, Welsh, Place, & Ribot, 2014; Oller, Zurer Pearson, & Cobo-Lewis, 2007; Pearson, Fernandez, Lewedeg, & Oller, 1997; E. Thordardottir, 2011). In the context of simultaneous bilingual acquisition, the bilinguals’ exposure is naturally divided between the two languages, e.g., the mother’s language and the father’s language, or the L1 (i.e., language acquired first, often also a home, heritage, or

² Henceforth L1 will indicate the first language (in the order of acquisition), and will often be juxtaposed with L2 which will indicate the second acquired language.

minority language) and the L2 (i.e., language acquired second, often also a community, or majority language³). As a result, bilingual children receive less input in each of the languages in comparison to the amount of input received by monolingual peers (De Houwer, 2014; Montrul, 2008; Pearson, Fernández, & Oller, 1993). Though the variations in the amount of language input are one of the factors that lead to subsequent differences between bilingual and monolingual children, many of the early developmental trajectories are similar across the two groups.

3.1.1. Theory of Mind development in bilingual children

Let us first point out that there is still an ongoing debate whether the developmental trajectory of theory of mind (ToM) is the same across different cultures. On one hand, Callaghan and collaborators (2005) that investigated false-belief understanding in five countries: Canada, India, Peru, Samoa, and Thailand, and found that children from all the studied cultures passed the false-belief task by approximately 5 years of age. Liu, Wellman, Tardif, and Sabbagh (2008) found that the general developmental trajectories were similar in Chinese and North American typically developing children. Specifically, all studied groups (Chinese mainland, Hong Kong, American and Canadian children) developed from below-chance to above-chance performance on false-belief tasks between 3- and 5- years old. However, there were also some locale-specific differences between the Canadian and Hong Kong children: the Canadian children reached the above-chance performance on false-belief tasks at the age of 40 months, two years earlier than the Hong Kong sample. Still, there were no such differences between the mainland Chinese and the American children who were found to develop synchronously. In general, results from the false-belief tasks performed by children with various native languages show that many 4- and 5- year olds perform successfully while

³ Henceforth, the majority language will indicate the language spoken by the majority population in a given country (e.g. English in the UK), and will often be juxtaposed with home language which will indicate a different language spoken at home (e.g. Polish spoken at home of a Polish family living in the UK).

many 3-year olds do not (Polish: Białecka-Pikul, 2012; English in Canada, Spanish in Peru, Samoan, Thai: Callaghan et al., 2005; Mandarin Chinese: Goetz, 2003; Romanian: Kovács, 2009). Even so, studies comparing ToM performance between notably distant cultures (e.g. American vs. Chinese, American vs. Japanese), have indeed pointed to some cross-cultural differences. Specifically, Wellman, Fang, Liu, Zhu, and Liu, (2006) found that Chinese children understand knowledge-ignorance before diverse beliefs: they understand another person's ignorance about the contents of a container when they know what is in the container before being able to understand that two persons (they and someone else) may have different beliefs about the same object, when they do not know which belief is true or false. However, the pattern is opposite for the American children: they understand diverse beliefs before understanding knowledge ignorance. Gut (2016) and Afek and Gut (2018) look for explanations of this in the history, traditional patterns of parenting and philosophy of the Chinese. They propose that these have shaped the current parenting styles and indirectly determine developmental differences found in the ToM performance of children from Chinese and Western cultures.

Regardless of whether language-specific differences affect ToM performance, the bilinguals' performance on false-belief tasks may be aided through their bilingual upbringing. The specific areas that are improved by bilingualism include metalinguistic skills, selective attention and sociolinguistic competence, and all are related to ToM performance. For example, bilingual children are consistently exposed to two labels (from two languages) denoting the same concept (e.g. PL: "pies", ENG: "dog"). Thus, their metalinguistic skills may be improved by the early recognition that there is more than one way of naming things (Au & Glusman, 1990). Indeed, compared to monolinguals, bilinguals exhibit a greater linguistic flexibility, i.e. they are more likely to agree that if the sun were called the "moon" and the moon were called the "sun", then the "sun" would be up at night and it would be dark (Bialystok, 1987; Friesen & Bialystok, 2012). This metalinguistic awareness is correspondent to the issue of

metarepresentation found in the false-belief tasks: there is more than one way of viewing an object or an event, and it can be viewed differently either by the same person and by different people (Gopnik & Astington, 1988). For instance, in the appearance-reality task (Flavell, Green, Flavell, Watson, & Campione, 1986), children are shown a sponge that looks like a rock. Passing the task requires children to understand that an object can appear to be one thing, when it really is something different (Flavell et al., 1986). Importantly, performance on the false-belief tasks is significantly correlated with the metalinguistic tasks, such as understanding homonyms (i.e. two word-forms that sound the same but have different meaning) (Doherty, 2000), understanding synonyms (i.e. two different word-forms with the same meaning) (Doherty & Perner, 1998), and tasks that measure understanding of synonymy, ambiguity, grammatical function and phonological segmentation (Longobardi, Spataro, & Renna, 2014). As follows, it is possible that bilingual children outperform monolingual peers on the false-belief tasks due to their enhanced metalinguistic skills.

Another reason why bilingual children might gain an advantage in false-belief tasks lies in the fact that their two languages are constantly active in their minds, which is visible e.g. when they transfer or borrow words from one language to another (for a review, see Kroll, Bogulski, & McClain, 2012). This dual language activation forces bilinguals to constantly exert control over the two languages, e.g. when speaking or even reading (and thus activating) one language, they need to inhibit the other language to keep it in a dormant state, i.e. not affecting their language production or comprehension (Bialystok, Craik, Green, & Gollan, 2009; Green, 1998). This constant control over the activation/inhibition processes trains the bilingual minds. Indeed, bilingual children are shown to have a better developed inhibitory control (Bialystok, Craik, Klein, & Viswanathan, 2004; Ellen Bialystok & Martin, 2004; Carlson & Meltzoff, 2008; Martin-Rhee & Bialystok, 2008). The bilingual advantage is especially visible in those tasks

that require interference suppression⁴, i.e. ignoring (or suppressing) a faulty response evoked by a misleading or conflicting information (Bialystok, Craik, Green, & Gollan, 2009; Esposito, Baker-Ward, & Mueller, 2013). This type of inhibitory control is strongly related to the performance on false-belief tasks in preschool children (Carlson & Moses, 2001; Leslie, German, & Polizzi, 2005). Specifically, passing a false-belief task requires the child to suppress (or override) a dominant tendency to refer to reality (Carlson & Moses, 2001; Carlson, Moses, & Hix, 1998).

Importantly, bilinguals outperform monolinguals not only on tasks directly focused on inhibitory control, but also on tasks more resembling theory of mind, for example those that require overriding an egocentric bias, i.e. the tendency to attribute own knowledge to others (Leslie et al., 2005). Fan, Liberman, Keysar, and Kinzler (2015) tested monolingual and bilingual 4- to 6-year old children on a social communication task that required perspective taking to interpret a speaker's intended meaning. The children were asked by the experimenter to move objects around a 4 × 4 grid shelf. Four grid squares were obstructed so that the child saw their contents, but the experimenter, who stood on the other side of the grid, did not. In the critical condition, the experimenter's instruction to "move the small car" could refer to a mutually visible target object (a medium-sized car), or to a distractor object, a smaller car that was visible only from the child's egocentric perspective. To succeed, the children had to take the experimenter's perspective and choose the mutually visible target. The authors found that the bilingual group regularly took the experimenter's perspective (in 77% of all the cases), while the monolingual children were at chance in selecting between the target and the distractor,

⁴ "Interference suppression" is here juxtaposed with „response inhibition“, i.e. the ability to refrain a habitual or prepotent response. While the former is seen in conflict tasks (e.g. Simon task) with two conflicting cues, each suggesting a different response, one of which has to be ignored, the latter is employed in delay tasks (e.g. Day/Night task) with univalent (single) cue, when the conflict is present only at the response level. Bilinguals show an advantage, relative to monolinguals, on conflict, but not delay tasks (for more details see Esposito, Baker-Ward, and Mueller (2013) and Martin-Rhee and Bialystok (2008)).

i.e. they selected the correct object in 50% of all the cases. Thus, bilinguals were shown to be more attuned to the perspective of the speaker, relative to their monolingual peers.

Last but not least, bilingual children need to adjust the language they speak according to their various interlocutors (e.g. English with the English monolingual teacher, Polish with the Polish grandparents). Thus, through recognition that other people cannot understand one of their languages, bilinguals come to admit the linguistic competence of others, and thus develop their own sociolinguistic competence (Goetz, 2003). There is evidence showing that bilingual children are able to appropriately match their language to others by their third birthday (Lanza, 1992). This ability to make informed language choices may be linked to the growing understanding that other people may have different knowledge than their own, and it is possible that bilingual children develop this ability during their third year.

In accordance with the reasoning mentioned above (bilingual advantage in metalinguistic awareness, inhibitory control, sociolinguistic competence), research shows that bilingual children perform equally well (Goetz, 2003) or significantly better than monolingual peers on false-belief tasks (Farhadian et al., 2010; Goetz, 2003; Kovács, 2009). For example, Goetz (2003) administered false-belief tasks to Chinese monolingual, English monolingual, and Chinese-English bilingual children (speaking Chinese at home and English at school) in groups of 3- and 4-year olds. The battery included a series of four tasks: (1) an appearance-reality task that measured the children's understanding that an object might look like one thing (rock) but be another (sponge); (2) a perspective-taking task that checked whether the children understood that a pictured object might look like one thing or another, depending on its orientation (a turtle standing on its feet or laying on its back); (3) an unexpected contents task which measured the children's understanding that they may know the contents of a box (a box of M&M's contained a toy car) but somebody who has not seen inside the box will hold a false belief about its contents; (4) an unexpected transfer task in which the children listen to a story about a boy who

put a candy in a blue drawer, but while he is away, his mother transfers the candy into a red drawer and they have to decide where will the boy look for the candy where he comes back home. The results revealed that both 3- and 4-year old bilinguals outperformed the monolingual peers on the appearance-reality, perspective-taking, and unexpected contents tasks. However, there was no bilingual advantage observed for the unexpected transfer task. Goetz suggests that the discrepancy in the results may be caused by the fact that the unexpected transfer task was the only task where children did not have to respond verbally (they pointed to the drawer where the boy believed the candy was). Specifically, she presupposes that the increased inhibitory control of bilinguals is particularly linked to the linguistic performance and thus bilinguals show an advantage only in verbal tasks.

Kovács (2009) tested 3-year old Romanian-Hungarian bilinguals from dual language families (i.e. mother and father speaking different languages to the child on daily basis) and Romanian monolinguals. The children were administered two theory of mind tasks: a standard unexpected transfer task (such as the one described in Goetz's 2003 study above), and a modified ToM task. The modified ToM task was meant to mimic the dual-language situation which bilingual children could often experience: the children had to recognize another person's false-belief by considering that this person does not speak their two languages. The children were presented a story in which two characters, a bilingual and a monolingual puppet want to buy ice-cream. There were two stands, one selling ice-cream and the other one selling sandwiches. The ice-cream seller called to the puppets in a language that the monolingual puppet did not speak (Hungarian) that he run out of ice-cream but that the sandwich-seller had some. The utterance was translated and it was pointed out to the participants that the monolingual puppet did not understand what the seller said. Then the children were asked: "Where will the monolingual puppet go to buy ice-cream?" It turned out that twice as many bilinguals performed successfully on both the standard unexpected transfer task and the

modified ToM task (59% of bilinguals and 25% of monolinguals answered correctly the standard ToM task, and 47% of bilinguals and 19% of monolinguals answered correctly the modified ToM task). Kovács attributes this advantage to the extensive practice those early simultaneous bilinguals have had in selecting and monitoring their two languages, which improved their inhibitory abilities and gave them an advantage in the false-belief tasks that require inhibitory control.

Moreover, the bilingual advantage on the ToM tasks is found also in tasks other than false-belief. Gordon (2016) investigated theory of mind in 4-year old English monolingual and bilingual English-Spanish children with balanced receptive vocabulary skills in both languages. Importantly, the children were tested not only with standard false-belief tasks (unexpected contents, unexpected transfer), but also with tasks that assessed their understanding of diverse beliefs, i.e. that they could hold one belief (e.g. about a location of a cat), while another person could have a different belief, or understanding of emotions, e.g. that a person could have a happy expression on their face while really feeling sad (real-apparent emotions task). The results revealed significant group differences only on the explicit false belief task (unexpected transfer) and the diverse desires task. Interestingly, bilinguals outperformed monolinguals on the diverse desires task (understanding that a person may have a belief other than their own), but scored significantly below the monolingual group on the unexpected transfer task.

Thus, though research evidence indicates a general advantage for the bilingual children relative to their monolingual peers, the results are not universal. Moreover, there is no evidence on whether the possible bilingual advantage on the performance on the false-belief tasks actually translates to their use and understanding of internal state terms.

3.1.2. Vocabulary skills in bilingual children

There is substantial evidence showing that bilingual children who acquire two languages simultaneously follow the same developmental trajectories as their monolingual peers. For

example, Pearson and collaborators (1993) found that 25 bilingual and 35 monolingual children produced similar numbers of words at six timepoints from the age of 1;4 to 2;3, when the bilinguals' total conceptual vocabulary (summed in both languages) was considered. These results were confirmed by Hoff and colleagues (2012) on a sample twice as big (almost 50 bilinguals and over 50 monolingual children aged 1;10 to 2;6). They also found bilingual and monolingual children start to combine words at the same time: almost all children in the sample (no differences between the bilingual and monolingual groups) combined words into utterances by their second birthday (Hoff et al., 2012). Notably, these group similarities are observed when both languages of a bilingual are taken into account, e.g. when the total vocabulary (L1 + L2) is considered.

However, when the bilinguals' vocabulary size is examined in each language separately, bilinguals are consistently reported to have smaller vocabulary than their monolingual peers. Bialystok, Luk, Peets, and Yang (2010) tested receptive vocabulary knowledge in a sample of over 1 700 children (over half bilingual) between 3 and 10 years old. The bilingual children had various L1s but all acquired English as their L2: it was the language of their schooling and the language they were tested in. The results showed that bilinguals of all age groups scored significantly below their monolingual peers on the receptive vocabulary test in the majority language, English. Similar findings have been reported by other researchers, e.g. Leseman (2000) and Oller and collaborators (2007). The vocabulary setbacks are found in bilinguals speaking different language pairs (English and another language: Bialystok et al., 2010; Turkish-Swedish, German-Swedish: Bohnacker, Lindgren, & Öztekin, 2016; Polish-Norwegian, Polish-English: Hansen et al., 2017; Russian-German: Klassert, Gagarina, & Kauschke, 2014). Similar results, showing the bilinguals' lower performance compared to

monolinguals, were also found in the bilinguals' L1 (home language) in Polish-English bilinguals (Haman et al., 2017; Mieszkowska et al., 2017)⁵.

Oftentimes, proficiency in the two languages is asymmetrical due to the predominance of one language in the child's environment. This predominance is manifested in a larger amount of exposure to that language, and/or the better quality of input in that language (Hoff, Welsh, et al., 2014; E. Thordardottir, 2011). Such unbalanced bilingualism is relatively common (Hoff, 2017; Hoff et al., 2012). The two languages may serve different purposes or be used in different domains of life, e.g. L1 used at home, L2 used at school. As a result, bilinguals may develop domain-specific vocabulary in one language, but not the other. For example, Bialystok and colleagues (2010) found that bilingual children who communicated in a given language at home and used English (their L2, majority language) at school, showed similar knowledge of school words (e.g. astronaut, rectangle) in English as their English monolingual peers, but did not know as many home words in English (e.g. squash, horrified), as the English monolinguals. In discussing their results, the authors point out that English (the majority language) is not used as extensively in bilingual homes as in the monolingual ones. Consequently, bilingual children receive less exposure to the home words in their L2, as compared to monolinguals, which accounts for lower vocabulary scores in this specific vocabulary domain.

Though the impact of dual language input on general vocabulary knowledge is well evidenced, it remains unknown whether such input, divided between the two languages, influences that peculiar area of the Internal State Lexicon. It is possible that, since general vocabulary size has been found by some to be positively correlated with the comprehension of internal state verbs (Antonietti et al., 2006), the relatively smaller vocabulary of bilingual children might also include fewer internal state terms (as compared to monolinguals). However,

⁵ The two studies were carried out on partially overlapping samples of children tested within the Bi-SLI-PL project.

the links between general vocabulary knowledge and the ISL are not irrefutable (see e.g. Charman & Shmueli-Goetz, 1998; Meins et al., 2006 and subsection 2.5.). Moreover, the relationship has not been tested before in bilingual children.

To recapitulate, children who are bilingual from early age follow the same developmental trajectories as their monolingual peers, especially when their total conceptual vocabulary is considered (e.g. Pearson et al., 1993; Hoff et al., 2012). However, when the bilinguals' vocabulary size is examined in each language separately, bilinguals are consistently reported to have smaller vocabulary than their monolingual peers (e.g. Haman et al., 2017; Oller et al., 2007). Moreover, it is common that bilinguals develop domain-specific vocabulary in one language, but not the other due to exposure differences, e.g. they receive less exposure to the specific words in their L2 (Bialystok et al., 2010). It is still to be seen whether such characteristics of bilingual upbringing may influence the acquisition of the Internal State Lexicon.

3.1.3. Narrative skills in bilingual children

The development of narrative competence is the end result of increasing language skills and the child's pragmatic awareness (Reese, Suggate, Long, & Schaughency, 2010). Research on bilingual narratives is still scarce and restricted to several language pairs: English-Spanish/ Hebrew/ Slovak/ Swedish, Swedish-Finnish, Russian-German/ Hebrew/ Norwegian, Turkish-German and Greek-Albanian (for a review see Gagarina (2016) and Gagarina, Klop, Tsimpli, & Walters (2016)). Importantly, narratives can be used to explore the bilingual child's developing language in detail, as they yield more qualitative data. In fact, Gagarina and collaborators (2016) and Iluz-Cohen and Walters (2011) argue that narrative tasks can be richly exploited in examining children's language development and can complement the results from psychometric measures used in child language assessment.

The present thesis explores the narrative skills in relation to the narrative macrostructure, i.e. the ability to tell a coherent story. Macrostructure is related to the concept of story structure which includes referring to the setting (time and place), and the episode structure. The episode structure is composed of an initiating event that triggers the protagonist's response, internal responses (the internal state of the protagonist as a response to the initiating event), the goal of the protagonist, the attempt to reach the goal, and the outcome of the undertaken actions (Gagarina, 2016; Stein & Glenn, 1975). Thus, the references to internal states partially feed the story structure, and the macrostructure as well (Gagarina et al., 2012; Kunnari, VäLimaa, & Laukkanen-Nevala, 2016; Maviş, Tunçer, & Gagarina, 2016). On that account, macrostructure may moderately reflect the child's capacity to view the protagonists as mental agents with goals and intentions.

Children generally begin to construct simple narratives at the age of 2, starting with local narration, i.e. describing story events in isolation, but around the age of 4 they start to chain events sequentially (Berman, 1988). At the age of 4 they also start to include internal state terms as initiating events (i.e. an event that triggers the protagonist's action) (Berman & Slobin, 1995). Between the ages of 6 to 10 children proceed to the global action-structure of the narrative, centring their stories around a general goal or event (Berman, 1988; Berman & Slobin, 1995). In short, the trajectories of a narrative development involve a shift from a local to a global level of information structuring and organization, and from a bottom-up to a top-down structure. Importantly, these developmental trajectories are shared by children from different language backgrounds (English, German, Spanish, Hebrew, and Turkish: Berman & Slobin, 1995).

The comparisons of narrative abilities between bilingual and monolingual children reveal no significant differences between the groups in the area of macrostructure measures: neither in the story structure, i.e. the setting and episode structure (initiating event, goal,

attempt, outcome), nor the use of internal state terms in a narrative (Bonifacci, Barbieri, Tomassini, & Roch, 2017; Kunnari et al., 2016). Moreover, there seem to be no differences in the area of macrostructure between narratives told in the two languages of bilinguals. The macrostructure of narratives is shown to be correlated across the two languages of a bilingual child, regardless of the specific ways of coding (Fiestas & Peña, 2004; Gagarina et al., 2015; Pearson, 2002; Uccelli & Páez, 2007).

It seems that telling a narrative with a coherent structure is a language task that requires a cognitive component. Thus, discourse abilities probably tap into bilingual children's language-general capacities, and not only language-specific skills (Gagarina, Klop, Tsimpli, & Walters, 2016; Paradis, Genesee, & Crago, 2011). It is possible that there is a carry-over of the macrostructure elements across the two languages even if the child's linguistic abilities in one of the languages are weaker (Gagarina et al., 2016). Possibly, the invariant structure of narratives across the bilingual child's two languages illustrate what Cummins (1979) named the Linguistic Interdependence Hypothesis: that some of the bilingual children's skills and knowledge gained in the L1 can be transferred into their L2. In that view, the development of narrative abilities in the L2 may depend on such competence already developed in the L1.

3.2. Internal State Lexicon in bilingual children

There are not many studies that have investigated the production of internal state terms in bilingual children. Many of those that did, explored the ISL as a secondary issue, treating it as one of the indices of the narrative macrostructure. These studies suggest that bilingual children use a similar amount of internal state terms as their monolingual peers when telling a story either in their home language (Otwinowska et al., 2018) or the language of majority (Bonifacci et al., 2017). However, these studies counted the overall number of internal state terms, and did not differentiate between the particular subclasses of internal state terms. Two studies looked closely into the bilingual use of the different subclasses of internal state terms.

Both were relatively small studies, with sample size of c.a. 20-25 participants. Shiro, Hoff, Ribot and Shanks, (2017) and Shiro, Hoff, and Shanks (in press) studied the production of internal state terms of 30 month old American English monolinguals (n = 24) and Spanish-English bilinguals (n = 23). The children's production of the terms was measured during a 30-minute structured play interaction with the mother. The authors distinguished between three subclasses of the ISL: emotion terms (including references to feelings, physical states, likes and dislikes), volition terms (including references to intentions, desires, wishes, promises), and cognition terms (including references to belief, certainty and uncertainty, perceptions, mental processes). In general, bilinguals and monolinguals referred equally often to internal states during their interaction with the mother (internal state terms constituted c.a. 10% of total speech in both groups). When the overall number of internal state terms was considered, bilingual and monolingual children did not differ: the children produced a similar total amount of types (the number of unique terms) and tokens (the number of occurrences of each type of a term). Interestingly, bilinguals often code-switched between the two languages, producing the terms both in English and Spanish. The analyses of the subclasses of internal state terms showed that monolinguals used more types when referring to cognition and volition states than bilinguals (either in English or Spanish). The authors conclude that bilinguals exhibit a smaller internal state vocabulary than monolinguals, in reference to cognitive and volition terms.

Altman, Armon-Lotem, Fichman, & Walters (2016) examined bilingual cross-language production of 6 subclasses of internal state terms (verbs only: perceptual, mental, motivational, emotion, physiological, linguistic, and consciousness verbs) during a picture book narration. The participants were English-Hebrew bilingual preschool children aged 5 – 6.5. Additionally, the bilinguals were either typically developing (n = 19) or with a diagnosis of Specific Language Impairment (n = 12). The cross-language comparisons of the overall number of produced internal state terms revealed that both groups (TD and SLI bilinguals) used significantly more

internal state terms in their L2 (Hebrew) narratives than their L1 (English) narratives. Interestingly, there were no differences in the total amount of internal state terms between the two groups of bilingual children (TD vs. SLI). A closer look at each of the subclasses of the ISL revealed that mental and consciousness verbs were used differently in the L1 and the L2: all children's narratives contained significantly more mental verbs (e.g. "think", "know") in the L2 (Hebrew) than in the L1 (English), but the L1 (English) narratives contained significantly more consciousness verbs (e.g. "asleep", "awake") than the L2 (Hebrew) stories. However, the authors point out that the L2 narratives were shorter, and lower frequencies of internal state terms could have led to higher ratios in the L2 than in the L1. They also suggest that the use of internal state terms in the bilingual's two languages may be mediated by the effects of exposure. It was possible that the mental terms were more frequent in the L2 because they were also more frequent in the L2 input towards the child: the children were attending Hebrew preschools (for at least two years, at the time of the study), where mental verbs such as "think", "know", "forget", "decide", "believe" were assumed to be used on a daily basis. On the other hand, the authors assume that the consciousness verbs (such as "asleep", "awake") may be more characteristic of the home (L1, English) setting.

In a previous analysis of a partially overlapping sample, Otwinowska and collaborators (2018) compared the overall amount of internal state terms produced by Polish-English bilingual children (aged 3 to 7 years old) in their Polish and English narratives. The results showed that children produced similar overall amount of internal state terms in both languages. However, this study did not differentiate between the particular subclasses of internal state terms. The limited number of studies focusing on the ISL in bilingual children, and the mixed results coming from these studies suggest an evident need for more research, particularly of the kind that focuses on the use of internal state terms in both languages of a bilingual child.

Chapter 4: The current study – analysis of ISL in bilingual and monolingual children’s narratives

The present chapter describes the analysis of production of internal state terms in Polish-English bilingual children at pre- and early school age. It starts by identifying the pending and open questions about the ISL in bilingual children and sketching a preliminary overview of the research questions and hypotheses of the present study (Section 4.1.). The next Section (4.2.) specifies the methodological presumptions that underlie the present analyses and which are based on the literature reported in the previous chapters. Section 4.3. describes the methods. First, it states in detail the research questions, their operationalization and the hypotheses (Subsections 4.3.1 and 4.3.2.). Since the thesis is a secondary analysis of existing data, i.e. the current analyses are based on a large dataset collected within the Bi-SLI-Poland project, the next Subsection (4.3.3.) describes the source of the data: the original projects and their research questions. It will be clear that the aims and the research questions of the present analysis are distinct from the aims and research questions of the original projects. Subsection 4.3.4. goes on to provide details on the participants (bilingual and monolingual children). Subsection 4.3.5. describes the general procedure for conducting the study and the tools used in the study.

4.1. Pending questions about ISL in bilingual children

This section presents an overview of the current study rationale. A detailed description of the research questions and hypotheses is provided at the beginning of Section 4.3. The ISL production in children has been equally an interest and a challenge to researchers studying child development. The nature of the ISL – the abstractness and polysemy of words, the fact that these words constitute a small proportion of the child’s lexicon – has made them difficult to study and analyse. At the same time, the use of internal state terms in children may inform us in more depth about their mentalizing abilities, or their readiness to demonstrate those abilities.

This is meaningful as the small or no use of the ISL is characteristic of atypical development, e.g. children with ASD use fewer cognitive terms than typically developing peers (Rumpf et al., 2012).

The present thesis explores the use of ISL in Polish-English bilingual children. There is an evident lack of research on the ISL in bilinguals, even though their general linguistic and cognitive development is comprehensively examined. In fact, bilingual upbringing may shape the bilingual use of the ISL to look differently than the ISL in monolingual children. First, bilingual children receive less input in each of the languages in comparison to their monolingual peers (e.g. De Houwer, 2014). And second, there is a dual language activation in the bilingual minds which forces bilinguals to constantly exert control over the two languages. Each of these phenomena result in clear differences between the bilingual and monolingual children. Bilinguals are shown to have smaller vocabularies in each of their languages, but at the same time they are reported to outperform monolinguals on false-belief tasks, commonly used to measure theory of mind (Farhadian et al., 2010; Goetz, 2003; Kovács, 2009). The source of bilingual advantage on the false-belief tasks lies partially in their better developed interference suppression (ignoring or suppressing a faulty response, see e.g. Bialystok, Craik, Green, & Gollan, 2009) and enhanced ability to override an egocentric bias (e.g. Leslie et al., 2005). These are related to the performance on the false-belief tasks in preschool children (Carlson & Moses, 2001; Leslie, German, & Polizzi, 2005, see Subsection 3.1.1). These two contrasts (smaller vocabulary size and better ToM) might influence the ISL acquisition in bilingual children, as compared to their monolingual peers. They may also influence the bilinguals' ISL use across their two languages. Neither of the matters have been conclusively explored by research so far.

The current analysis aims to answer the following questions. First, it evaluates whether language status (bilingual vs. monolingual) may influence the use of ISL. Specifically, it will

be checked whether bilinguals and monolinguals differ in their use of internal state terms when telling a fictional story based on a set of pictures. Previous research has indicated some group differences only in particular subclasses of ISL. Shiro et al. (2017, in press) found similarities between bilinguals and monolinguals in the overall number of internal state terms used during play time with the mother, but found that the monolinguals referred more frequently to the cognition and volition states than the bilinguals. In theory, the potential differences in the ISL between bilinguals and monolinguals could be an effect of lower language abilities of bilinguals: especially vocabulary has been shown to influence the use of ISL (Antonietti et al., 2006), and thus smaller vocabulary of bilingual children might negatively affect their production of internal state terms. On the other hand, these constraints might be levelled out by a general advantage in theory of mind reported for bilinguals (Farhadian et al., 2010; Goetz, 2003; Kovács, 2009).

Second, the present analysis investigates whether the ISL develops comparably in both languages of the bilingual child. Altman, Armon-Lotem, Fichman, and Walters (2016) showed that bilingual pre-school children used more mental terms in their L2 (Hebrew) narratives than their L1 (English) narratives. The authors explain the higher frequency of mental terms in the L2 in the light of higher exposure to these words in the L2 (majority language) input. However, it cannot be ruled out that the use of internal state terms across the bilingual's two languages is highly correlated, if not interdependent. In fact, if the words are present in the bilingual child's input in both the home language (i.e. from family members), and in the majority language (i.e. from caretakers, school teachers, books), then the child might acquire the ISL in both languages. Moreover, the child's readiness to refer to internal states of story protagonists, as a component of mentalizing ability, should be primarily related to the child's development of theory of mind, and subsequently expressed via language. Thus, the use of internal state terms should rely first on the development of the awareness that people's actions are governed by internal states,

which is invariant of the child's language proficiency. Second, the use of IST relies on the acquisition of the terms. In case of bilingual children, these terms should be acquired in both languages. Once a bilingual child has developed the mentalizing ability, and the expression of this ability is not hindered by the lack of knowledge of a word in either language, he or she might use internal state terms similarly in both languages.

Third, the current analysis aims to identify the best predictors of the ISL production in the narratives of bilingual children in Polish and English. The possible predictors include the children's age, theory of mind performance, vocabulary size in each language, receptive grammar knowledge in each language, the number of total words produced in a story, and the general ability to tell a coherent story (story structure). In general, studies on monolingual preschoolers found a steady increase over time in the number of internal state terms understood (Booth & Hall, 1995; Astington & Pelletier, 2003) and used by children (Becker Razuri, Hiles Howard, Purvis, & Cross, 2017b; Rudek & Haden, 2005). However, as to other variables, research has yielded largely inconclusive results: some studies found no statistically significant relation between the amount of the ISL and general vocabulary knowledge (Meins et al., 2006) or theory of mind performance (Charman & Shmueli-Goetz, 1998; Meins et al., 2006), while others did (Antonietti et al., 2006; Symons et al., 2005). This issue will be discussed in detail in Subsection 4.3.2. (The hypotheses).

Last, the current analysis aims to assess the narrative context as a potential source of the ISL and a medium of ISL enhancement. These two additional aims are focused on the narrative tools, instead of the bilingual population as such. First, it will be investigated whether the ISL production in a narrative context can be immediately improved by presenting a child with a model story told by an adult. Previous research found that when retelling stories after a given model, children tend to address more goals of the story characters and use more elaborate vocabulary (Isbell et al. 2004; Peterson and McCabe 1992). Retelling stories after a model has

been used also as a strategy to train children's understanding of the internal states of the story protagonists (Lewis et al., 1994). In fact, a model story may sensitize children to the internal states of the characters and stimulate them to refer more to those states. Second, it will be checked whether the use of ISL when telling a story differs from the ISL use in a comparatively more interactive context, i.e. when children are explicitly asked about the internal states of the story protagonists and why particular characters did what they did. Meins et al., 2006 found that children use more internal state terms during an interview when they are asked to describe their friend, compared to when they simply tell a story based on a picture book. It is possible that the tendency to spontaneously label internal states of story protagonists is independent of the child's ability to simply observe those states (i.e. without commenting on them).

To summarize, the present analysis will explore the productive Internal State Lexicon in bilingual children, by comparing their production to that of monolinguals and by comparing the production across the two languages of the bilingual children. Additionally, the ISL in story-telling context will be juxtaposed with the ISL use after a model and the ISL use in a relatively more interactional context. The planned analyses are based on some specific presumptions derived from a careful review of the literature background presented above. Let us now turn to specify each of the methodological presumptions of this thesis.

4.2. Methodological presumptions of the present analysis

The present study is based on a series of presumptions established in the course of previous research on the ISL in children, as outlined in the first three chapters of the thesis. Thus, the present study investigates two most common subclasses investigated in other studies: the emotional terms and the mental terms (that include both cognitive and desire terms). Additionally, it explores the use of perceptual terms. Though the perceptual terms are rarely studied, children may use them apart from mental terms to attribute knowledge, or lack of it. In the present thesis, the Internal State Lexicon is understood to include all words related to the

internal states, independently of the grammatical category: verbs and nouns, adjectives and adverbs (Hall & Nagy, 1979). It is also assumed that ISL constitutes a small portion of child's overall lexicon and thus it is characterized by a relatively lower frequency in child-produced speech (Bartsch & Wellman, 1995; Slaughter et al., 2007; Tardif & Wellman, 2000). Moreover, children are assumed to use internal state terms both with a mentalistic and a non-mentalistic meaning, i.e. as an idiomatic or conversational device without indicating the actual mental states (Nielsen & Dissanayake, 2000; Shatz et al., 1983). The present analysis is focused on the mentalistic use of internal state terms and excludes the instances when children use the terms or as an idiomatic or conversational device (e.g. "and they lived happily ever after"). As such, it is assumed that production of internal state terms can serve as an index of the ISL acquisition and of the general mentalizing ability in children. That is, it is understood that children who use internal state terms attend to internal states of self and others (Carpendale & Lewis, 2004). Last, it is assumed that narrative tasks in which children are asked to tell a story based on a specific picture-book or a set of pictures is an adequate technique for eliciting internal state terms. That is because: (1) the story context draws the child's attention not only to the action itself, but also to how this action is perceived by the individual story protagonists (Bokus, 2013; Bruner, 1986); (2) the story context is a natural and common context for interaction in families; (3) story-telling should reveal real-life variability of the responses, i.e. some children may use many internal state terms but since the story can be well told with a sole focus on the landscape of action, some children may use no internal state terms in their stories (Symons et al., 2005). These are assumptions treated as established research-evidenced arguments on which further research questions can be built. Therefore, we can now turn to the description of the aims and detailed research questions taken up in the present thesis.

4.3. The methods

4.3.1. The aims and the research questions of the current study

The present analysis was designed primarily to investigate the productive Internal State Lexicon (ISL) in Polish-English bilingual children at pre- and early school age. The bilinguals' performance is explored across their two languages and their performance in Polish is compared to that of matched Polish monolinguals peers. The results should inform us whether language status (bilingual vs. monolingual) may influence the use of Internal State Lexicon, and whether the ISL develops comparably in both languages of the bilingual child. The analysis will also attempt to identify the predictors of the ISL production in the narratives in Polish and English.

In addition, the analysis aimed to assess the narrative context as a potential source of the ISL and a medium of the ISL enhancement. To this end, two additional aims were elaborated, ones that were not focused on the bilingual population per se, but on the nature of the used tool. These two aims will investigate whether the ISL production in a narrative context can be immediately improved by presenting a child with a model story told by an adult, and whether the ISL use in the narrative non-interactive contexts differs from the ISL use in relatively more interactive context. In order to investigate whether the effects of retelling and the effects of interactive vs. non-interactive context are global, the analyses for these aims will be carried out on both bilingual and monolingual children, and on narratives in both languages of the bilingual children.

The following research questions are posed:

RQ1: Do bilingual and monolingual children differ in the amount of internal state terms (IST) used when telling a story in Polish?

RQ2: Do bilingual children differ in the amount of IST used when telling the stories in Polish and English?

RQ3: What are the best predictors of the use of IST in the narratives told in Polish and English?

RQ4: Do children differ in the amount of IST when telling a story, as compared to when they retell a story immediately after listening to a model story told by an adult?

RQ5: Do children differ in the use of IST when explicitly engaged in a conversation about the internal states of story protagonists, as compared to the use of IST in their spontaneous telling?

4.3.2. The hypotheses

Each hypothesis corresponds to the relevant research question specified above, e.g. H1 corresponds to RQ1. After each hypothesis, a short rationale is given, based on the literature reviewed in the previous chapters.

H1: The bilingual and monolingual children do not differ in the amount of internal state terms used when telling a story in Polish.

This hypothesis was based on results from Shiro and collaborators (Shiro et al., 2017, in press), showing that bilinguals and monolinguals do not differ in the overall amount of internal state terms produced in interactions with mothers. However, Shiro found differences between the groups in the number of types when referring to the specific subclasses of internal states: monolinguals used more types when referring to cognition and volition states than bilinguals (either in English or Spanish). Hence, the hypothesis was drawn with considerable caution. Even more so as there is ample evidence showing that bilingual children have smaller vocabularies in each of their languages, when compared to monolingual peers (for comparisons with majority language peers see: Bialystok, Luk, Peets, & Yang, 2010; De Houwer, Bornstein, & Putnick, 2014; Hoff et al., 2012; Hoff, Rumiche, Burrige, Ribot, & Welsh, 2014; E. Thordardottir, 2011; for comparisons with home language peers see: Haman et al., 2017; Mieszkowska et al., 2017). Moreover, some studies on pre-school children point to a direct link between the child's receptive vocabulary knowledge and their understanding of internal state

terms (Antonietti et al., 2006). Thus, the relatively smaller vocabulary of bilingual children might negatively affect their knowledge of internal state terms (as compared to monolinguals): not only understanding but also production of the ISL. On the other hand, bilingual preschoolers are shown to perform better than monolingual peers on the theory of mind tasks (Farhadian et al., 2010; Goetz, 2003; Kovács, 2009). Importantly, performance on the false-belief tasks in early school children has been shown to correlate with their use of internal state terms (Nielsen & Dissanayake, 2000). Thus, it cannot be completely ruled out that bilinguals in the present analysis will use the same amount or more internal state terms than monolinguals, due to better mentalizing abilities.

H2: Bilingual children do not differ in the amount of internal state terms used when telling the stories in Polish and English.

This hypothesis is in line with the Interdependence Hypothesis (Cummins, 1979) which suggests that bilingual children's skills, metalinguistic and pragmatic knowledge gained in the L1 can be transferred into their L2. Also, previous analyses on a sample partly overlapping with the present one (Otwinowska et al., 2018) showed that bilinguals used a similar overall amount of internal state terms when telling their stories in Polish and English. However, the results obtained by Altman, Armon-Lotem, Fichman, & Walters (2016) showed that bilingual preschool children used significantly more cognitive (or mental) terms in their L2 (Hebrew) narratives than their L1 (English) narratives. Though the results from Altman et al. (2016) could be due to some unavoidable limitations (small sample, the fact that the L2 narratives were shorter, and lower frequencies of internal state terms could have led to higher ratios in the L2 than in the L1), they could also reflect the effects of exposure: the higher frequency of mental terms in the L2 (majority language) narratives was assumed to be due to higher exposure to these words in the L2 input towards the child. If the latter were true, we might observe some different cross-language tendencies in regards to the number of particular subclasses of internal

state terms, e.g. mental terms might be more frequent in the L2 stories (English, majority language) than in the L1 stories (Polish, home language).

H3: There is a positive relationship between the use of internal state terms in a narrative and the children's age, vocabulary size and receptive grammar knowledge in each language, theory of mind development, the number of total words produced in a story, and the general ability to tell a coherent story.

This hypothesis presupposes a general positive relationship between the use of internal state terms and each of the mentioned variables, but does not determine beforehand the actual best model predicting the number of IST in a Polish or English narrative. In general, studies on monolingual preschoolers found a steady increase over time in the number of internal state terms understood (Booth & Hall, 1995; Astington & Pelletier, 2003) and used by children (Becker Razuri, Hiles Howard, Purvis, & Cross, 2017b; Rudek & Haden, 2005). However, evidence regarding the link between ISL and the remaining variables is still largely inconclusive. Antonietti et al. (2006) found a positive correlation between receptive vocabulary knowledge and understanding of internal state terms in 4-8 year old children. However, Meins et al. (2006), who studied children at the age of 7-9 years old and explored the use of internal state terms while narrating a picture story and describing a friend, failed to establish any link between the children's receptive vocabulary knowledge and the use of internal state terms. Moreover, in their study, the use of internal state terms was not significantly correlated with the performance on the ToM task. Similarly, Charman and Shmueli-Goetz (1998) also failed to establish a significant relationship between 7-year olds' theory of mind performance and their use of internal state terms when narrating a picture book. Contrary to those results, Nielsen and Dissanayake (2000) found a significant positive correlation in slightly younger children at preschool age: the children's performance on four false-belief tasks was strongly and positively correlated with their use of internal state terms during free play with a parent. No previous

studies have directly explored the relation between receptive grammar knowledge and understanding or production of internal state terms. Importantly, to the best of the author's knowledge, no studies have explored the relation between the above mentioned variables and the ISL production in narratives told in both languages of bilingual children.

H4: Children use more internal state terms when they retell a story immediately after listening to a model story told by an adult, as compared to when they tell a story by themselves.

This hypothesis was drawn on findings from studies on monolingual children showing that when retelling stories after a given model, children tend to address more goals of story characters and use more elaborate vocabulary (Isbell et al. 2004; Peterson and McCabe 1992). Retelling stories after a model has also been used as a strategy to train children's narrative abilities (Spencer & Slocum, 2010) and their understanding of the internal states of the story protagonists (Lewis et al., 1994). Thus, a model story told by an adult may draw the children's attention to the internal states of the characters and encourage them to use more internal state terms. The hypothesis is tested on both bilingual and monolingual children, and in the Polish and English stories of bilinguals. This way it will be established whether the retelling effect on the ISL is global and observable in both groups and across the bilinguals' languages.

H5: Children use more internal state terms when explicitly engaged in a conversation about the internal states of story protagonists, as compared to the amount of internal state terms in their told stories.

This hypothesis was based on the evidence showing that monolingual children tend to use a lot of internal state terms when interacting with peers and adults (Brown et al., 1996; Hughes, Lecce, & Wilson, 2007; Pinto et al., 2016), but they may use comparatively fewer internal state terms when the context is non-interactive, e.g. when telling a story (Meins et al., 2006). This may suggest that developing theory of mind does not necessarily translate to spontaneously using the ToM abilities to describe, explain and interpret the behaviour of others

(Meins et al., 2006). To test whether this effect is global, the analysis is performed on both bilinguals and monolinguals and on both languages of the bilingual children.

4.3.3. The source study

As indicated in the Introduction, the data used for the current analysis comes from two research projects. The source study was conducted within the Bi-SLI-Poland project and its data was further coded and analysed from various angles within the WLRB project. It is essential to note that the analyses presented in this thesis are far beyond the scope of either of the two projects. The Bi-SLI-Poland project “*Cognitive and language development of Polish bilingual children at the school entrance age - risks and opportunities*” was designed to examine and describe the typical developmental patterns of Polish-English bilingual children living in the UK, which would eventually help to identify Polish-English bilingual children at risk of Specific Language Impairment. Specifically, Bi-SLI-Poland aimed to: (1) create a profile of a typical language development of Polish children acquiring English language in a natural context of bilingualism; (2) define chances and potential threats stemming from bilingualism and working out methods of avoiding those threats; (3) identify problems in the linguistic development of Polish bilingual children which meet SLI criteria to enable early intervention; (4) develop methods and research techniques which could possibly become a base for diagnostic tools for probing linguistic development of Polish mono- and bilingual children; (5) spread knowledge of typical development of bilingual children and SLI among parents and specialists in education and speech therapy. Over 180 bilinguals were tested in the UK⁶ with a battery of tasks that measured their linguistic development (e.g. vocabulary, grammar knowledge, phonological processing, narrative abilities) and cognitive development (e.g. working memory, executive functioning, theory of mind). The task battery and the sample were

⁶ The testing in the Bi-SLI-Poland and WLRB projects included also a group of over 40 Polish-English bilinguals tested in Poland (children attending bilingual pre-schools and schools), a group of over 300 Polish monolingual peers (tested in Poland) and a group of 30 English monolingual peers (tested in the UK).

large (in order to enable description of typical development in these bilinguals) and the accumulated data could be used to explore more research questions than those specified in the Bi-SLI-Poland project. Thus, another research project was designed, “*Phonological and Morpho-syntactic Features of Language and Discourse of Polish Children Raised Bilingually in Migrant Communities in Great Britain*” (WLRB) with the aim to test a small control group of English monolingual children and to describe in detail the phonological and morpho-syntactic system as well as the discourse features of the same Polish-English bilingual children tested in the Bi-SLI-Poland project. One of the results of the WLRB project was a corpus of speech samples of bilingual and monolingual (Polish and English) children. This subsequently underwent a detailed linguistic analysis in order to investigate possible differences between the phonological and morpho-syntactic competences and discourse abilities of bilingual and monolingual children.

The aims and the research questions of the present analysis (specified below) are distinct from the aims and research questions of either of the two projects (Bi-SLI-Poland and WLRB). Consequently, the analyses carried out for the purposes of the present thesis were done in addition to the analyses carried out within the two projects.

4.3.4. Participants

The sample studied in the current thesis comes entirely from the Bi-SLI-Poland project. In the Bi-SLI-Poland, there were altogether over 180 Polish-English bilinguals tested in the UK. The children were aged 4.5 to 7 years old, so they were pre- and early school children. Most of the bilingual children in the sample were attending primary school (61 children, 81%). According to their age, they were either enrolled in Key stage 1 (5 year olds and older) or in Reception (4 year olds). Six of the 4 year old bilinguals (8%) were attending nursery. The formal primary education in the UK starts when the child turns 5 years old (Key stage 1). Before that, 4 year olds are entitled to 15 hours of free nursery education for 38 weeks of the year or

they can start a Reception year in the primary school⁷. In Poland, in 2013/2014, when most of the monolinguals were tested, parents could decide whether their 6 year old children were to pursue obligatory pre-school education (“0” grade) or to start primary school (1st grade). Not all parents participating in the study reported what school their children were attending, but of those who did (47 parents, 63%), all but one reported their children to attend preschool. In fact, in 2013/2014, only 15,5% of 6 year olds were enrolled in the first grade of the primary school⁸.

The bilingual families were recruited through Polish Saturday schools, Polish catholic churches in London, Cambridge and Manchester, via the internet, social media, and through the word of mouth from parents who already took part in the study. A written parental consent was obtained for all the children participating in the study. Children also expressed their oral consent to take part in the testing. The whole procedure included a battery of over 20 tasks that measured the children’s linguistic development in Polish and English (e.g. vocabulary, grammar knowledge, phonological processing, narrative abilities) and cognitive development (e.g. working memory, executive functioning, theory of mind). The current study included data from all the bilingual children who have narrated a story in both languages, and have done the following tasks: the non-verbal IQ test (Raven’s Colored Progressive Matrices, Raven, 2003), reflection on thinking test (Białecka-Pikul, 2012), and the expressive vocabulary tasks in both Polish and English (for a detailed description of each task, see the next Section 4.3.6.). There were 96 participants fulfilling the above criteria. Some exclusions from this group were needed: 6 bilingual children were excluded as they were reported by the parents to have had hearing problems in the past. 15 children were excluded because they turned to be effectively trilingual (i.e. had one parent of neither Polish nor English origin, who interacted with the child always

⁷ An overview of the National Curriculum in the UK can be found here: <https://www.gov.uk/national-curriculum>.

⁸ GUS Statistics, <https://stat.gov.pl/obszary-tematyczne/edukacja/edukacja/oswiata-i-wychowanie-w-roku-szkolnym-20132014,1,8.html>

in his/her native language). As a result, the present analysis included data from 75 Polish-English bilingual children living in the UK. The bilingual children were born to families with at least one Polish parent, and 49 out of 75 bilingual children had both Polish parents. Out of the remaining 26 children, 9 had English fathers and 17 had fathers of other nationality: Italian, Nigerian, Algerian, Bamar, Jordanian, Hungarian. However, most (13 of the children) were reported to have scarce contact with the third (father's) language, and the remaining 4 children were reported to have low fluency in the third language. These four children were left in the sample because their language experience accurately reflects the diversity of the population. All the bilinguals were exposed to Polish from birth, and their average onset of acquisition of English was around their first birthday ($M = 13$ months, $SD = 15$ months, range: 0-48 months).

Next, 75 Polish monolingual children living in Poland were carefully matched one-to-one to the bilingual group. The matching was done in the following way: for each child from the bilingual group ($n=75$), a comparable monolingual child was found from a larger sample of 93 Polish monolingual children who performed the same tests in Polish. The comparison was made in terms of the children's gender, age, and non-verbal IQ score (equal weights)⁹. Thanks to the matching procedure, the two groups were comparable in the potential interfering factors such as gender, age, and non-verbal IQ. The groups were also similar in socio-economic status (SES) as measured by the number of years of maternal education (see Table 1 for details).

⁹ The matching was performed by the author using a script in Python created by dr Jakub Szewczyk (Jagiellonian University).

Table 1.

Characteristics of the participants: gender, age (in months), non-verbal intelligence score and maternal education (in years) across bilinguals and Polish monolinguals.

	PL-EN bilinguals (n = 75)	PL monolinguals (n = 75)	Between-groups comparison:
Gender	45 f + 30 m	45 f + 30 m	$\chi^2(1) = 0, p = 1$
Age (months) M \pm SD [range]	68 \pm 9 [53-83]	68 \pm 8 [53-83]	$t(148) = 0.01,$ $p = 0.99$
Raven (raw score) M \pm SD [range]	23 \pm 5 [13-34]	22 \pm 5 [11-32]	$t(148) = 0.55,$ $p = 0.58$
Years of maternal education M \pm SD [range]	16 \pm 3 [11-24]	17 \pm 3 [12-24]	$t(148) = -1.11,$ $p = 0.27$

4.3.5. General procedure and tools

The children participating in the Bi-SLI-Poland project were tested with a large battery of over 20 tasks performed in Polish and English, measuring their vocabulary size (receptive, productive), grammar abilities (receptive, productive), phonological processing (non-word repetition task), narrative abilities (narratives), and an array of cognitive abilities, including their executive functioning, working memory (verbal and non-verbal), and theory of mind development. The bilingual children were tested mostly at home, while the monolingual children were tested mostly at schools and preschools. The difference in the testing place of the two groups was a result of organizational difficulties. The bilingual children were recruited through the internet and the word of mouth, thus it was the parents signing up the children for the testing. The children came from different schools and thus it was impossible to test them in one place. The monolingual children were recruited through schools and preschools that enabled us (with parental consent) to test the children at their premises. The experimenters who

performed the testing were carefully trained in all the task procedures and had to carry out a training testing, evaluated by the research coordinators. The experimenters for the Polish testing (of bilinguals and monolinguals) were Polish native speakers. The experimenters for the English testing (of bilinguals) were either native speakers of English, or Poles that emigrated to the UK and were fully proficient in English. Each child was tested individually. The bilinguals were tested across c.a. 8 sessions in both languages, by two experimenters, one using only Polish with the child, the other using English. Each session lasted no more than an hour (including breaks). The tasks that measured cognitive development, including the reflection on thinking task, were performed in the dominant language of a bilingual, as reported by parents. The monolingual testing was spread to 4 sessions, as all the language tasks were performed in Polish only. Each session could be terminated or paused upon the child's request, or if the experimenter saw that the child was tired or distracted. To the best of the researchers' efforts, the whole testing was to be carried out within 4 weeks from the date of the first session. There were two schemas of testing that differed in the order of the tasks. The experimenters were also allowed to change the order of the tasks in their assigned scheme, e.g. due to technical difficulties or if the child refused to perform a particular task. The present analysis makes use of four of the measures used in the Bi-SLI-Poland project: grammar knowledge, vocabulary knowledge, theory of mind, and narratives.

Receptive grammar knowledge. The children's understanding of grammatical structures was tested with the Test of Reception of Grammar– TROG-2 in English (Bishop, 2003) and with its Polish translation (Smoczyńska, 2008, unpublished). The English version of TROG was standardized on 792 children aged 4 to 16 years old. The TROG measures comprehension of 20 grammar constructions, including e.g., negatives, singular and plural inflection, object and subject relative clauses. Both language versions consist of 80 items each, i.e. 80 target grammar structures. The child is presented with a board of four pictures at a time, hears a

sentence containing the target structure and is asked which of the pictures best represents what he/she has heard. One of the pictures illustrates the target grammatical structure and three constitute the lexical and grammatical foils to this structure. The structures are presented with rising difficulty. For each correct answer the child scores one point, and the maximum number of points is 80.

Receptive vocabulary. The children's receptive vocabulary in Polish was measured with *Obrazkowy Test Słownikowy, OTSR* (The Picture Vocabulary Test – Comprehension; Haman, Fronczyk, & Łuniewska, 2012). This is a published test, normed on a Polish monolingual population (children aged 2;0 to 6;11) and designed to assess the comprehension of nouns, verbs, and adjectives. The OTSR includes two fully comparable versions of the test (A and B) to allow more data points in the assessment. Each version includes 88 items that are ordered with rising difficulty. Each test item is presented in a board of four pictures: one picture depicts the target word, one phonetic foil, one semantic foil, and one thematic foil. The children were presented with one board at a time and asked to point to the one of the four pictures that appropriately depicted the target word. The children did both versions of the test, with the order of the versions counterbalanced. Depending on the child's age, the easier, initial items were skipped in each version. The procedure in each version was terminated after four consecutive errors. A child could score a maximum of 88 points in each version of the test (one point for each correct answer). The present analysis considered only one of the test versions, for which a child obtained a higher score. This was done because sometimes after performing one version of the task, the children grew impatient during the second version and asked to terminate the task prematurely. Also, if a child performed only one of the versions, the score from that version was used for the analyses.

The bilingual children's receptive vocabulary in English was tested with *British Picture Vocabulary Scale – 3rd edition (BPVS-3; Dunn, Dunn, & Styles, 2009)*. The BPVS-3 is a

published test, normed on the British population (children aged 3;0 to 16;11). It is designed to assess the child's comprehension of nouns, verbs, and adjectives. The BPVS-3 includes 168 items, grouped in 12 sets ordered with rising difficulty. Similarly to the OTSR, each test item was presented in a board of four pictures and the children were asked to point to the one that appropriately depicted the heard target word. The testing started with a set specified by the child's age, i.e. the easier, initial items were skipped depending on the child's age. The testing was terminated when the child made 8 or more errors in a set.

Expressive vocabulary. The children's expressive vocabulary in Polish was assessed with *Zadanie Nazywania Obrazków* (ZNO, The Picture Naming Task; Haman & Smoczyńska, 2010, unpublished). It is designed to investigate the child's expressive (productive) knowledge of nouns and verbs. The ZNO consists of 53 color pictures depicting 32 nouns and 21 verbs. The children were presented with all the 53 pictures (one at a time) and asked to name each picture with one word. The test was administered from the first to the last item, regardless of the number of errors made by the child. A child could score a maximum of 53 points (one point for each correct answer). The correct answer included the target word, its close synonym, or a dialectal variant.

The bilingual children's productive vocabulary in English was assessed with the Expressive Vocabulary Test – 2nd edition (EVT-2; Williams, 2007). The EVT-2 is a published test, normed on the monolingual American population (aged 2;6 to 90 years old and older). It is designed to investigate the child's productive (expressive) knowledge of nouns, verbs and attributes. The EVT-2 consists of two parallel versions (Form A and B), of which each includes 190 items. The children were tested with only one version of the test, Form A. The children were presented with one picture at a time and asked to name the picture with one word. Depending on the child's age, the easier, initial items were skipped. The testing was terminated

after 5 consecutive errors. The correct answer included the target word, its close synonym, or a dialectal variant.

Theory of mind. The children's theory of mind performance (more specifically: the reflection on thinking) was measured with the Test Refleksji nad Myśleniem (TRM, Reflection on Thinking Test; Białecka-Pikul, 2012). The TRM was developed for children over 4 years old and constitutes a battery of 9 tasks in the form of illustrated stories that describe the actions of two protagonists. The events relate situations where the protagonists hold different beliefs about some objects (e.g. their identity, their location), which requires inferring about the minds of others. More specifically, the tasks assess the child's understanding of 1st and 2nd order beliefs, understanding of deception, ambiguity and interpretation, and understanding of surprise. The test was performed only in one language of the bilingual children: the one that parents reported the child to be more at ease in (59 children in the sample took the test in Polish and 16 children took the test in English). In line with the standard procedure of the TRM, the children listened to the stories told by the experimenter (aided by pictures) and after each story, they were asked to explain or predict the behaviour of one of the protagonists. The TRM yields two indices based on the children's answers: appropriateness of reflection (response accuracy) and the intensity of reflection. The appropriateness of reflection index is a sum of points scored in the questions concerning the behaviour, thoughts or emotions of one of the protagonists (e.g. "what will she do?", "what will she think", "how will she feel?"). A child could score 1 point for a correct answer and 0 points for an incorrect answer, a "don't know" answer or a lack of an answer. A child could score a maximum of 8 points on the appropriateness index. The index of intensity of reflection is a sum of points scored in the why-questions. Of interest here is the level to which the child referred to the mental states of the protagonists (their knowledge, belief, thoughts) when interpreting their behaviour. For example, in one of the stories, a girl-protagonist, Maya, relocates a book, while Hollie is not watching. The children were first asked

“Where will Hollie think the book is?” (appropriateness of reflection question), and then: “Why will Hollie be looking there?” (intensity of reflection). To the letter, a child could respond “I don’t know” (0 points), “because she put it there/ because she saw it there/ because she wants to find it” (1 point), or “because she thinks it is there/ because she doesn’t know Maya put it somewhere else” (2 points). A child could score a maximum of 16 points on the intensity of reflection index.

Narrative abilities. Narrative abilities in all children were measured with the Multilingual Assessment Instrument for Narratives (MAIN, English version: Gagarina et al., 2012; Polish version: Kiebzak-Mandera, Otwinowska, Białecka-Pikul, 2012), designed to assess narrative skills in bilingual and multilingual children aged from 3 to 10 years old. The MAIN provides guidelines for evaluating the production and comprehension of narratives with the use of four compatible picture stories controlled for cognitive and linguistic complexity, parallelism in macrostructure and microstructure, as well as for cultural appropriateness and robustness (see Appendix B for the pictures). The testing procedure for the narrative task was the same in each language (Polish and English) and consisted of three stages: (1) Warm-up, (2) Narrative Telling (MAIN: Baby Birds/ Baby Goats, counterbalanced) and comprehension questions, (3) Narrative Retelling (MAIN: Dog/ Cat, counterbalanced) and comprehension questions. Throughout the testing, the experimenter and the child were seated next to each other. At the beginning, the child was asked several warm up questions, e.g. “Do you like listening to stories and fairy tales? Do you know what a story or a fairy tale always begins with/ends with?”. If the child did not know the answer, the experimenter explained how stories could begin and end. The child was also prompted to tell any story he or she wanted. Their own stories were not analysed. Then the experimenter presented the child with three envelopes and informed the child that each contained a different story. In fact, all envelopes contained the same picture story, in accordance to the testing scheme, but this was done in order to strengthen the child’s

belief that the experimenter was not familiar with the stories. The child was asked to choose one envelope.

In the telling mode that followed, the child was asked to take the picture story from the envelope, look at the pictures, and tell a story without showing the pictures to the experimenter (the child was explicitly asked not to do that). This was done to ensure the 'non-shared attention' condition, as the experimenter was only the listener and the child had to narrate alone. The experimenter prompted the child gently only if he or she could not begin, or if there was a long pause. The experimenter did not interrupt or otherwise intervene in the narrative, even if the child had problems naming the characters. Then the experimenter asked the child comprehension questions concerning the events of the story and the protagonists' internal states. Specifically, the questions targeted the protagonists' emotions (e.g. How do the baby birds feel? Imagine that the dog can see the birds. How does the dog feel?) and the protagonists' motivation for specific action (e.g. Why is the bird flying away? Why is the cat climbing the tree? Why did the dog grab the cat's tail?). For the full list of comprehension questions for each story, see Appendix B.

In the retelling mode, when the child had chosen the envelope, the experimenter and the child viewed the pictures together. First, the experimenter told the model story to the child in a friendly manner, following the script and pointing to the pictures (for story scripts see Appendix B). The model stories contained the same specific internal state terms: in each model story, there were three references to three emotional states (3/180 words): "glad", "playful", "cheerful"; six references to four mental states (6/180 words): "want", "surprised", "think", "decide"; and five references to two perceptual states (5/180 words): "see" and "notice". Altogether, the internal state terms amounted to 14 out of 180 words in each story, giving a ratio of 7.8%. After listening to the model story, the child was asked to retell the story while viewing the pictures together with the experimenter in the 'shared attention' manner. After the

retelling, the child was also asked a set of comprehension questions, comparable to those used in the telling mode.

The MAIN narratives were used to extract two measures: the story structure score (that indicates the general ability to tell a coherent story), and the amount of internal state terms. The story structure score was a maximum of 14 points: 2 points for expressing a setting, a total of 9 for the three episodes of each story: within each of three episodes 3 points were given for each element of the Goal-Attempt-Outcome (GAO) sequence (max. 3 episodes * 3 elements = 9 points); additionally, 1 point could be obtained for a full GAO sequence in each of the three episodes (max. 3 episodes * 1 full GAO sequence = 3 points). Thus, the total maximum number of points was $14 = 2$ (for the setting) + 9 (for the GAO elements) + 3 (for full GAO sequences). It is important to note that the story structure score used in the present analysis did not include the points from the references to the internal states (either in the initiating event or as a response to the initiating event), commonly included in the story structure score (see Section 3.1.3). Thus, the story structure score and the points for the internal state terms were separate. This was done to ensure that the story structure score which would be entered as one of the predictors if the use of internal state terms is not artificially inflated, i.e. that the same internal state terms will not be counted twice: once as points in the story structure, and the second time as points in the number of internal state terms produced in a narrative.

Internal state terms. The second measure extracted from the MAIN was the amount of internal state terms. Children's production in the MAIN task was coded for the use of internal state terms (types and tokens). These were used in the subsequent analyses as the number of internal state terms (the raw amount of internal state terms tokens) and the proportion of internal state terms (the number of internal state terms tokens divided by the total number of words in a story, transformed into a percentage). The internal state terms were coded in the telling mode, the answers to the comprehension questions after the telling mode, and in the retelling mode.

At the beginning of the coding, two lists of possible internal state terms were prepared, in Polish and in English, based on the terms included in the MAIN manuals (Gagarina et al., 2012; Kiebzak-Mandera, Otwinowska, Białecka-Pikul, 2012). As the coding of the narratives progressed, new items (used by the children) were added to the “dictionary lists” of possible items: items that children used in the present study. The two lists of possible items, in Polish and English, grew in parallel, e.g. if an item was added to the Polish list, its translation (with possible synonyms) was added to the English list as well. The terms were automatically counted with the use of Excel formulas. Once the coding was done, the lists in the two languages were trimmed to contain only the items that the children did actually use. For example, an item was included in the Polish list if at least one child used it in the Polish task, but it was not included in the Polish list if it was used only in the English stories. Eventually, the final Polish list contained 79 internal state terms: 11 perceptual terms, 40 emotion terms and 28 mental terms. The final English list contained 67 internal state terms: 8 perceptual terms, 36 emotion terms and 23 mental terms. The slight differences in the number of internal state terms in each list may be due to the fact that the Polish narratives were collected from 150 children (both Polish monolingual and Polish-English bilingual), while the English narratives were collected from 75 children (only Polish-English bilinguals). In the English list, 44 out of 67 items (66%) were internal states found in the Polish list as well. The most overlapping subclass were the perceptual terms (7 out of 8 items, 88% of correspondence between the lists), the mental terms showed 65% of inter-list correspondence (22 out of 36 items), and the emotional terms showed 61% of inter-list correspondence (15 out of 23 items). The two final lists can be found in Appendix C.

A second rater coded internal state terms in 15% of the narratives, i.e. 33 out of 225 narratives: 11 Polish narratives from the Polish monolinguals, 11 Polish narratives from the bilinguals, and 11 English narratives from the bilinguals. The inter-rater agreement was

calculated for the told stories, retold stories, and children’s answers to the comprehension questions. The agreement was thus calculated for 18 sets of data. The ICC values indicated satisfactory agreement in 4 sets, good agreement in 6 sets and excellent agreement in 8 sets (see Table 2).

Table 2.

Intraclass Correlation Coefficients (ICC) values indicating reliability of inter-rater agreement calculated for coding of the IST in the told stories, retold stories, and children’s answers to the comprehension questions in Polish and English.

	Polish			English		
	Telling	Comprehension Questions	Retelling	Telling	Comprehension Questions	Retelling
emotional	0.57 (s)	0.84 (g)	0.92 (e)	0.81 (g)	0.86 (g)	0.92 (e)
mental	0.94 (e)	0.76 (g)	0.75 (g)	0.95 (e)	0.7(s)	0.92 (e)
perceptual	0.93 (e)	0.65 (s)	0.97 (e)	0.96 (e)	0.63 (s)	0.77 (g)

Values less than 0.5, between 0.5 and 0.75, between 0.75 and 0.9, and greater than 0.9 are indicative of poor, satisfactory (s), good (g), and excellent (e) reliability, respectively (Koo & Li, 2016).

Chapter 5: The results

This chapter presents the results of the analysis of the productive Internal State Lexicon in Polish-English bilingual children at pre- and early school age. The bilinguals' performance is explored across their two languages and compared to that of matched Polish monolinguals peers. The chapter starts with introductory analyses (Section 5.1.): before exploring the research questions specified above, the bilinguals and monolinguals were compared in their general language performance in Polish and their theory of mind performance. These comparisons serve as a background for exploring the children's use of the ISL. Then, the bilinguals' language performance was compared across their two languages. This was done to explore their language dominance, which again could serve as a backdrop for the ISL use across the bilinguals' languages. Then, the main analyses are described (Section 5.2.). Each research question is reported in one subsection. Finally, a summary of the results is provided in Section 5.3.

Whenever the data from the tests were analyzed (e.g. lexical tests, grammar tests, ToM test), mostly the raw scores were used for the analyses. The raw scores were used in the comparisons of the bilingual and monolingual groups because then the children's performance on the same test was compared (Section 5.1.1.). In the cross-language comparisons in the bilingual children (Section 5.1.2), the raw scores were also used when grammar abilities were compared across the bilinguals' languages, since both language versions of the TROG have the same score scale. However, when receptive lexical knowledge across the bilinguals' two languages was compared, the raw scores had to be converted into percentiles. This was because the two tests have different raw score scales (e.g. the maximum possible score on the OTSR is 88, and 168 in the BPVS-3). In this case, percentiles were used since the OTSR and the BPVS-3 were normalized on respective monolingual populations. No valid language comparison could be performed for the expressive vocabulary skills: the two tests (the ZNO and the EVT-2) had different raw score scales and thus raw scores could not be used. Percentiles could not be

calculated as neither of the tests was normalized on the relevant populations: the ZNO was not normalized at all, and the EVT-2 was normalized on an American population which could misrepresent the results of bilinguals raised in the UK. Finally, z-scores could not be used to statistically compare the bilinguals' mean scores on expressive vocabulary tests across the two languages, since the mean scores would be close or equal to 0, as a result of converting raw scores into z-scores.

In the main analyses (Section 5.2.), the number of IST and the proportion of IST were used. Whenever other measures were introduced into analyses, e.g. in RQ3 (identifying predictors of the ISL use), raw scores from the tests were used in all the statistical tests. However, z-scores of expressive and receptive vocabulary scores were used to graphically present correlational analyses in RQ3, i.e. in order to show patterns from both languages on a single graph. In that case, the mean scores and the standard deviations for the z-scores were calculated on the merged populations (monolingual and bilingual Polish for the Polish tests, and bilingual English for the English tests).

5.1. Introductory analyses of the bilingual sample

5.1.1. Bilingual and monolingual comparison

In order to investigate the potential group differences, the two groups were compared on their language performance in Polish and their theory of mind. The language performance included the measures of receptive vocabulary in Polish (as measured by Obrazkowy Test Słownikowy, OTSR, The Picture Vocabulary Test – Comprehension; Haman & Fronczyk, 2012), expressive vocabulary in Polish (as measured by Zadanie Nazywania Obrazków, ZNO, Picture Naming Task; Haman & Smoczyńska, 2010, unpublished), and the receptive grammar skills in Polish (as measured by the Test of Reception of Grammar – TROG-2 Polish translation by Smoczyńska, 2008, unpublished). The theory of mind development was evinced by the indices on the Test of Reflection on Thinking (TRM, Białecka-Pikul, 2012): appropriateness of

reflection and the intensity of reflection. As to the children's language performance, there was a significant difference between the groups on the receptive vocabulary size in Polish, both when considering their raw score and the percentile score. In the raw score, the difference was significant, $t(145) = -14.72, p < 0.001, d = 0.775$ (medium effect size), with bilinguals scoring lower ($M = 62, SD = 12$) than monolinguals ($M = 71, SD = 11$). The same pattern was visible when comparing the children's percentile scores: $t(145) = -4.11, p < 0.001, d = 0.68$ (medium effect size), with bilinguals scoring lower ($M = 28^{\text{th}}$ percentile, $SD = 24$) than monolinguals ($M = 46^{\text{th}}$ percentile, $SD = 29$). This also means that the bilinguals in the sample showed worse receptive vocabulary skills than Polish monolinguals studied in the normalization: only 28% of Polish monolingual children (studied in the normalizations) scored lower than the bilinguals in the present sample. The difference on the test of receptive grammar in Polish was on the verge of significance, $t(128) = -1.95, p = 0.05$, with bilinguals scoring lower ($M = 61, SD = 10$), than monolinguals ($M = 64, SD = 11$). The two groups also differed on the expressive vocabulary test in Polish, $t(145) = -10.62, p < 0.001, d = 1.734$ (large effect size), with bilinguals scoring lower ($M = 34, SD = 8$) than monolinguals ($M = 45, SD = 4$). There was also a significant difference between the groups on both indices of the Reflection on Thinking Test: the appropriateness of reflection, $t(146) = 2.73, p = 0.007, d = 0.446$ (small effect size), and the intensity of reflection, $t(146) = 3.58, p = 0.000, d = 0.584$ (medium effect size). However, this time bilinguals obtained significantly higher scores both in the appropriateness of reflection ($M_{bilinguals} = 4.97, SD_{bilinguals} = 1.81; M_{monolinguals} = 4.12, SD_{monolinguals} = 2.02$) and the intensity of reflection ($M_{bilinguals} = 7.53, SD_{bilinguals} = 2.93; M_{monolinguals} = 5.75, SD_{monolinguals} = 3.18$). Thus, bilinguals showed lower language skills (receptive and expressive vocabulary size, receptive grammar knowledge) than the Polish monolinguals, but outperformed their monolingual peers on the test measuring theory of mind. Of interest was also whether the two groups differed in the general verbosity when telling a narrative, as measured by the total

number of words (tokens) produced when telling a story in Polish. The results from the *t*-test revealed no group differences: $t(143) = 1.27, p = 0.21$. Thus, bilinguals and monolinguals produced a similar total number of words when telling stories in Polish. Additionally, stories told by both groups were similarly saturated with ISL: internal state terms constituted on average 3% of all tokens in both bilingual and monolingual Polish stories (see Table 3).

Table 3.

The ISL saturation of the Polish stories told by bilinguals and monolinguals.

	Bilinguals (M ± SD)	Monolinguals (M ± SD)
Total Number of Words (TNW) in a story	66.3 ± 26.6	60.3 ± 29.3
All IST tokens	2.03 ± 1.51	1.84 ± 1.85
Mean % of IST / TNW	3%	3%

5.1.2. Bilingual comparison in Polish and English

In order to explore their language dominance, the bilinguals' receptive vocabulary and receptive grammar knowledge was compared across the two languages. The receptive vocabulary skills were measured by the BPVS-3 in English, and the OTSR in Polish. The receptive grammar skills were measured by the TROG in English (Bishop, 2003) and in Polish (translation by Smoczyńska, 2008, unpublished). As mentioned above, the bilinguals' results on the expressive vocabulary tests could not be compared as the two tests differed in the score scale, and were not normalized on relevant populations, hence the raw scores could not be converted into percentiles. The bilingual children's results on the receptive vocabulary tests were analysed in percentiles because the two tests differed in the score scale. Bilinguals' scores were on the receptive vocabulary tests on average in the 29th percentile in English and in the 28th percentile in Polish. The difference was statistically non-significant, $t(42) = -0.39, p = 0.7$.

This shows that the bilinguals' receptive skills were similar across the two languages, but it also confirms that bilingual children performed worse than their monolingual peers, i.e. only 29% and 28% of respective monolingual children (studied in the normalizations) scored lower than the bilinguals in the present sample. The bilinguals scored comparably on the receptive grammar knowledge test (TROG), $t(42) = -1.81$, $p = 0.08$, in Polish: $M = 60$, $SD = 10$, in English: $M = 59$, $SD = 10$. Since the English TROG-2 was standardized on English monolinguals aged 4 to 16 years, the bilinguals' scores were translated into percentiles. Bilinguals scored on average in the 55th percentile ($SD = 30$, range: 0 - 99), which shows they performed normally and comparably to English monolingual children in the standardization sample. However, it is important to note that the bilinguals' individual scores varied widely and ranged between 0 and 99th percentile. To conclude, bilingual children exhibited similar receptive performance in their two languages, though in most tasks, they performed below their monolingual peers.

Of interest was also whether bilinguals differed in the general verbosity when telling a narrative in Polish and English, as measured by the overall amount of tokens (total number of words) produced when telling a story. The results from a paired sample t -test comparing the means of the number of tokens in Polish and English revealed a difference that was on the verge of significance: $t(74) = -1.98$, $p = 0.05$, $d = -0.285$ (small effect size), with the English narratives containing slightly more tokens ($M = 76$, $SD = 40$) than the Polish stories ($M = 66$, $SD = 27$). However, the bilinguals' stories in both languages were similarly saturated with internal state terms: these constituted on average 3% of all tokens produced in the Polish and English stories (see Table 4).

Table 4.

The ISL saturation of the Polish and English stories told by bilinguals.

	Polish (M ± SD)	English (M ± SD)
Total Number of Words (TNW) in a story	66.3 ± 26.6	75.9 ± 39.85
All IST tokens	2.03 ± 1.51	1.95 ± 1.67
Mean % of IST / TNW	3%	3%

To recapitulate, the bilinguals studied here showed lower linguistic skills in Polish, compared to their monolingual peers. They attained significantly lower scores on all three language tests: that of receptive vocabulary, expressive vocabulary and receptive grammar. However, bilinguals outperformed their monolingual peers on the theory of mind test: they showed better on the scores of the appropriateness of reflection and the intensity of reflection. Moreover, bilinguals seemed to be quite balanced across their two languages: they scored comparably on the receptive vocabulary and the receptive grammar tests (no comparison on expressive vocabulary was possible). Last, bilinguals and monolinguals did not differ in their general verbosity when telling stories in Polish (as measured by the total number of words, TNW). There was a difference on the verge of significance between the TNW in the Polish and English narratives of the bilinguals, with the English stories containing slightly more tokens than the Polish stories. However, this difference might be attributed to the general cross-linguistic differences between Polish and English, rather than to the bilingual's proficiency in the two languages, which seems to be comparable.

5.2. Main analyses related to ISL

Let us now turn to the presentation of the main analyses related to the ISL in the bilingual children. In the present section, the bilinguals' use of ISL in a narrative will be compared to that of monolingual peers, and across the bilinguals' two languages. Next, we will search for

the best predictors of the use of ISL in the Polish and English narratives. Last two Subsections will present the results of analyses of the narrative task: the effect of retelling after a model on the inclusion of ISL, and the comparison of ISL in a told story and in an interview focused on the protagonists' internal states.

The values of the dependent variables, the IST number and the IST proportion were generally low: the distributions of the variables were positively skewed, i.e. there were many scores on the left of the distribution (near zero). The skewness ranged from 0.5 to 2.6 for the particular IST subclasses. Thus, the data did not follow the Gaussian distribution. Therefore, the main analyses were performed with the use of non-parametric tests. Whenever group comparisons were made between the bilinguals and the monolinguals, a Mann-Whitney U Test was used. Mann-Whitney U Test is an independent samples rank test for the difference between two population medians, instead of means (used by a parametric t-test). Accordingly, when language comparisons were made (e.g. between two languages of the bilinguals, or between told and retold stories), a Wilcoxon Signed Rank Test was used. Wilcoxon Signed Rank Test is a non-parametric rank test used to compare mean ranks in related (paired) samples. Another non-parametric test used in the present analysis was the Spearman's rho test (also known as Spearman rank correlation) used to investigate the relation between the IST amount in the Polish and English stories of the bilinguals and the possible children's age, vocabulary size and receptive grammar knowledge in each language, theory of mind development, the story structure, and the total number of words produced in a story. Finally, to explore the effects of group, language, mode (Telling, Retelling), and context (Telling vs. Questions) on the use of IST, a series of non-parametric two-way ANOVAs for trimmed means was conducted. A trimmed mean, similarly to a median, is a statistical measure of central tendency and considered a useful estimator in non-parametric tests because it is less sensitive to outliers than the mean but yields a reasonable estimate of central tendency for many statistical models. The test that

was performed with the use of “t2way” function in R from WRS2 package, a collection of robust statistical methods based on Wilcoxon's WRS functions (for more information on the package, see: Mair & Schoenbrodt, 2017; Mair & Wilcox, 2016; for comparison of different methods for non-parametric two-way ANOVAs see Feys, 2016). “T2way” calculates a trimmed mean by discarding 20% of the data points from the high end and 20% of the data points from the low end of the probability distribution. Post hoc comparisons for the trimmed-mean analyses were done with a corresponding “mcp2atm” function from the WRS2 package (Mair & Schoenbrodt, 2017).

5.2.1. RQ1: Do bilingual and monolingual children differ in the amount of internal state terms used when telling a story in Polish?

It was hypothesized that the bilingual and monolingual children do not differ in the use of internal state terms when telling a story in Polish. Most children from both groups did use at least one IST in their narratives. The exact number of children who used at least one internal state term in their narratives can be found in Table 5. The chi-square test showed there was no significant difference between the groups (bilingual, monolingual) on the distribution of children who used no IST or at least one IST, $\chi^2 = 2.54, p = 0.11$. This means that the use or no use of IST is not associated with the group status. The descriptive statistics for the amount of internal state terms across the bilingual and monolingual groups can be seen in Table 6 (number of IST) and Table 7 (proportion of IST in the total number of words in a story). A Mann-Whitney U Test was conducted to compare the number of all internal state terms (tokens) used in Polish by the bilingual and monolingual group, and revealed no statistically significant difference between the groups, $U = 3143.5, p = 0.2$. Another Mann-Whitney U Test was performed on the proportion of all IST. Again, there were no significant differences between the groups, $U = 3119, p = 0.25$. Thus, the bilingual and monolingual stories contained a similar number of IST and were also similarly saturated with internal state terms. All internal state

terms (tokens) constituted on average 3% of the total number of words in both the bilinguals' and monolinguals' stories (see Table 7). Finally, a non-parametric two-way ANOVA was run to compare the effect of group (bilingual, monolingual) and the effect of internal state term subclass (emotional, mental, perceptual) on the proportion of IST in children's Polish told stories. The results revealed no significant effect of group, Test statistic = 1.01, $p = 0.32$. However, there was a significant main effect of IST subclass, Test statistic = 25.33, $p = 0.001$, with all children (bilingual and monolingual) using a similar number of mental and perceptual terms ($p = 0.56$) but significantly fewer emotional than mental terms ($p = 0.003$), and significantly fewer emotional than perceptual terms ($p = 0.000$). The effect of the interaction between group and IST subclass was non-significant, Test statistic = 2.17, $p = 0.34$.

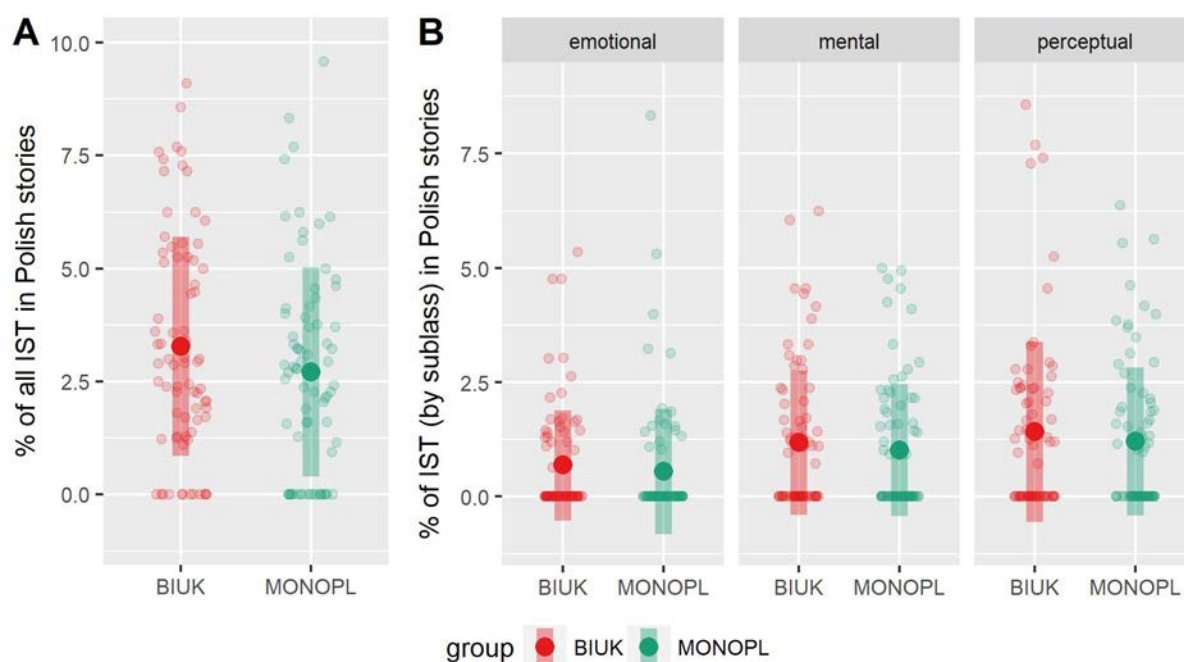


Figure 1. The IST proportion (IST tokens/TNW, in %) in bilingual and monolingual children's told stories: (A) all internal state terms, (B) internal state terms by subclass. The small dots represent the individual children, the solid dots represent the means, the lineranges represent the means \pm 1SD.

Table 5.

The number of children who used no internal state terms in their told stories (IST = 0) and the number of children who used at least one internal state term (IST > 0).

	Bilinguals (No. of children)	Monolinguals (No. of children)
IST = 0	12	20
IST > 0	63	55

Table 6.

The descriptive statistics for the number of internal state terms (tokens) used in Polish stories of bilingual and monolingual children: medians, means and standard deviations for the number of the IST in each group (bilingual, monolingual) and each subclass (emotional, mental, perceptual).

	Bilinguals		Monolinguals	
	Median	(M ± SD)	Median	(M ± SD)
All IST (number)	2	2.03 ± 1.51	1	1.84 ± 1.85
Emotional IST (number)	0	0.47 ± 0.76	0	0.29 ± 0.71
Mental IST (number)	0	0.75 ± 0.97	0	0.76 ± 1.13
Perceptual IST (number)	1	0.81 ± 0.97	0	0.79 ± 1.08

Table 7.

The descriptive statistics for the IST proportion (IST/TNW, in %) in Polish stories of bilingual and monolingual children: medians, means and standard deviations for each group (bilingual, monolingual) and each subclass (emotional, mental, perceptual).

	Bilinguals		Monolinguals	
	Median	(M ± SD)	Median	(M ± SD)
All IST (%)	3	3.28 ± 2.43	3	2.90 ± 2.83
Emotional IST (%)	0	0.68 ± 1.2	0	1.56 ± 1.36
Mental IST (%)	0	1.18 ± 1.58	0	1.15 ± 1.82
Perceptual IST (%)	1	1.42 ± 1.96	0	1.2 ± 1.62

Additionally, it was checked which of the terms were most common in the Polish narratives of bilinguals and monolinguals (see Table 8). The groups used similar terms. The emotional terms that were used most by the children were “wystraszony” (scared), “szczęśliwy” (happy), and “smutny” (sad). The most frequently used mental terms were “chcieć” (want), “próbować” (try), and “móc” (can/be able to). The most frequently used perception terms were “zobaczyć” (see), “widzieć” (see) and “patrzeć” (“look”). Also, a few monolinguals used the term “zauważyć” (notice), while almost none of the bilinguals did (see Appendix C for the full list of terms).

The terms used by many of the children were: “chcieć” (want) – used by 43% of bilinguals and 40% of monolinguals, and “zobaczyć” (see Table 8) – used by 25% of bilinguals and 23% of monolinguals. A closer inspection of the narrative content suggested that children may use both cognitive and perception terms to refer to the other’s state of knowledge, as evinced by the following examples:

An excerpt from a narrative told by a monolingual boy, aged 6;10:

- *CHI: Była=Był@ sobie mama ptak.
[There was a mommy bird]
- *CHI: I urodziła dwa...
[And she gave birth to two...]
- *CHI: Yy, dwa...
[Yyy, two...]
- *CHI: Odleciała ich nakarmić.
[She flew away to feed them.]
- *CHI: A kot ich **zobaczył**.
[A cat **saw** them.]
- *CHI: **Chciał** je zjeść.
[It **wanted** to eat them.]
- *CHI: A wtedy mama przyleciała.
[Then mommy bird came.]

A bilingual girl, aged 5;5:

- *CHI: A potem jakaś wielka kózka **zobaczyła**, że on tam wpadł.
[And then some big goat **saw** that he fell in.]
- *CHI: Sobie skoczyła=skoszyła i wypchnęła=wypchnęł@ go już=jus głową.

[She jumped and pushed him out with her head.]

*CHI: I potem jakiś=jakis lis **podglądał**.

[And then some fox **watched** them.]

Moreover, the excerpts also show that children at pre- and early school age may focus more on those internal states that allow them to explain the protagonist's actions, and give less attention to referring the protagonists' emotions. This would confirm the results from the two-way ANOVA showing that both groups used significantly more mental and perceptual terms than emotional terms.

Table 8.

The most frequent terms used by the bilingual and monolingual children in the told stories. The values indicate the number and the percentage of children (in the bilingual and monolingual groups) that used the specific terms at least once. The table includes three most frequent terms in each subclass. The full list of terms used by the children is given in Appendix C. The terms in the table are given in the order of decreasing number of occurrences in Polish (first column).

Most frequent terms in the Polish told stories (telling mode)							
			No of occurrences (tokens) in all children	Bilinguals (No of children who used the term)	Bilinguals (% of children who used the term)	Monolinguals (No of children who used the term)	Monolinguals (% of children who used the term)
emotional	wystarszony	(scared)	20	12	16%	6	8%
	szczęśliwy	(happy)	10	7	9%	2	3%
	smutny	(sad)	5	3	4%	2	3%
mental	chcieć	(want)	95	32	43%	30	40%
	móc	(able to)	7	6	8%	1	1%
	próbować	(try)	7	3	4%	4	5%
perception	zobaczyć	(see)	49	19	25%	17	23%
	widzieć	(see)	25	12	16%	8	11%
	patrzeć	(look)	22	9	12%	9	12%

5.2.2. RQ2: Do bilingual children differ in the amount of internal state terms used when telling the stories in Polish and English?

It was hypothesized that bilingual children do not differ in the use of internal state terms when telling the stories in Polish and English. The number of children who used at least one internal state term in their narratives can be found in Table 9. The McNemar test for dependent samples showed that the proportion of bilinguals who used no IST in their Polish stories was different from the proportion in the English stories, $\chi^2 = 39.58$, $p = 0.0$. This means that the children who did not use any IST in their Polish stories were not necessarily the same children that did not use any IST in English (see Table 9). However, it must be noted that the children generally used few IST: on average only 2 internal state terms per story (see Table 10 and Table 11). Thus, the difference between using zero or one or two internal state terms could be mediated by factors other than the language being spoken (e.g. child's focus on the task). The descriptive statistics for the amount of internal state terms across languages can be seen in Table 10 (number of IST) and Table 11 (proportion of IST in the total number of words in a story).

First, a Wilcoxon Signed Rank Test was conducted to compare the number of all internal state terms used in Polish and English, and revealed no statistically significant difference between the languages, $V = 765.5$, $p = 0.66$. Then, another Wilcoxon Signed Rank Test was performed on the proportion of all IST. Again, there were no significant differences between the languages, $V = 1584$, $p = 0.13$. Thus, the bilingual stories contained a similar number of IST and were similarly saturated with internal state terms in both languages. All internal state terms (tokens) constituted on average 3% of the total number of words in both the bilinguals' and monolinguals' stories (see Table 11). Next, a non-parametric two-way ANOVA was run to compare the effect of language (Polish, English) and the effect of internal state term subclass (emotional, mental, perceptual) on the number of internal state terms (tokens) in children's told stories. The results revealed no significant effect of language, Test statistic = 0.78, $p = 0.38$.

However, there was a significant main effect of the IST subclass, Test statistic = 23.6, $p = 0.001$, with mental and perceptual terms being significantly more frequent than emotional terms, regardless of language used for storytelling (emotional vs. mental, $p = 0.003$; emotional vs. perceptual, $p = 0.001$). The effect of the interaction between group and IST subclass was insignificant, Test statistic = 0.17, $p = 0.92$.

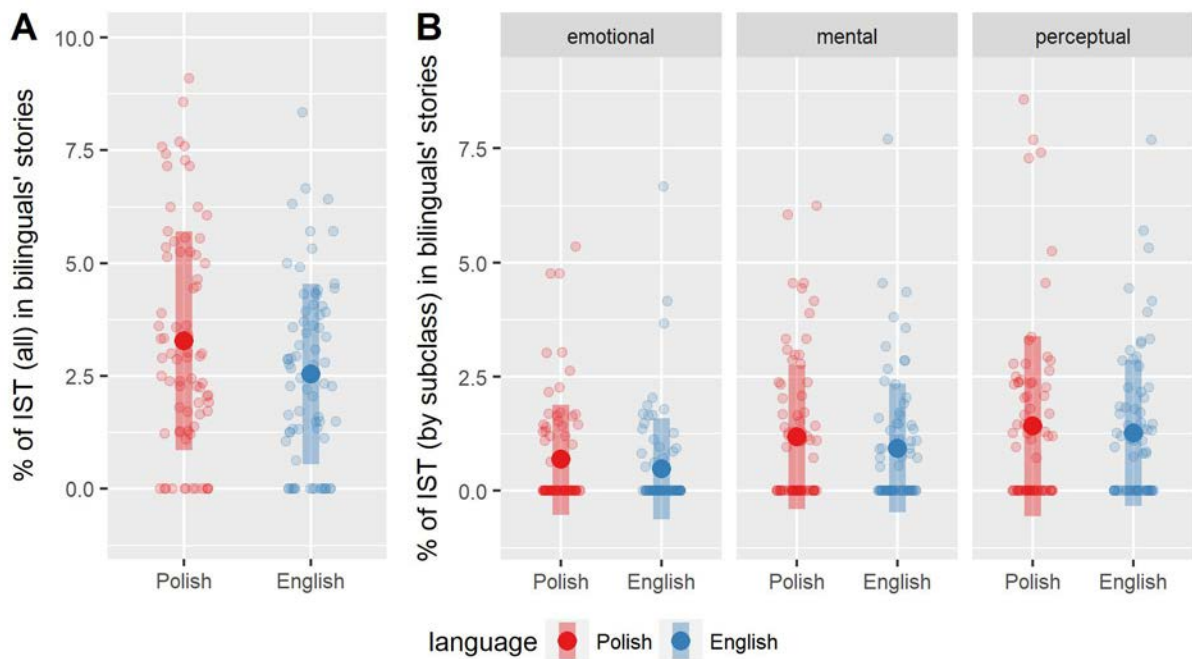


Figure 2. The IST proportion (IST tokens/TNW, in %) in stories told by bilinguals in Polish and in English: (A) all internal state terms, (B) internal state terms by subclass. The small dots represent the individual children, the solid dots represent the means, the lineranges represent the means $\pm 1SD$.

Table 9.

The number of bilingual children who used no internal state terms in their told stories (IST = 0) and the number of children who used at least one internal state term (IST > 0). The first column shows the number of children who used/did not use IST in Polish stories only, the second column shows the number of children who used/did not use IST in English, and the last column shows the number of children who used/did not use IST in Polish and English stories.

	Polish (No. of children)	English (No. of children)	Polish & English (No. of children)
IST = 0	12	16	3
IST > 0	63	59	72

Table 10.

The descriptive statistics for the number of internal state terms (tokens) used in Polish and English stories of bilingual children: medians, means and standard deviations for the number of the IST in each language and each subclass.

	Polish		English	
	Median	(M ± SD)	Median	(M ± SD)
All IST (number)	2	2.03 ± 1.5	2	1.95 ± 1.67
Emotional IST (number)	0	0.47 ± 0.76	0	0.39 ± 0.78
Mental IST (number)	0	0.75 ± 0.97	0	0.6 ± 0.7
Perceptual IST (number)	1	0.81 ± 0.97	1	0.96 ± 1.09

Table 11.

The descriptive statistics for the IST proportion (IST/TNW, in %) in Polish and English stories of bilingual children: medians, means and standard deviations for each language and each subclass.

	Polish		English	
	Median	(M ± SD)	Median	(M ± SD)
All IST (%)	3	3.28 ± 2.43	3	3.4 ± 5.51
Emotional IST (%)	0	0.68 ± 1.2	0	0.63 ± 1.96
Mental IST (%)	0	1.18 ± 1.58	0	1.23 ± 2.52
Perceptual IST (%)	1	1.42 ± 1.96	1	1.55 ± 2.9

Then, a cross-language correlation was carried out for the bilingual population in order to investigate whether the use of internal state terms in both languages is interrelated. Since all the values were positively skewed and did not follow the Gaussian distribution, the non-parametric Spearman's rank correlations were used to investigate the relation between the use of internal state terms in the Polish and English stories of the bilinguals. The results revealed no significant cross-language correlations in the overall IST proportion in the told stories: $\rho = -0.08$, $p = 0.52$. A closer look at the particular ISL subclasses revealed no significant cross-language correlations for any of the subclasses: emotional terms, $\rho = 0.07$, $p = 0.54$, mental terms, $\rho = 0.12$, $p = 0.29$, perceptual terms, $\rho = 0.09$, $p = 0.45$.

Additionally, it was checked which of the terms were most common in the Polish and English narratives (see Table 12). The most frequently emotional terms that were used by the bilinguals in the Polish stories were also the most frequently used terms used in the English stories: "wystraszony" (scared), "szczęśliwy" (happy), and "smutny" (sad). The most frequently used mental terms were "chcieć" (want), "próbować" (try), and "móc" (can/be able to). The perception terms that were used most by the children were "zobaczyć" (see), and "patrzeć" (look).

A closer look at the stories of particular children seemed to confirm a general similarity in the use of ISL

A bilingual boy, aged 5;6, his English narrative:

- *CHI: One little goat did fell in the water.
- *CHI: And yy the the big, the big goat did push him on the grass.
- *CHI: And then a fox came and, and he **wanted** to to eat the little baby amm goat.
- *CHI: And he wanted to eat the little baby goat.
- *CHI: And he chased him.
- *CHI: And and and the and the baby goat didn't **see** him.
- *CHI: And and, is this a wolf?
- *CHI: No, it's a fox.

- *CHI: And the fox did, did, did@ catch him but, but he didn't eat him.
- *CHI: Because, because the bird went on his way and, and he and he and he did, rescued the goat.
- *CHI: And then, and then he went back, back to his family.
- *CHI: And the fox.
- *CHI: And the bird was chasing the fox.

The same boy, his Polish narrative:

- *CHI: Ptaszki siedziały w gniazdku i śpiewały.
[Birds were sitting in a nest and singing.]
- *CHI: Później jeden odleciał a później przyleciał.
[Later, one of them flew away and later came back.]
- *CHI: A jak wcześniej odleciał to kot tam był i **chciał** wejść na drzewo.
[But before, when he flew away, a cat was there and **wanted** to climb the tree.]
- *CHI: No i jeden ptaszek odleciał i przyleciał.
[And then one bird flew away and came back.]
- *CHI: I kot wszedł... i zaczął wbiegać na drzewo.
[And the cat went... started running up the tree.]
- *CHI: A później pies **zobaczył** kota i **chciał** go zjeść, czyli też...
[And then a dog **saw** the cat and **wanted** to eat him, so he also...]
- *CHI: Czyli też pobiegł za kotkiem.
[So he also run after the cat.]
- *CHI: A później na trawie się gonili.
[And they were chasing each other on the grass.]
- *EXP: Aha, dobrze.
[Aha, good.]
- *CHI: I skończone.
[And done.]

Table 12.

The most frequent terms used by the bilingual children in their Polish and English told stories. The values indicate the number and the percentage of children that used the specific terms at least once. The table includes three most frequent terms in each subclass. The full list of terms used by the children is given in Appendix C. The terms in the column are given in the order of decreasing number of occurrences in Polish (first column).

		No of occurrences (tokens) in Polish	Bilinguals in Polish (no of children who used the term)	Bilinguals in Polish (% of children who used the term)		No of occurrences (tokens) in English	Bilinguals in English (no of children who used the term)	Bilinguals in English (% of children who used the term)
emotional	wystarszony	13	12	16%	scared	10	8	11%
	szczęśliwy	7	7	9%	happy	7	6	8%
	smutny	3	3	4%	sad	2	2	3%
mental	chcieć	45	32	43%	want	19	18	24%
	móc	6	6	8%	able to	4	5	7%
	próbować	3	3	4%	try	18	15	20%
perception	zobaczyć	25	19	25%	see	48	30	40%
	widzieć	17	12	16%				
	patrzeć	11	9	12%	look	14	10	13%
	zauważyć	2	2	3%	notice	0	0	0%

5.2.3. RQ3: What are the best predictors of the use of internal state terms in the narratives told in Polish and English?

It was hypothesized that there is a positive relationship between the use of internal state terms in a narrative and the children's age, vocabulary size and receptive grammar knowledge in each language, theory of mind development, the story structure, and the total number of words produced in a story. The hypothesis did not presuppose which of these factors will constitute the best model predicting the number of internal state terms in the narratives told in Polish and in English.

First, a series of non-parametric Spearman's rank correlations were carried out in order to investigate how the number the internal state terms in the told stories was related to age, receptive and expressive vocabulary score in each language, receptive grammar knowledge in each language, performance on ToM test, the story structure score, and the total number of words produced in a story. The chosen tests were non-parametric due to a significant skewness of the number of internal state terms. The correlational analyses revealed which of the factors were significantly related to the number of internal state terms produced in the stories. Also, a correlational matrix was created in order to observe if any of the factors are cross-related to each other. Finally, two regression analyses were performed in order to identify the model that best predicted the use of internal state terms in the stories told in each language.

Age. The Spearman's rank correlation between age (in months) and the number of all internal state terms (tokens) in Polish told stories was non-significant either for the monolinguals, $\rho = 0.1, p = 0.41$, or bilinguals, $\rho = 0.13, p = 0.27$. However, the relationship was positive and significant for age and the number of all internal state terms (tokens) in bilingual stories told in English, $\rho = 0.38, p = 0.001$ (see Figure 3.A). When looking closely at the correlations between age and the specific subclasses of English internal state terms, a significant positive correlation was found only for age and the number of English perceptual

terms, $\rho = 0.41$, $p = 0.000$. The remaining subclasses of the English internal terms did not correlate significantly with age, English emotional terms: $\rho = 0.09$, $p = 0.45$, English mental terms: $\rho = 0.11$, $p = 0.37$ (see Figure 3.B).

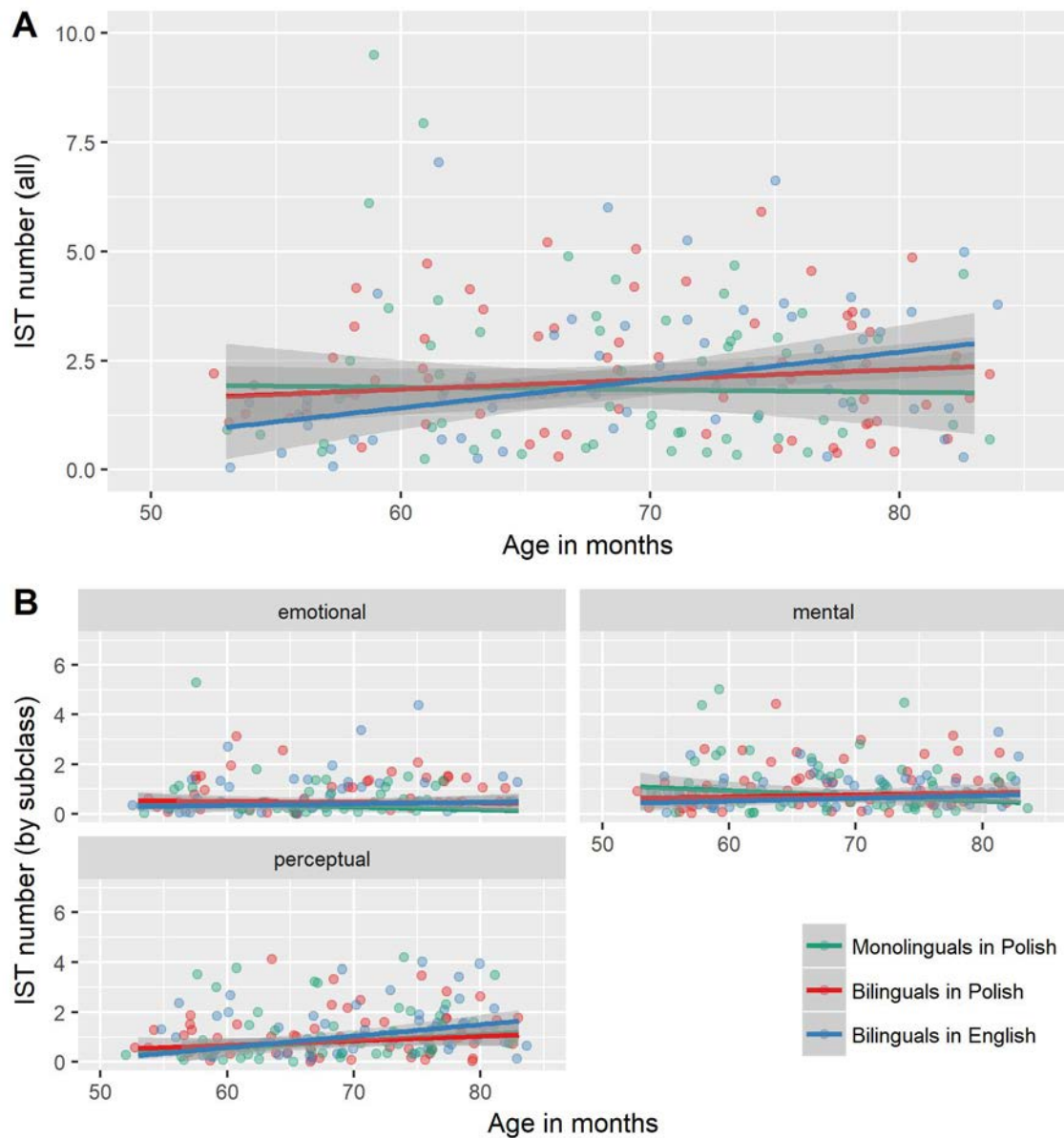


Figure 3. The number of (A) all internal state terms (tokens) and (B) IST tokens by subclass in the Polish and English told stories of bilinguals and monolinguals plotted against the children's age (in months).

Receptive vocabulary size. There was no statistically significant correlation between the Polish receptive vocabulary scores and the number of all internal state terms in the Polish stories, either in monolingual children, $\rho = 0.14$, $p = 0.23$, or in bilingual children, $\rho = 0.19$, $p = 0.11$. However, the relation between the bilinguals' English receptive vocabulary scores and the number of all internal state terms (tokens) in the English stories was positive and significant, $\rho = 0.26$, $p = 0.03$ (see Figure 4.A). A closer look at the particular subclasses of ISL in the bilinguals' stories revealed a positive significant correlation only between the number of English perceptual terms and English receptive vocabulary, $\rho = 0.29$, $p = 0.02$, but the remaining subclasses of internal state terms showed a non-significant correlation with English receptive vocabulary: English emotional terms, $\rho = 0.14$, $p = 0.26$, mental terms, $\rho = 0.02$, $p = 0.84$. The correlations are pictured in Figure 4.B.

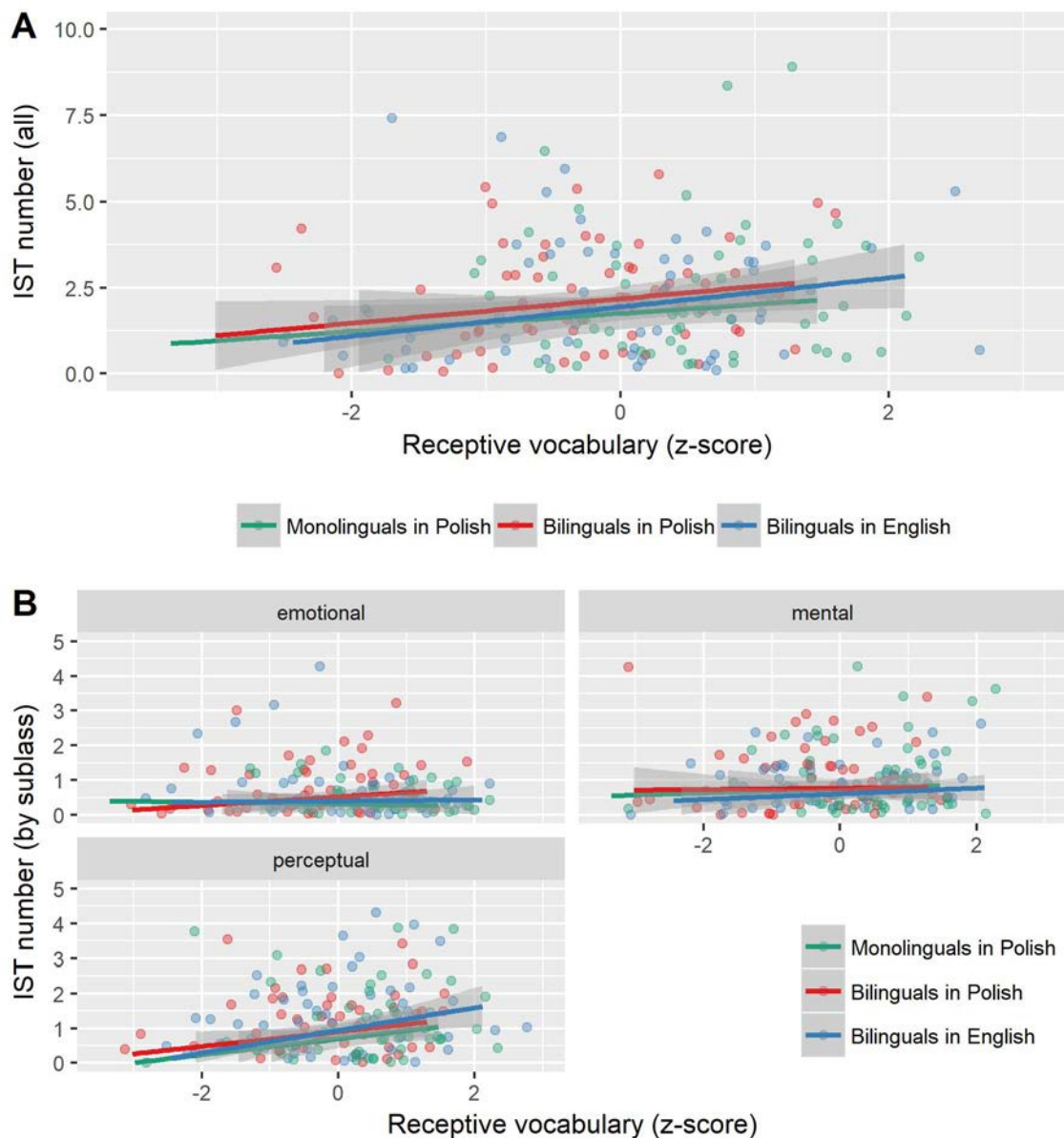


Figure 4. The number of (A) all internal state terms (tokens) and (B) tokens by subclass in the Polish and English told stories plotted against receptive vocabulary scores (z-scores) in Polish and English. The z-scores were calculated solely for the graphical purposes, to show two languages on a single graph. The mean score and the standard deviations for the z-scores were calculated on the merged populations (monolingual and bilingual Polish for the Polish vocabulary tests, and bilingual English for the English vocabulary tests).

Expressive vocabulary size. The correlations between the expressive vocabulary scores and the number of internal state terms were non-significant either for the monolinguals, $\rho = 0.08$, $p = 0.52$, or for the bilinguals, either in Polish, $\rho = 0.02$, $p = 0.9$, or in English, $\rho = 0.17$, $p = 0.14$ (see Figure 5). Since no significant correlations were found between

expressive vocabulary size and all the internal state terms taken together, no further correlations were performed on the particular subclasses of internal state terms.

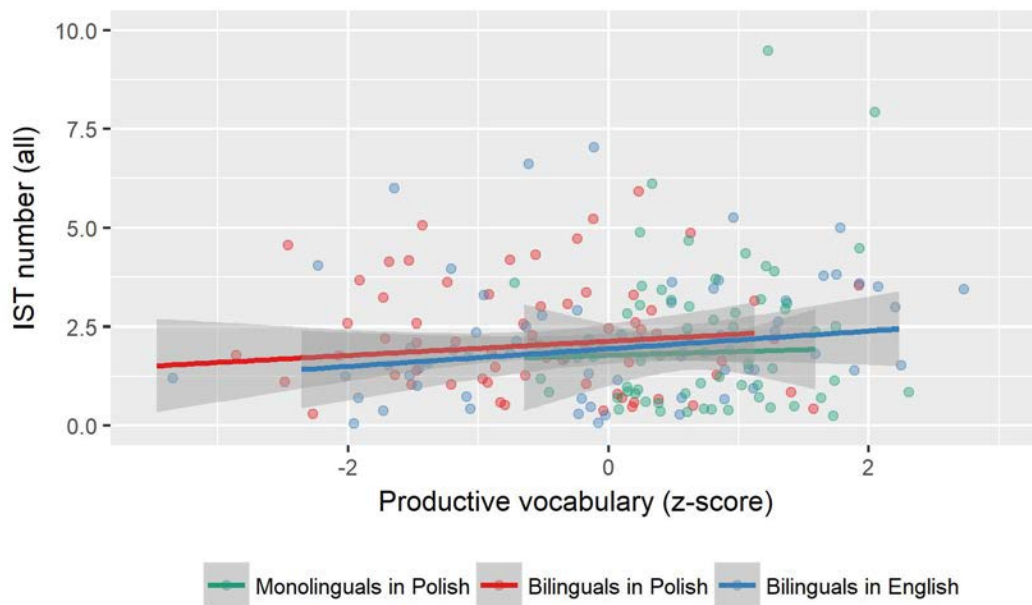


Figure 5. The number of internal state terms (tokens) in the Polish and English told stories plotted against productive vocabulary scores (z-scores). The z-scores were calculated solely for the graphical purposes, to show two languages on a single graph. The mean score and the standard deviations for the z-scores were calculated on the merged populations (monolingual and bilingual Polish for the Polish vocabulary tests, and bilingual English for the English vocabulary tests).

Receptive grammar knowledge. The correlations between the receptive grammar knowledge and the overall number of internal state terms were non-significant for the monolinguals, $\rho = 0.11$, $p = 0.36$, but were significant for the bilinguals in Polish, $\rho = 0.29$, $p = 0.02$. A closer look at the particular subclasses of ISL revealed significant correlations only between Polish grammar skills and the number of Polish perceptual terms (tokens), $\rho = 0.31$, $p = 0.02$. There were no significant correlations between Polish grammar skills and either mental terms, $\rho = 0.19$, $p = 0.14$, or emotional terms $\rho = 0.12$, $p = 0.37$. In English, the correlation between English grammar skills and the number of all English ISTs was non-significant, $\rho = 0.04$, $p = 0.8$ (see Figure 6).

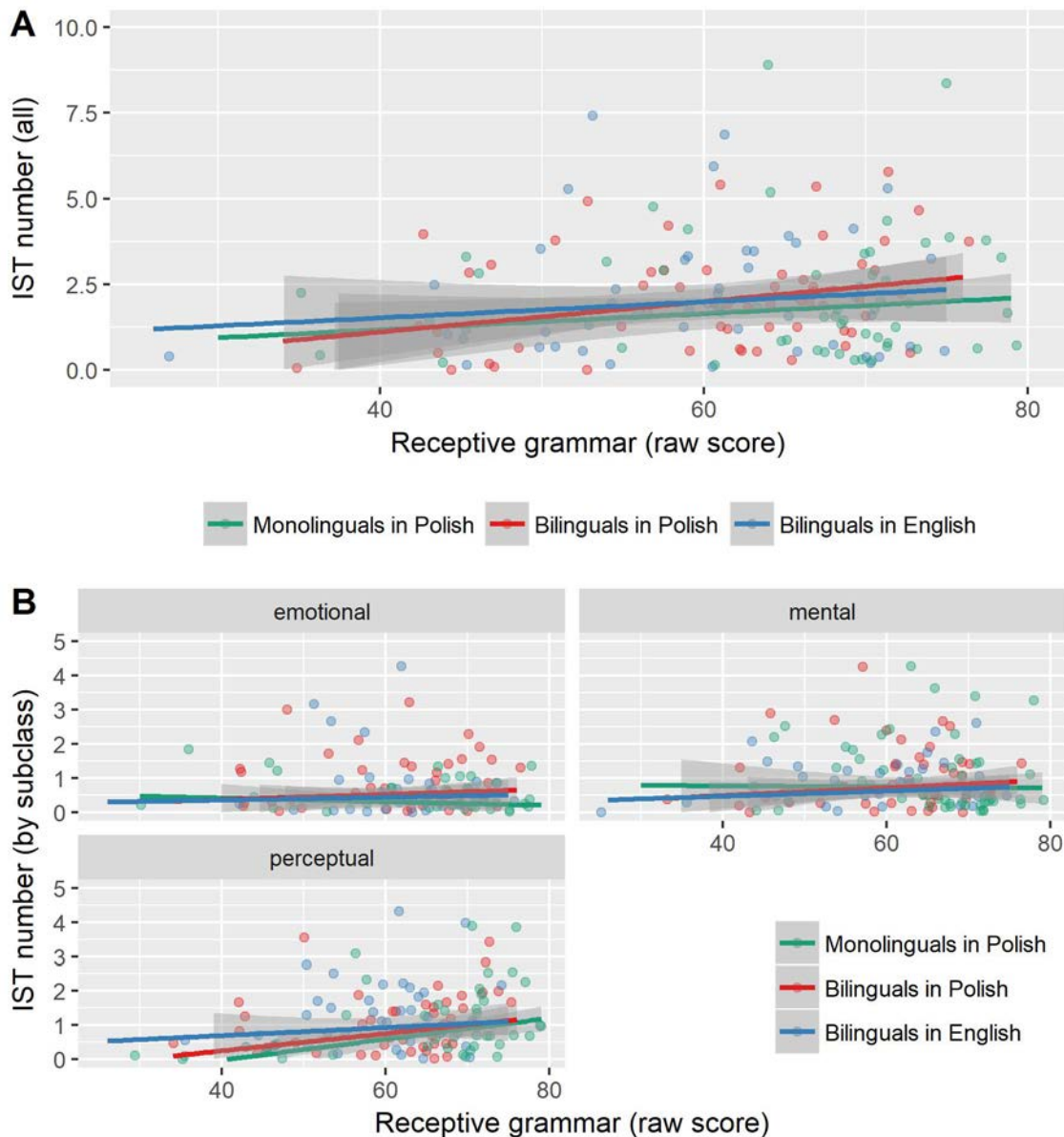


Figure 6. The number (A) all internal state terms (tokens) and (B) tokens by subclass in the Polish and English told stories plotted against receptive grammar test score (raw score).

Theory of mind. The task used to measure theory of mind performance yields two indices: appropriateness of reflection (response accuracy) and the intensity of reflection. The correlation between the appropriateness on the ToM task and the number of internal state terms was non-significant either in the monolinguals, $\rho = 0.06$, $p = 0.62$, or in the bilinguals, either in Polish, $\rho = 0.18$, $p = 0.16$, or in English, $\rho = 0.17$, $p = 0.15$ (see Figure 7.A). No further correlations between ToM appropriateness and specific subclasses of IST were performed.

However, the relationship between the intensity of reflection and the number of internal state terms showed a different pattern (see Figure 7.B). Here, the correlation with Polish internal state terms was non-significant in either monolinguals, $\rho = 0.18$, $p = 0.12$, or bilinguals, $\rho = 0.11$, $p = 0.37$, but there was a significant positive relationship between the intensity of reflection and the number of English internal state terms (all categories taken together), $\rho = 0.28$, $p = 0.02$. The investigation of the specific subclasses of internal state terms revealed a significant correlation between the intensity of reflection on the ToM task and the number of English perceptual terms, $\rho = 0.27$, $p = 0.02$. The relationship was non-significant for the English emotional terms, $\rho = 0.05$, $p = 0.66$ and English mental terms, $\rho = 0.17$, $p = 0.16$ (see Figure 8).

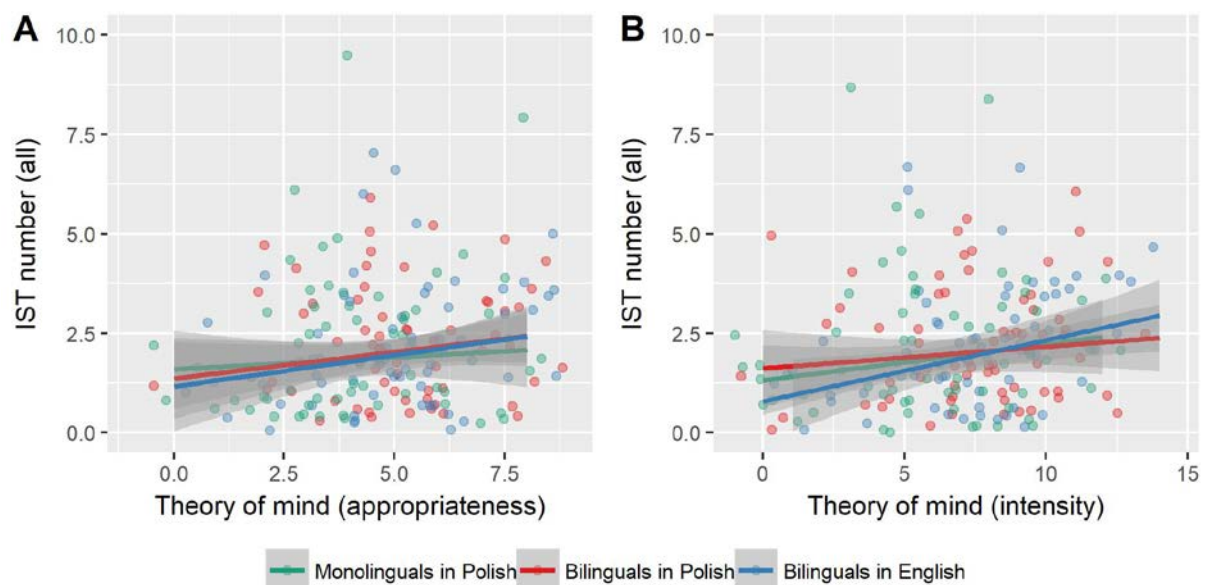


Figure 7. The number of all internal state terms (tokens) in the Polish and English told stories plotted against the score on Reflection on Thinking Task (A) appropriateness of reflection and (B) intensity of reflection.

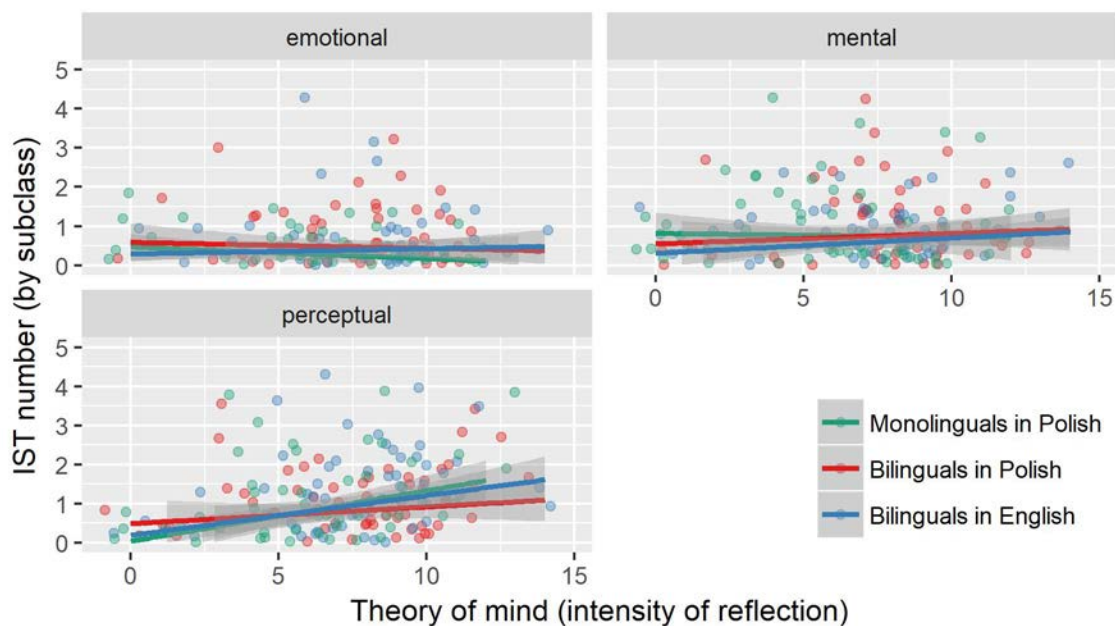


Figure 8. The number of internal state tokens by subclass in the Polish and English told stories plotted against the score on the intensity of reflection in the Reflection on Thinking Task.

Story structure. It is important to note that the story structure did not include the points usually given for the references to the internal states (Gagarina et al., 2012). The results from the Polish stories revealed positive significant correlations between story structure scores and the number of all internal state terms both for the Polish monolinguals, $\rho = 0.47$, $p = 0.000$, and the bilinguals', $\rho = 0.23$, $p = 0.05$. The English revealed no significant correlations, $\rho = 0.14$, $p = 0.22$ (see Figure 9.A). A closer look at the particular subclasses of ISL in Polish revealed positive significant correlations between the story structure score and two subclasses of IST, and that only for the Polish monolingual group: mental terms, $\rho = 0.4$, $p = 0.001$, and perceptual terms, $\rho = 0.33$, $p = 0.005$. The correlation between the monolinguals' story structure score in Polish and the number of emotional terms was non-significant, $\rho = 0.08$, $p = 0.51$. No significant correlations were found between the number of particular subclasses of IST in Polish and the story structure score in the bilingual group: emotional terms, $\rho = 0.11$, $p = 0.35$, mental terms, $\rho = 0.21$, $p = 0.07$, perceptual terms, $\rho = 0.11$, $p = 0.36$. The correlations are pictured in Figure 9.B.

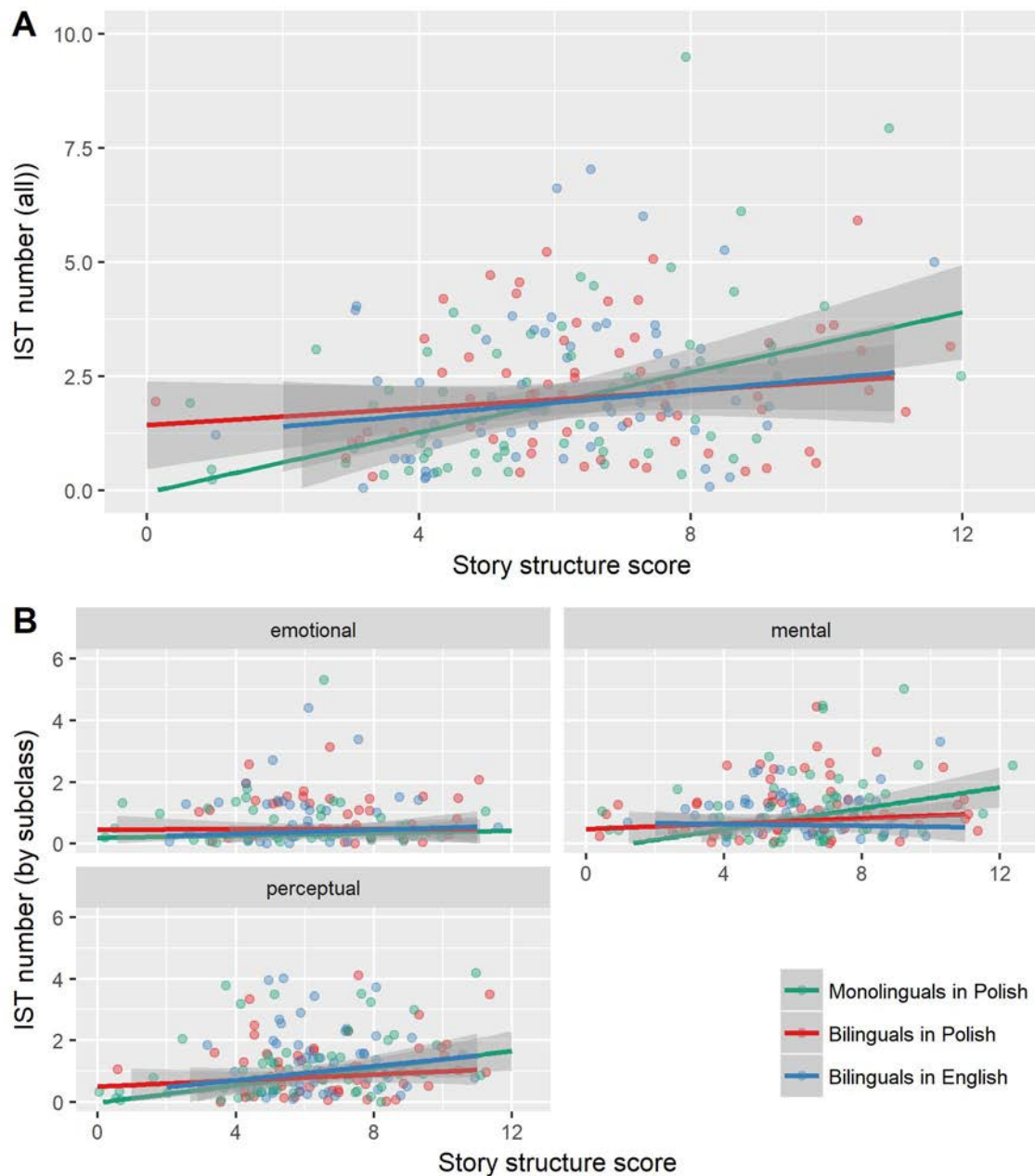


Figure 9. The number of (A) all internal state terms (tokens) and (B) tokens by subclass in the Polish and English told stories plotted against story structure score in Polish and English.

Total number of words in a story. When the overall number of all internal state terms (tokens) was considered, it correlated significantly with the total number of words (TNW) in both the Polish monolinguals, $\rho = 0.55$, $p = 0.000$, and the bilinguals, both in Polish, $\rho = 0.32$, $p = 0.005$, and in English, $\rho = 0.49$, $p = 0.000$ (see Figure 10.A). A closer look at the particular ISL subclasses revealed a different pattern for each of the two groups: the total

number of words (TNW) in the Polish stories of monolinguals correlated significantly with the number of mental, $\rho = 0.34, p = 0.000$ and perceptual terms, $\rho = 0.38, p = 0.001$ but not the emotional terms, $\rho = 0.08, p = 0.53$. On the other hand, TNW in the Polish stories of bilinguals did not correlate significantly with neither of the ISL subclasses: emotional terms, $\rho = 0.19, p = 0.09$, mental terms, $\rho = 0.17, p = 0.15$, perceptual terms, $\rho = 0.21, p = 0.08$. TNW in the English stories of bilinguals correlated significantly with the number of emotional terms, $r(73) = 0.47, p = 0.0001$, and perceptual terms, $\rho = 0.42, p = 0.000$, but not the mental terms, $\rho = 0.17, p = 0.14$. The correlations are pictured in Figure 10.B.

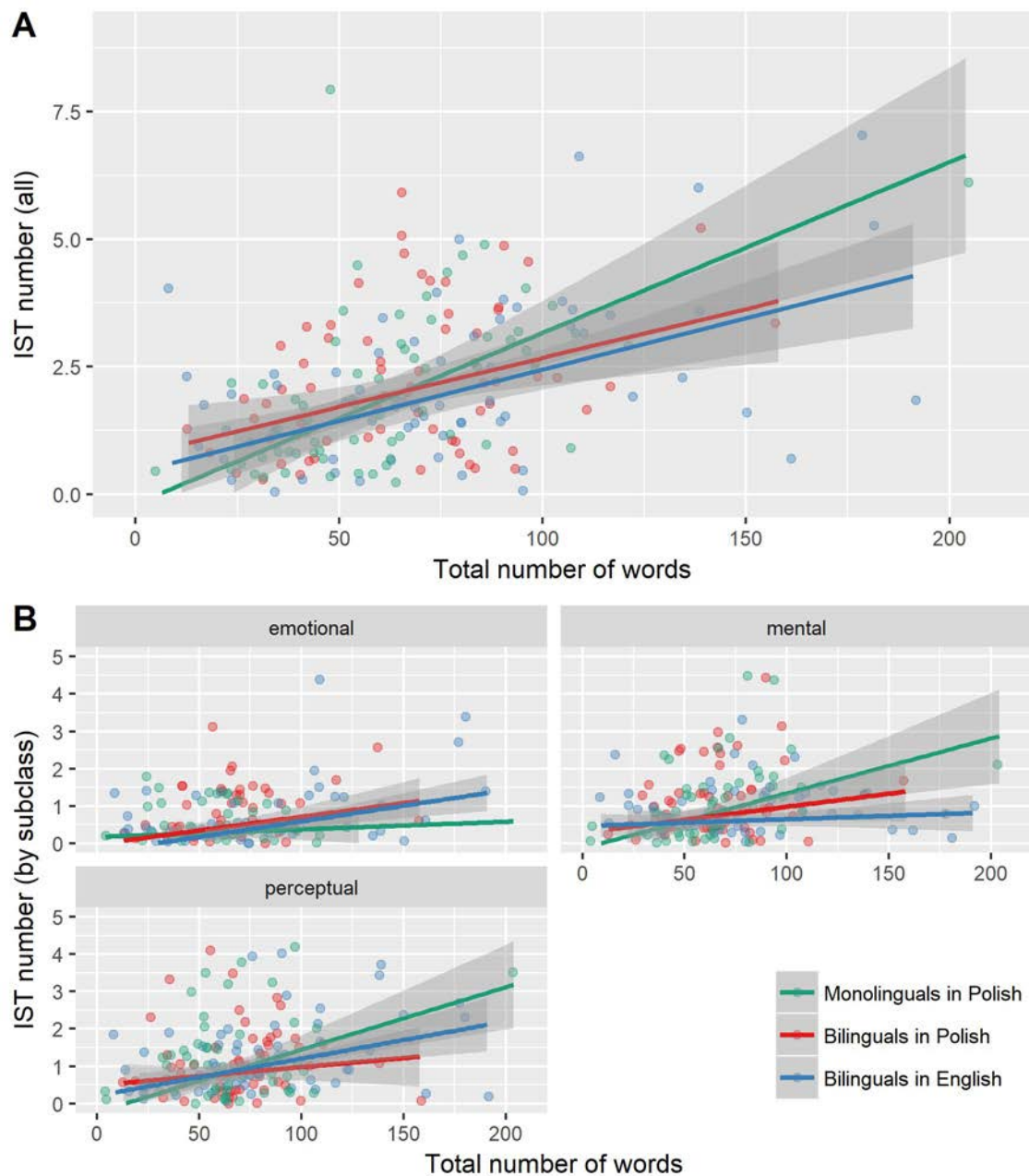


Figure 10. The number of (A) all internal state terms (tokens) and (B) tokens by category in the Polish and English told stories plotted against the total number of words in the Polish and the English stories.

Table 13.

Spearman correlation coefficients from the correlational analyses between different factors and the number of all internal state terms (tokens) in the stories told by monolinguals and bilinguals. Only significant correlations are reported.

	Number of internal state terms (all tokens)		
	Monolinguals in Polish	Bilinguals in Polish	Bilinguals in English
Age	ns	ns	0.38***
Receptive vocabulary	ns	ns	0.26*
Expressive vocabulary	ns	ns	ns
Receptive grammar	ns	0.29*	ns
ToM - appropriateness of reflection	ns	ns	ns
ToM - intensity of reflection	ns	ns	0.28*
Story structure	0.47***	0.23*	ns
Total number of words (TNW)	0.55***	0.32**	0.49***

* = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$.

The results from all the correlational analyses are summarized in Table 13. Next, two correlational matrices were created in order to investigate relations between all the measures: one for the Polish set of data ($n=150$, bilinguals and monolinguals merged) presented in Table 14 and one for English set of data ($n=75$, bilinguals), presented in Table 15. The correlational matrices show that in both Polish and English narratives, the total number of words correlates positively with the story structure (Polish: $\rho = 0.44$, $p \leq 0.001$, English: $\rho = 0.28$, $p \leq 0.05$). This inter-correlation between the independent variables warns us to treat the subsequent regression analyses with caution: the estimated impact of one independent variable (e.g. TNW) on the amount of IST while controlling for the other variable may be less precise than if predictors were uncorrelated with one another.

Table 14.

Spearman correlation matrix for the Polish data (n=150, bilinguals and monolinguals merged).

	Age	PL TNW	PL all IST number	PL IST/TNW proportion	PL receptive vocab	PL expressive vocab	PL receptive grammar	ToM – appropriateness	ToM - intensity
Age									
PL TNW	0.01								
PL all IST number	0.05	0.44***							
PL IST/TNW proportion	-0.01	-0.04	0.82***						
PL receptive vocab	0.44***	0.04	0.14	0.10					
PL expressive vocab	0.26**	-0.04	0.01	0.01	0.65***				
PL receptive grammar	0.38**	0.16	0.18*	0.10	0.59***	0.38***			
ToM - appropriateness	0.42***	0.00	0.12	0.14	0.36***	0.08	0.47***		
ToM - intensity	0.46***	0.08	0.14	0.09	0.37***	0.05	0.52***	0.80***	
PL story structure	0.14	0.44***	0.37***	0.16	0.32***	0.18*	0.27**	0.12	0.23**

* = $p \leq 0.5$, ** = $p \leq 0.01$, *** = $p \leq 0.001$.

Table 15.

Spearman correlation matrix for the English data (n=75, bilinguals).

	Age	EN TNW	EN all IST number	EN IST/TNW proportion	EN receptive vocab	EN expressive vocab	EN receptive grammar	ToM – appropriateness	ToM - intensity
Age									
EN TNW	0.04								
EN all IST number	0.38***	0.49***							
EN IST/TNW proportion	0.31**	-0.22	0.50***						
EN receptive vocab	0.55***	0.06	0.26*	0.16					
EN expressive vocab	0.57***	0.09	0.13	0.06	0.80***				
EN receptive grammar	0.45***	0.25	0.04	0.15	0.58***	0.53***			
ToM - appropriateness	0.38***	0.14	0.17	0.05	0.58***	0.56***	0.56***		
ToM - intensity	0.51***	0.14	0.28*	0.19	0.60***	0.52***	0.62***	0.80***	
EN story structure	0.22	0.23*	0.14	0.00	0.25*	0.23*	0.37**	0.26*	0.37**

* = $p \leq 0.5$, ** = $p \leq 0.01$, *** = $p \leq 0.001$.

Following the correlational analyses, two multiple regression analyses were carried out to identify the best model predicting the use of internal state terms in the children's told stories in both languages. When investigating the Polish stories, both the bilinguals and monolinguals' scores were taken into account ($N = 150$). This was done because the correlational analyses revealed similar patterns for both groups when it comes to their Polish stories. When investigating the English stories, only the bilinguals' scores were taken into account ($n = 75$). In both regression models, the all-subsets method was used. The all-subsets method (here: with *regsubsets()* function in the *leaps* package in R, Lumey & Miller, 2004) performs an exhaustive search for the best regression model, containing a subset of predictors used in the maximal model. The maximal model for each language contained the predictors that were significantly correlated with the number of internal state terms in a narrative in a given language. The Polish maximal model contained the receptive grammar test score, story structure score and the total number of words produced in a story. The English maximal model contained age, receptive vocabulary score, theory of mind score – the intensity of reflection, and the total number of words produced in the English narrative.

Table 16 presents the best regression model that predicted the overall number of internal state terms used in the Polish narratives told by bilingual and monolingual children. The significant predictors were the total number of words produced in a narrative and the story structure score: the larger number of internal state terms in a narrative was related to the larger number of total words in a story, and the higher story structure score. Such a model explained 22% of the variance in the overall number of internal state terms used, $F(1,148) = 21.21$, $p < 0.0001$, *Adj. R squared* = 0.22.

Table 16.

The best regression model predicting the number of internal state terms (all tokens) in the Polish narratives of bilingual and monolingual children.

	<i>Estimate</i>	<i>SE</i>	<i>t</i>	<i>P</i>
Intercept	-0.32	0.39	-0.81	0.419
Total Number of Words	0.02	0.01	4.35	0.000
Story structure	0.16	0.06	2.43	0.017

However, the relationship was also partially true the other way around: when a maximal model containing story structure score and the number of IST (i.e. two significant correlates of TNW, see Table 14) was used to predict the total number of words in a story, it showed these two constituted the best model to predict the TNW. Such model explained 30% of variance, $F(2,143) = 29.1$, $p < 0.0001$, *Adj. R squared* = 0.3 (see Table 17). The fact that the relationship was found both ways implies the data do not yield conclusive results as to the directionality of the influence. Therefore, it must be concluded that the total number of words in a story, the story structure score and the number of internal state terms in a Polish narrative are interrelated, with no implication of the directionality of the relationship.

Table 17.

The best regression model predicting the total number of words in the Polish narratives of bilingual and monolingual children.

	<i>Estimate</i>	<i>SE</i>	<i>t</i>	<i>P</i>
Intercept	27.75	5.77	4.81	0.000
All IST (tokens)	5.44	1.25	4.36	0.000
Story structure	4.23	0.98	4.3	0.000

Table 18 presents the best regression model that predicted the overall number of internal state terms used in the English narratives told by bilingual children. The significant predictors were age and the total number of words in the English narrative: the larger number of internal state terms in the English narratives was related to the larger number of words in a story and the bilinguals' age. The model explained 36% of the variance, $F(2,72) = 21.99$, $p < 0.0001$, $Adj. R squared = 0.36$.

Table 18.

The best regression model predicting the number of internal state terms (all tokens) in the English narratives of bilingual children.

	<i>Estimate</i>	<i>SE</i>	<i>t</i>	<i>P</i>
Intercept	-0.32	1.23	-3.09	0.003
Total Number of Words	0.02	0.01	5.54	0.000
Age	0.06	0.02	3.44	0.001

However, the opposite relation was also partially true: when the same maximal model containing story structure score and the number of IST (i.e. two significant correlates of TNW, see Table 15) was used to predict the total number of words in a story, it showed that the overall number of internal state terms (tokens) produced in the English narrative was the best predictor of the TNW, $\beta = 0.53$, $t(73) = 5.29$, $p = 0.000$, explaining 27% of variance, $F(1,73) = 27.98$, $p = 0.000$, $Adj. R squared = 0.27$ (see Table 19). Since the relationship was bidirectional, it must be concluded that the total number of words in a story and the number of internal state terms in the English narrative are interrelated, with no implication of the directionality of the relationship.

Table 19.

The best regression model predicting the total number of words in the English narratives of bilingual children.

	<i>Estimate</i>	<i>SE</i>	<i>t</i>	<i>P</i>
Intercept	51.41	6.08	8.46	0.000
All IST	12.58	2.38	5.29	0.000

5.2.4. RQ4: Do children differ in the amount of internal state terms when telling a story, as compared to when they retell a story after listening to a model story told by an adult?

It was hypothesized that the children produce more IST-saturated narratives when they retell a story immediately after listening to a model story told by an adult as compared to when they tell a story by themselves. In the model story, internal state terms amounted to 14 out of 180 words, giving the overall IST ratio of 7.8% (see Subsection 4.3.6.). In order to establish whether the retelling effect on the use of internal state terms in narratives is global, the between-modes comparisons were performed on bilinguals and monolinguals and in both languages of the bilingual children.

Comparison of Telling and Retelling modes in the bilingual and monolingual stories in Polish

The number of children who used at least one internal state term in their told and retold narratives can be found in Table 20. The descriptive statistics for the use of internal state terms by bilinguals and monolinguals in their told and retold stories in Polish can be seen in Table 21 (number of IST) and Table 22 (proportion of IST in the total number of words in a story). A non-parametric two-way ANOVA was conducted to compare the proportion of IST in the bilingual and monolingual told stories and stories retold after a model. The ANOVA investigated the main effects of mode (telling, retelling) and group (bilingual, monolingual), and the interaction effects of mode and group. The main effect of mode was significant, *Test statistic* = 63.15, $p = 0.001$, with more internal state terms produced in the stories retold after the model than in the spontaneously told stories. The main effect of group was non-significant, *Test statistic* = 0.09, $p = 0.76$. The effect of the interaction between mode and group was also non-significant, *Test statistic* = 1.84, $p = 0.18$. Thus, both bilinguals and monolinguals produced more IST-saturated narratives when retelling the story after the experimenter than when spontaneously telling the story.

A series of non-parametric Wilcoxon Signed Rank Tests investigated the differences between the two modes (telling, retelling) in the proportions of particular subclasses of IST. All children used significantly more mental terms in their retold stories, compared to their told stories: bilinguals, $V = 471, p < 0.001$, monolinguals, $V = 381, p < 0.001$. Also, all children used significantly more perceptual terms in their retold stories, compared to their told stories: bilinguals, $V = 463, p < 0.001$, monolinguals, $V = 312, p < 0.001$. However, the between-mode difference in the proportion of emotional terms was non-significant either in monolinguals, $V = 253.5, p = 0.22$, or bilinguals, $V = 586, p = 0.47$. The results are pictured in Figure 11.

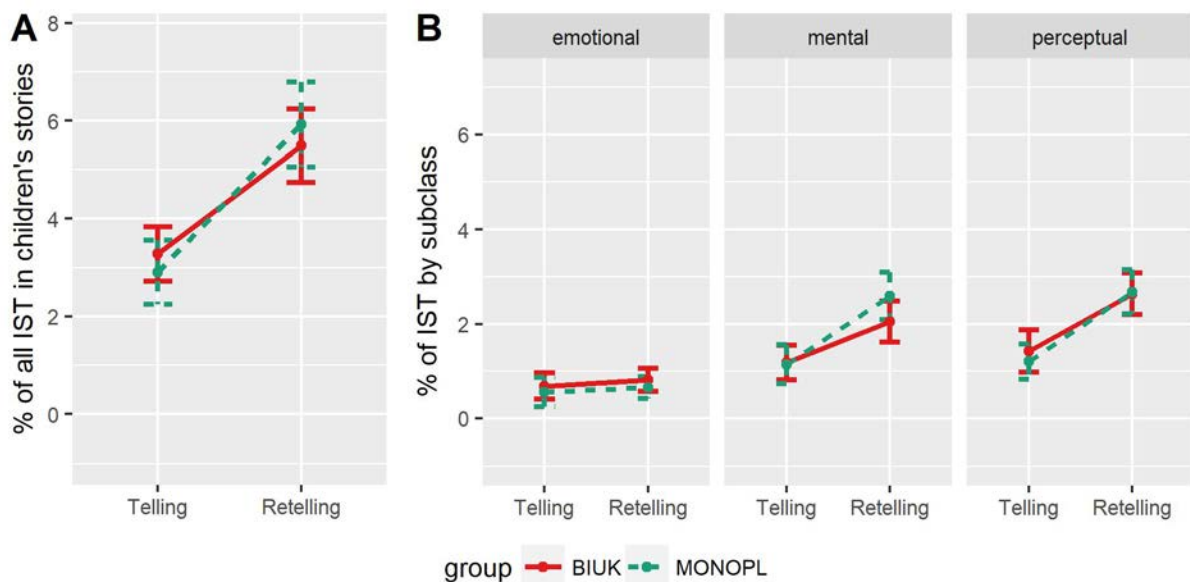


Figure 11. The proportion of internal state terms (tokens) in stories told and retold by bilinguals and monolinguals in Polish: (A) internal state terms by subclass, (B) all internal state terms. The error bars represent the 95% confidence interval (CI).

Table 20.

The number of children who used no internal state terms in their told and retold Polish stories (IST = 0) and the number of children who used at least one internal state term (IST > 0).

		Bilinguals (No. of children)	Monolinguals (No. of children)
Told stories	IST = 0	12	20
	IST > 0	63	55
Retold stories	IST = 0	6	8
	IST > 0	69	67

Table 21.

The descriptive statistics for the number of internal state terms (tokens) used in stories told and retold by bilinguals and monolinguals in Polish: medians, means and standard deviations.

	Bilinguals				Monolinguals			
	Telling		Retelling		Telling		Retelling	
	Md	(M ± SD)	Md	(M ± SD)	Md	(M ± SD)	Md	(M ± SD)
All IST (number)	2	2.03 ± 1.5	5	4.61 ± 2.59	1	1.84 ± 1.85	4	4.13 ± 2.48
Emotional IST (number)	0	0.47 ± 0.76	1	0.71 ± 0.83	0	0.29 ± 0.71	0	0.51 ± 0.72
Mental IST (number)	0	0.75 ± 0.97	1	1.72 ± 1.46	0	0.76 ± 1.13	2	1.72 ± 1.28
Perceptual IST (number)	1	0.81 ± 0.97	2	2.19 ± 1.51	0	0.79 ± 1.08	2	1.91 ± 1.38

Table 22.

The descriptive statistics for the proportion of internal state terms used in stories told and retold by bilinguals and monolinguals in Polish: medians, means and standard deviations.

	Bilinguals				Monolinguals			
	Telling		Retelling		Telling		Retelling	
	Md	(M ± SD)	Md	(M ± SD)	Md	(M ± SD)	Md	(M ± SD)
All IST (%)	3	3.28 ± 2.43	5	5.49 ± 3.27	3	2.90 ± 2.83	6	5.92 ± 3.78
Emotional IST (%)	0	0.68 ± 1.2	1	0.81 ± 1.05	0	1.56 ± 1.36	0	0.66 ± 1
Mental IST (%)	0	1.18 ± 1.58	2	2.04 ± 1.9	0	1.15 ± 1.82	2	2.59 ± 2.2
Perceptual IST (%)	1	1.42 ± 1.96	3	2.64 ± 1.93	0	1.2 ± 1.62	3	2.67 ± 2.04

Comparison of Telling and Retelling modes in the bilingual stories in Polish and English.

The number of children who used at least one internal state term in their Polish and English narratives can be found in Table 23. The descriptive statistics for the use of internal state terms in the bilinguals' told and retold stories in Polish and English can be seen Table 24 (number of IST) and Table 25 (proportion of IST). A non-parametric two-way ANOVA was conducted to compare the overall proportion of internal state terms (all subclasses) used by bilinguals when telling a story and when retelling the story after the model. The ANOVA investigated the main effects of mode (telling, retelling) and language (Polish, English), and the interaction effects of mode and language. The main effect of mode was significant, Test statistic = 39.58, $p = 0.001$, with more internal state terms produced in the retelling mode than in the telling. The main effect of language was non-significant, Test statistic = 0.35, $p = 0.56$. The effect of the interaction between mode and language was also non-significant, Test statistic = 0.44, $p = 0.51$. To confirm this, a series of Wilcoxon Signed Rank Tests was run. No cross-language differences were found in the use of IST in the retelling mode: there was no difference between Polish and English proportion of all IST, $V = 1217$, $p = 0.36$, or in the proportion of emotional, $V = 685$, $p = 0.2$, mental, $V = 1041$, $p = 0.24$, or perceptual terms, $V = 1075.5$, $p = 0.25$. Thus, bilinguals produced more IST-saturated narratives in the retelling mode than in the telling regardless of the language they were speaking.

A closer look at the particular subclasses of internal state terms revealed significant between-modes differences only in the number of mental and perceptual terms. Specifically, children used significantly more mental terms in their retold stories, compared to their told stories both in Polish, $V = 471$, $p < 0.001$, and in English, $V = 357.5$, $p < 0.001$. They also used significantly more perceptual terms in their retold stories, both in Polish, $V = 788$, $p = 0.29$, and in English, $V = 575.5$, $p = 0.01$. The between-mode difference was non-significant for the use

of emotional terms, either in Polish, $V = 586, p = 0.47$, or in English, $V = 263, p = 0.28$. The results are pictured in Figure 12.

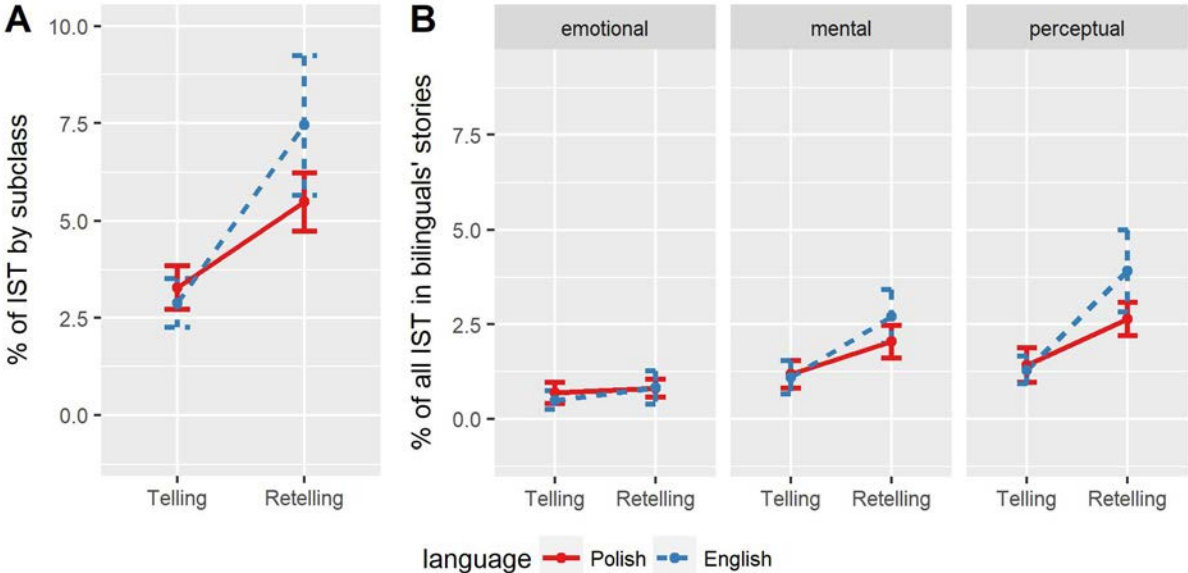


Figure 12. The proportion of internal state terms (tokens) in stories told and retold by bilinguals in Polish and in English: (A) all internal state terms, (B) internal state terms by subclass. The error bars represent the 95% confidence interval (CI).

Table 23.

The number of bilingual children who used no internal state terms in their told and retold Polish and English stories ($IST = 0$) and the number of children who used at least one internal state term ($IST > 0$).

		Polish (No. of children)	English (No. of children)
Told stories	IST = 0	12	16
	IST > 0	63	59
Retold stories	IST = 0	6	6
	IST > 0	69	69

Table 24.

The descriptive statistics for the number of internal state terms (tokens) used in the bilingual stories told and retold in Polish and English: medians, means and standard deviations.

	Polish				English			
	Telling		Retelling		Telling		Retelling	
	Md	(M ± SD)	Md	(M ± SD)	Md	(M ± SD)	Md	(M ± SD)
All IST (number)		2.03 ± 1.5		4.61 ± 2.59		1.95 ± 1.67		4.69 ± 2.66
Emotional IST (number)		0.47 ± 0.76		0.71 ± 0.83		0.39 ± 0.78		0.59 ± 0.9
Mental IST (number)		0.75 ± 0.97		1.72 ± 1.46		0.6 ± 0.7		1.91 ± 1.6
Perceptual IST (number)		0.81 ± 0.97		2.19 ± 1.51		0.96 ± 1.09		2.2 ± 1.46

Table 25.

The descriptive statistics for the proportion of internal state terms used in the bilingual stories told and retold in Polish and English: medians, means and standard deviations.

	Polish				English			
	Telling		Retelling		Telling		Retelling	
	Md	(M ± SD)	Md	(M ± SD)	Md	(M ± SD)	Md	(M ± SD)
All IST (%)	3	3.28 ± 2.43	5	5.49 ± 3.27	3	2.88 ± 2.71	5	7.45 ± 7.77
Emotional IST (%)	0	0.68 ± 1.2	1	0.81 ± 1.05	0	0.5 ± 1.1	0	0.83 ± 1.9
Mental IST (%)	0	1.18 ± 1.58	2	2.04 ± 1.9	0	1.1 ± 1.93	2	2.7 ± 3.11
Perceptual IST (%)	1	1.42 ± 1.96	3	2.64 ± 1.93	1	1.29 ± 1.6	3	3.91 ± 4.72

5.2.5. RQ5: Do children use more internal state terms when explicitly engaged in a conversation about the mental states of story protagonists, as compared to the amount of internal state terms in their spontaneous telling?

It was hypothesized that children use more internal state terms when explicitly engaged in a conversation about the mental states of story protagonists, as compared to the amount of internal state terms in their spontaneous telling. To test whether this effect is global, the analysis was performed on both bilinguals and monolinguals and on both languages of the bilingual children. Because the total number of words in the answers to the comprehension questions was not gathered, therefore proportions of IST to the TNW could not be calculated. Any analyses on the proportions of IST had to be excluded. Hence, the analyses performed for this research questions were done only on the number of internal state terms (altogether and by subclass).

Comparison of Telling and Questions contexts in the bilingual and monolingual stories in Polish.

The number of children who used at least one internal state term when telling stories and when answering comprehension questions in Polish can be seen in Table 26. The descriptive statistics for the number of internal state terms when bilinguals and monolinguals told stories (telling) and when they were explicitly asked about protagonists' internal states (comprehension questions) can be seen in and Table 27 (number of IST). A non-parametric two-way ANOVA was conducted to compare the overall number of internal state terms used by the bilinguals and monolinguals when telling a story and when answering comprehension questions. The ANOVA investigated the main effects of context (telling, comprehension questions) and group (bilingual, monolingual), and the interaction effects of context and group. The results revealed a significant main effect of context, Test statistic = 115.31, $p = 0.001$, with more internal state terms elicited in the questions than in the telling. The main effect of group was also significant, Test statistic = 6.37, $p = 0.01$, with

bilinguals using more internal state terms than monolinguals. However, a closer look at the particular subclasses revealed group differences only in the area of emotional terms and only in the answers to the comprehension questions: in their answers, bilinguals used significantly more emotional terms than monolinguals, $W = 3359.5$, $p = 0.04$. There were no differences between the groups on the use of mental terms, $W = 2957.5$, $p = 0.58$, and perceptual terms, $W = 2837$, $p = 0.91$, in their answers to the comprehension questions. The interaction effect of context and group was also non-significant, Test statistic = 1.33, $p = 0.25$. Thus, both bilinguals and monolinguals used more internal state terms when answering the comprehension questions than when spontaneously telling the story.

However, some more insight into the group-difference in the use of emotional terms when answering comprehension questions comes from a qualitative look at the Polish and English interviews of the children, as exemplified by a bilingual girl aged 6;5:

An excerpt from the English interview (comprehension questions) from a bilingual girl, 6;5:

- *EXP: So, why is mummy bird flying away here?
- *CHI: Because she wanted=wanna@ 0to find some food.
- *EXP: Okay.
- *EXP: And how do you think the baby birds feel?
- *CHI: Amm.
- *CHI: **Sad**.
- *EXP: Why do you think so?
- *CHI: &Becau- Because, because their=they@ mummy flew ... up to had\$ some food.
- *EXP: And why is the cat climbing the tree?
- *CHI: Because he wanted=wanna ... 0to get the bird and eat it.
- *EXP: Okay and how did the baby birds feel in this picture?
- *CHI: Amm **scared**.
- *EXP: And why do you think that?
- *CHI: Because, because a cat, &becaus- they didn't see=saw@ first the cat and he made them scared and then he wanted=wanned 0to catch them but he felt\$ so so scared.
- *EXP: Okay.
- *EXP: And why did the dog grab the cat's tail?
- *CHI: Because he wanted=wanna ... do it so he ... climbed=climb ... off the tree.
- *EXP: Okay, amm, imagine that the dog sees the birds.
- *EXP: How does the dog feel?
- *CHI: Amm, **happy**?
- *EXP: Why is that?
- *CHI: &Becau- because, amm, because he 0is gonna be very nice to the birds and the cat is gonna go, fly=flew@ away.

The same girl, Polish interview:

- *EXP: A dlaczego mała koza jest w wodzie?
[Why is the small goat in the water?]
- *CHI: Bo, bo chciała uratować małą kozę.
[Because she wanted to save the small goat.]
- *EXP: Mhm, a jak się czuje małe kozłátko?
[Mhm, and how does the baby goat feel?]
- *CHI: **Bardzo źle.**
[**Very bad.**]
- *EXP: A dlaczego?
[And why?]
- *CHI: Bo się utopiło w wodzie.
[Because it drowned in the water.]
- *EXP: Mhm, a dlaczego lis skacze w stronę kozłátka?
[Mhm, and why is the fox jumping towards the goat?]
- *CHI: Bo chce go zjeść.
[Because he wants to eat him.]
- *EXP: A jak się czuje małe kozłátko?
[And how does the baby goat feel?]
- *CHI: **Źle.**
[**Bad.**]
- *EXP: A dlaczego?
[And why?]
- *CHI: Bo, bo lis złapał za nogę.
[Because, because the fox caught [his] leg.]
- *EXP: Mhm, a dlaczego ptak złapał lisa za ogon?
[Mhm, and why did the bird catch the fox's tail?]
- *CHI: Bo chciał uratować małą kózkę=kozkę@.
[Because he wanted to save the baby goat.]
- *EXP: A...wyobraź sobie, że ptak widzi teraz kózki.
[And... imagine that the bird can see the goats.]
- *EXP: Jak się czuje?
[How does he feel?]
- *CHI: Ten, kto?
[He, who?]
- *EXP: Wyobraź sobie, że ptak widzi teraz kózki, jak się czuje?
[Imagine that the bird can see the goats now, how does he feel?]
- *CHI: **Dobrze?**
[**Good?**]
- *EXP: Mhm, a dlaczego?
[Mhm, and why?]
- *CHI: Bo, mhm, bo, bo on łapie lisa, który jest bardzo niegrzeczny.
[Because, mhm, because because he is catching the fox who is very naughty]

It can be seen that when answering similar questions in English (“how does X feel?”), the child refers to the protagonist’s emotions (“sad”, “happy”, “scared”), but when answering the same questions in Polish, the girl uses very general expressions (“feels bad”, “feels very bad”, “feels good”), which were not counted as instances of internal state terms. Moreover, the way this girl answers the Polish questions is in fact characteristic of the way Polish monolingual children answer as well, as evinced by an interview from a monolingual girl of the same age (6;5):

- *EXP: Dlaczego mama ptaszków odleciała.
[Why did the mommy bird fly away?]
- *CHI: Przynieść=przynieś im coś do jedzenia.
[To bring them something to eat.]
- *EXP: A jak się czują małe ptaszki?
[And how do the baby birds feel?]
- *CHI: Tak yyy, tak się trochę **boją**?
[They are **scared** a bit?]
- *EXP: A dlaczego?
[And why?]
- *CHI: Bo one są małe i, i boją się, że im coś się stanie.
[Because they are little, and and they are scared that something will happen to them.]
- *EXP: Mhm, a dlaczego kot wchodzi na drzewo?
[Mhm, and why is the cat climbing the tree?]
- *CHI: Bo chce=kce jedno ugryźć=ugryś i sobie złapać.
[Because he wants to bite one [bird] and catch [it].]
- *EXP: A jak się czuje mały ptaszek?
[And how does the baby bird feel?]
- *CHI: **Źle**.
[**Bad.**]
- *EXP: A dlaczego?
[And why?]
- *CHI: Bo bo bo bo chce=kce kot go zabrać na za jedzenie, a mama &odl- przyleciała i zobaczyła i nie nakarmiła, a a małe pisklątko chciał=kciało mmm yyy żyzy chciało zostać, ale ten kot mu nie pozwalał.
[Because because because the cat wants to take him for food and mommy flew back and saw [it] and didn’t feed [the birds], and and the baby bird wanted mmm wanted to stay but the cat wouldn’t let him.]
- *EXP: A Powiedz, dlaczego pies złapał kota za ogon?
[And tell me, why did the dog grab the cat’s tail?]
- *CHI: Żeby yyy ... bo on jest dobry i żeby i żeby on yyy &k- małego pisklątka nie zabrał.
[To ... because he is good and to, to, for the cat not to take the baby bird.]
- *EXP: Mhm.
- *EXP: A wyobraź sobie, że pies widzi teraz ptaszki.
[And imagine the dog can see the birds.]
- *EXP: Jak się czuje?
[How does he feel?]
- *CHI: Yyy ... **dobrze**.
[Yyy... **good.**]

- *EXP: Dlaczego?
[Why?]
- *CHI: Bo nie je małych ptaszków i jest dobry.
[Because he doesn't eat the small birds and he's good.]

There were also significant differences between the contexts in the number of particular subclasses of ISL used by all the children. Both groups used significantly more references to emotional states in the questions than in their told stories, bilinguals: $V = 160.5, p < 0.001$, and monolinguals: $V = 117.5, p < 0.001$, and more references to mental states, bilinguals: $V = 104.5, p < 0.001$, monolinguals: $V = 173.5, p < 0.001$. There was also a significant difference in the number of perceptual terms used by the children in the two contexts. Both groups used fewer perceptual terms when answering comprehension questions than when telling their stories: bilinguals, $V = 532.5, p = 0.001$, monolinguals, $V = 560.5, p = 0.004$. The results are pictures in Figure 13.

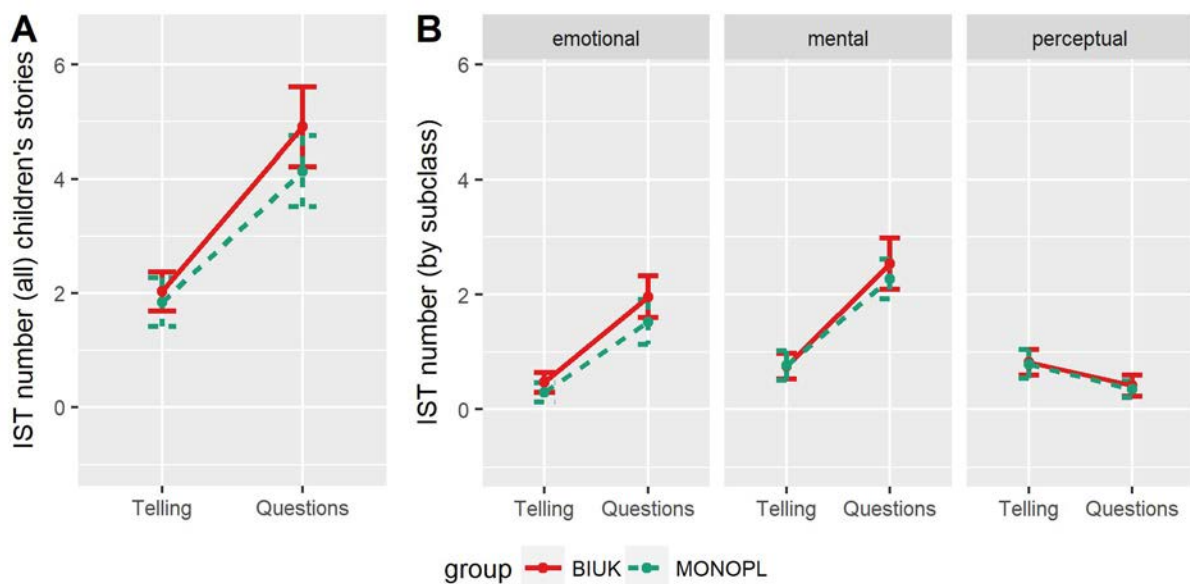


Figure 13. The number of internal state terms (tokens) when bilinguals and monolinguals tell stories (telling) and when they are explicitly asked about protagonists' internal states (comprehension questions): (A) all internal state terms, (B) internal state terms by subclass. The error bars represent the 95% confidence interval (CI).

Table 26.

The number of children who used no internal state terms when telling stories and when answering comprehension questions in Polish (IST = 0) and the number of children who used at least one internal state term (IST > 0).

		Bilinguals (No. of children)	Monolinguals (No. of children)
Told stories	IST = 0	12	20
	IST > 0	63	55
Answers to comprehension questions	IST = 0	9	9
	IST > 0	66	66

Table 27.

The descriptive statistics for the number of internal state terms (tokens) when bilinguals and monolinguals tell stories (telling) and when they are explicitly asked about protagonists' internal states (questions) in Polish: medians, means and standard deviations.

	Bilinguals				Monolinguals			
	Telling		Questions		Telling		Questions	
	Md	(M ± SD)	Md	(M ± SD)	Md	(M ± SD)	Md	(M ± SD)
All IST (number)	2	2.03 ± 1.5	5	4.91 ± 3.02	1	1.84 ± 1.85	4	4.13 ± 2.68
Emotional IST (number)	0	0.47 ± 0.76	2	1.96 ± 1.57	0	0.29 ± 0.71	1	1.52 ± 1.71
Mental IST (number)	0	0.75 ± 0.97	3	2.53 ± 1.94	0	0.76 ± 1.13	2	2.27 ± 1.5
Perceptual IST (number)	1	0.81 ± 0.97	0	0.41 ± 0.81	0	0.79 ± 1.08	0	0.35 ± 0.63

Comparison of Telling and Questions contexts in the bilingual stories in Polish and English.

The descriptive statistics for the number of Polish and English internal state terms when bilinguals told stories (telling) and when they were explicitly asked about protagonists' internal states (comprehension questions) can be seen in Table 28. A non-parametric two-way ANOVA was conducted to compare the overall number of internal state terms used by bilinguals when telling a story and when answering comprehension questions. The ANOVA investigated the

main effects of context (telling, comprehension questions) and language (Polish, English), and the interaction effects of context and language. The results revealed a significant main effect of context, Test statistic = 247.2, $p = 0.001$, with more internal state terms elicited in the questions than in the telling. A significant main effect of language was also found, Test statistic = 10.88, $p = 0.002$, with more internal state terms appearing in English than in Polish. This difference was also evinced by the content of the children's stories as mentioned in the subsection above. Finally, there was a significant effect of interaction between context and language, Test statistic = 14.54, $p = 0.001$. Thus, bilinguals used more internal state terms when explicitly asked about the protagonists internal states in English (as compared to Polish and as compared to the number of internal state terms in their spontaneously told stories).

A series of paired Wilcoxon Signed Rank Tests was conducted to explore the between-context differences in the particular ISL subclasses. There was a significant difference in the number of emotional and mental and perceptual terms in both languages. Specifically, children used more emotional terms when answering the questions both in Polish, $W = 1184$, $p < 0.001$, and in English, $W = 579.5$, $p < 0.001$. They also used more mental terms in their answers to questions than in their spontaneous tellings, in Polish, $W = 1237$, $p < 0.001$, and in English, $W = 983$, $p < 0.001$. However, they used significantly fewer perceptual terms in their answers to questions than in their told stories, both in Polish, $W = 3528.5$, $p = 0.002$, and in English, $W = 3456$, $p = 0.01$. The results are pictured in Figure 14.

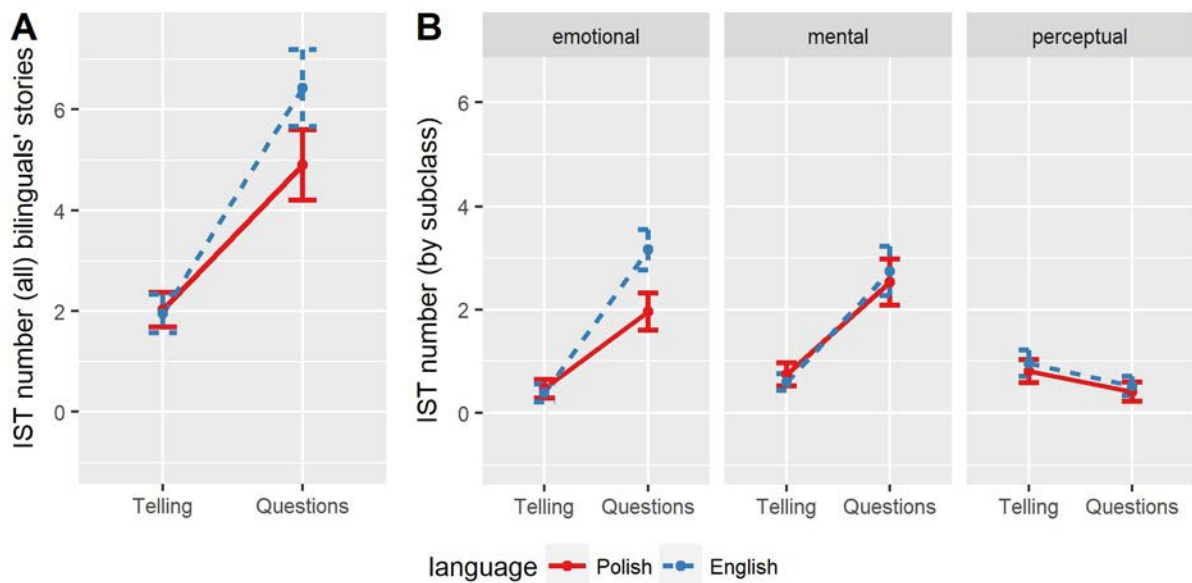


Figure 14. The number of internal state terms (tokens) when bilinguals tell stories (telling) and when they are explicitly asked about protagonists' internal states (comprehension questions): (A) all internal state terms, (B) internal state terms by subclass. The error bars represent the 95% confidence interval (CI).

Table 28.

The descriptive statistics for the internal state terms (tokens) when bilinguals tell stories (telling) and when they are explicitly asked about protagonists' internal states (Questions) in both languages: medians, means and standard deviations.

	Polish				English			
	Telling		Questions		Telling		Questions	
	Md	(M ± SD)	Md	(M ± SD)	Md	(M ± SD)	Md	(M ± SD)
All IST (number)	2	2.03 ± 1.5	5	4.91 ± 3.02	2	1.95 ± 1.67	7	6.43 ± 3.34
Emotional IST (number)	0	0.47 ± 0.76	2	1.96 ± 1.57	0	0.39 ± 0.78	3	3.16 ± 1.71
Mental IST (number)	0	0.75 ± 0.97	3	2.53 ± 1.94	0	0.6 ± 0.7	3	2.75 ± 2.08
Perceptual IST (number)	1	0.81 ± 0.97	0	0.41 ± 0.81	1	0.96 ± 1.09	0	0.52 ± 0.81

5.3. The summary of the results

The present section contains a summary of the results. A detailed discussion of the results with references to previously published findings is the focus of the next chapter. The introductory analyses presented in this chapter revealed that the bilingual children performed significantly lower than the monolinguals in all the language tasks: receptive vocabulary, expressive vocabulary and receptive grammar. They also seemed balanced across their two languages: they scored comparably on both the receptive vocabulary and the receptive grammar tests. They also outperformed their monolingual peers on the theory of mind test: they showed better appropriateness of reflection and better intensity of reflection. These results may serve as a background of the subsequent outcomes.

The main analyses investigated the use of ISL in bilingual children as compared to their monolingual peers and between the two languages of the bilinguals. The first research questions considered the use of ISL across the bilingual and monolingual children. Taking all internal state terms together, bilingual and monolingual children used a similar amount of internal state terms (number and proportion) when telling a story in Polish. All internal state terms (tokens) constituted on average 3% of the total number of words in both the bilinguals' and monolinguals' stories. When considering the particular subclasses of internal state terms (emotional, mental, perceptual), there were no statistically significant differences between the bilingual and monolingual groups in the proportion of the particular terms used. Hence, the bilingual and monolingual narratives were similarly saturated with the particular subclasses of IST. However, both groups used fewer emotional terms than mental or perceptual terms in their Polish told stories. This lack of difference between the two groups is particularly interesting in the view of introductory analyses, namely the language gaps between the bilingual and monolingual children, and the bilingual advantage in the area of reflection on thinking. This matter is taken up in more detail in the Discussion (Chapter 6).

The second research question was focused on the ISL use across the two languages of the bilinguals. Considering all internal state terms together, bilingual children used a similar amount of internal state terms (number and proportion) when telling their stories in Polish and in English. Thus, the bilingual stories contained a similar number of IST and were similarly saturated with internal state terms in both languages. All internal state terms (tokens) constituted on average 3% of the total number of words in both the bilinguals' and monolinguals' stories. When considering the particular subclasses of internal state terms (emotional, mental, perceptual), there were no statistically significant differences between the languages in the amount of the particular terms used. Hence, the proportions of the particular subclasses of IST were similar between the bilinguals' languages. However, both Polish and English stories of the bilinguals contained fewer emotional terms than mental or perceptual terms. There were no significant cross-language correlations in the amount of internal state terms used by bilinguals in the Polish and English stories. It is also worth noting that the bilinguals studied here were fairly balanced in their languages (though their language performance was lower than that of Polish monolingual peers). The matter is discussed in detail in the next chapter, but this may indicate that the ISL develops comparably in both languages of the bilingual child or that there may be a cross-language transfer of skills and knowledge.

The third research question focused on the predictors of the ISL use in the narratives told in Polish and in English. The two languages yielded slightly different results. In Polish, the number of internal state terms in the narratives was related to the (bilingual and monolingual) children's total number of words produced in a story, their story structure score, and their receptive grammar score. The regression analysis showed that the larger number of internal state terms in a narrative was related to the larger number of total words in a story, and the higher story structure score. However, the opposite relation was also true, i.e. the total number of words was significantly predicted by the overall number of Polish internal state terms and

the Polish story structure score. The number of internal state terms in the English narratives was related to the total number of words produced in a story, and the bilingual children's age. The regression analysis showed that the larger the number of total words in a story, and the older the age, the larger the overall number of internal state terms in the English narrative. However, the opposite relation was also true: the total number of words in a narrative was significantly predicted by the overall number of internal state terms in a story. Interestingly, the children's vocabulary size (receptive or productive) and performance on the theory of mind task did not emerge as significant predictors of internal state terms production in a narrative context.

Next, two research questions focused on the narratives as a tool for eliciting and improving the ISL in children. The fourth research question investigated the effect of adult modelling on the use of IST. The results showed that both bilinguals and monolinguals produced more IST-saturated narratives when retelling the story after the experimenter than when spontaneously telling the story. A closer look at the particular subclasses of ISL revealed an increase (in the retold stories) of mental and perceptual terms. However, the between-mode difference in the use of emotional terms was non-significant. The results also showed that bilinguals produced more IST-saturated narratives in the retelling mode than in the telling regardless of the language they were speaking. This was true for the use of all internal state terms, and the mental and perceptual subclasses. However, the use of emotional terms did not change between the modes, either in Polish or in English.

The last research question compared the use of IST when telling a story and when answering comprehension questions that focused on the protagonist's internal states and reasons for particular actions. When considering the overall number of internal state terms, bilingual and monolingual children generally referred more to the internal states of the story protagonists when answering the comprehension questions about their internal states, as

compared to when they spontaneously told a story. A closer look at the particular subclasses revealed that in their answers, bilinguals used significantly more emotional terms than monolinguals. No such differences were found for the perceptual or mental terms. When explicitly engaged in a conversation about the internal states of story protagonists, all children, regardless of the language they were speaking, used more emotional and mental terms, and fewer perceptual terms.

The next chapter includes a detailed discussion of the results presented above as related to previously published findings, as well as conclusions and practical implications for parents and practitioners working on language development in bilingual children.

Chapter 6: The discussion

The present thesis set out to investigate the use of internal state terms in Polish-English bilingual children at pre- and early school age (4.5-7 years old). This was done by comparing the bilinguals' IST production to a group of monolingual peers, and across the bilinguals' two languages. There is an evident lack of research on the Internal State Lexicon in bilinguals, even though their general linguistic and cognitive development is comprehensively examined. As a matter of fact, bilingual upbringing may shape bilingual use of ISL to look differently than ISL in monolingual children. First, bilingual children hear less of each language, when compared to monolinguals, which may translate to lower language performance (e.g. Haman, et al., 2017). On the other hand, bilingual upbringing yields also some advantages due to dual language activation in the mind. Having two languages constantly active in the mind forces bilinguals to exert constant control over the two languages. As a result, bilinguals are often reported to outperform monolinguals on tasks tapping into executive functions and cognitive abilities, including theory of mind (Farhadian et al., 2010; Goetz, 2003; Kovács, 2009). Internal State Lexicon might be placed at the intersection of the language and ToM development: the internal state terms need to be learned any other words in the lexicon, and their use is an indicator of the children's mentalizing abilities. The matter of ISL use in this specific population will thus complete the picture of the bilingual child development.

One of the primary aims of the present analysis was to investigate whether language status (bilingual vs. monolingual group) may influence the use of Internal State Lexicon. Bilingual and monolingual children in the analysed sample produced a similar amount of internal state terms when telling a story in Polish. The groups were similar both in the number of internal state terms, and in the proportion of internal state terms in the total number of words in a story. Moreover, this was true both for the overall amount of IST, and for each of the studied subclasses (emotional, mental, perceptual terms). Thus, the results suggest that the language

status (bilingual vs. monolingual) does not influence the amount of internal state terms produced in a story-telling context at pre- and early school age. This is in accordance with studies that examined the overall amount of internal state terms in bilinguals' narratives. They indicate that bilingual children use a similar amount of internal state terms as their monolingual peers when telling a story in the language of majority (L2) (Bonifacci et al., 2017). Also, previous analyses on a sample partly overlapping (Otwińska et al., 2018) showed that bilinguals use a similar overall amount of internal state terms when telling their stories in Polish (their home language, L1), as their monolingual peers. The present analysis addressed the limitation of the previous studies by not only considering the overall amount of internal state terms but also exploring the particular subclasses of ISL, and including not only the raw numbers of IST but also their proportions in the number of total words in a story. The present study also confirms and extends the results from a study by Shiro, Hoff, Ribot and Shanks (2017, in press) that investigated the ISL production in the context of a structured mother-child play. Shiro and collaborators found that, when the overall number of internal state terms was considered, bilinguals and monolinguals produced a similar total amount of types (the number of unique terms) and tokens (the number of occurrences of each type of a term). However, they also found that monolinguals used more types when referring to cognition and volition states than bilinguals (either in English or Spanish). This led them to assume that bilinguals in fact may have a smaller internal state vocabulary than monolinguals, in reference to particular subclasses of the ISL. Though the present analysis does not confirm this result, it must be noted that the sample studied here was considerably older (children aged 4.5-7) than the one in the study by Shiro and colleagues (children aged 2.5). Thus, the results from the two studies are not necessarily contradictory, as each may picture the state of the bilingual Internal State Lexicon at a different point in the child's development.

The bilingual and monolingual groups studied here differed significantly in their language performance in Polish, as measured by the expressive and receptive vocabulary size and receptive grammar skills. The bilinguals attained lower scores on all three language tests, compared to the monolingual peers. This is in line with the available evidence showing that bilingual children have smaller vocabularies in each of their languages, when compared to monolingual peers (for comparisons with majority language peers see: Bialystok, Luk, Peets, & Yang, 2010; De Houwer, Bornstein, & Putnick, 2014; Hoff et al., 2012; Hoff, Rumiche, Burrige, Ribot, & Welsh, 2014; Thordardottir, 2011; for comparisons with home language peers see: Haman et al., 2017; Mieszkowska et al., 2017). However, bilinguals also outperformed the monolingual peers on the theory of mind task that assessed the understanding of the 1st and 2nd order beliefs, understanding of deception, ambiguity and interpretation, and understanding of surprise. Though difficult to vouch for, the two differences might have evened each other out, leading to no overall difference in the use of ISL between the groups. But there is one thing of greater importance: the finding that the bilinguals' smaller vocabulary size in Polish (compared to monolinguals) did not translate to smaller internal state vocabulary, a point made by Shiro and colleagues (2017). Hence, the relatively smaller vocabulary of bilingual children does not necessarily impede their production of internal state terms (as compared to monolinguals).

It was also found that the internal state terms constituted 3% of all words in both bilinguals' and monolinguals' stories. This confirms previous investigations of spontaneous speech which found that cognitive terms appear in up to 5% of pre- and early school children's utterances (e.g. Bartsch & Wellman, 1995; Pascual et al., 2008; Tardif & Wellman, 2000). This low proportion might be also related to the nature of the task at hand, i.e. telling a fictional story based on a set of pictures. The narrative task includes two dimensions of a narrative: a landscape of action and a landscape of consciousness (Bokus, 2000, 2004, 2013; Bruner, 1986). This

duality offers an opportunity to describe and label internal states, but it is also possible to tell a story with a sole focus on the landscape of action without any internal references (Symons et al., 2005).

Finally, both groups used significantly more mental and perceptual terms than emotional terms. This again might be a result of the task used. The terms used by many of the children were: “chcieć” (want) – used by 43% of bilinguals and 40% of monolinguals, and “zobaczyć” (see) – used by 25% of bilinguals and 23% of monolinguals. The common use of the perception terms might indicate that children attribute knowledge or lack thereof based on whether a person saw or did not see something happen. Thus, children may use both cognitive and perception terms to refer to the other’s state of knowledge, as evinced by the examples from children’s narratives presented in Chapter 5 (Results). A relatively greater use of mental and perceptual terms over emotional terms shows that, in the context of telling a story, pre- and early school children generally focus more on those internal states that allow them to explain the protagonist’s actions, and give less attention to referring the protagonists’ emotions. Indeed, research shows that children this age usually start to include internal state terms as initiating events, i.e. an event that triggers the protagonist’s action (Berman & Slobin, 1995).

Another aim of the analysis was to check whether ISL develops comparably in both languages of the bilingual child. The results show clearly that bilingual children use a similar amount of internal state terms when telling their stories in Polish and in English. This was true both for the number of IST in a story, and for the proportion of IST to the total number of words in a story. There were also no cross-language differences in the particular subclasses of the terms (emotional, mental, perceptual). These results corroborate findings from a study done on a partly overlapping sample (coming from the same source project) of Polish-English bilingual children aged 3-7 years old which showed that bilinguals used the same overall amount of internal state terms when telling their stories in Polish and English (Otwinowska et al., 2018).

The explanation may be twofold. First, the ISL may develop comparably in both languages of the bilingual child or there may be some cross-language transfer of skills and knowledge (as suggested in the Interdependence Hypothesis, Cummins, 1979). However, the lack of difference could be also attributed to a fairly balanced proficiency in the two languages, as evinced by the children's largely similar performance in the tasks measuring receptive vocabulary size and receptive grammar abilities in the two languages. It would be interesting to investigate whether unbalanced bilinguals show a similar pattern of results, i.e. whether they produce a similar amount of internal state terms in both languages, regardless of their language proficiency. However, such an investigation was not possible with the sample studied within the Bi-SLI-Poland project, as the majority of the children were fairly balanced bilinguals.

It must be also noted that there was no significant cross-language correlation in the overall amount of internal state terms used by bilinguals in the Polish and English stories. An exploration of the particular subclasses of ISL revealed a significant (positive) correlation only in the amount of perceptual terms in the Polish and English narratives, i.e. children who used many perceptual terms in their narratives in one language, also used many perceptual terms in their narratives in the other language. The failure to establish a relation between the overall ISL production in the two languages might have been caused by very low values of internal state terms in general: children produced a similar number of internal state terms in both languages, and that is generally a low number (all internal state terms constituted 3% of the total number of words in both the Polish and English stories). Nonetheless, the reason might be more profound and reflect either the effects of exposure to the ISL in a particular language (as suggested by Altman, Armon-Lotem, Fichman, & Walters, 2016). Altman and colleagues (2016) found that bilingual preschool children used significantly more cognitive terms in their L2 (Hebrew) narratives than their L1 (English) narratives. While they point out that the L2 narratives in their study were shorter and lower frequencies of IST could have led to higher

ratios in the L2 than in the L1, they also caution that the result might be actually mediated by the effects of exposure. Specifically, mental terms could have been more frequent in the L2 because they could be more frequent in the L2 input from the Hebrew preschool, where mental verbs such as “think”, “know”, “forget”, “decide”, “believe” were probably used on a daily basis. Since the present study did not involve an examination of the children’s ISL input patterns, the results of the present analysis do not allow for a conclusive answer to Altman’s et al. argument. However, the present study found no difference in the proportion of mental terms between the two languages. Moreover, the most frequently emotional terms that were used by the bilinguals in the Polish stories were also the most frequently used terms used in the English stories, e.g. “wystraszony” (scared), “szczęśliwy” (happy), and “smutny” (sad). The case was the same for the mental and perception terms. A closer look at the stories of the particular children also seemed to point to a general similarity in the use of ISL. Thus, importantly, the ISL use across the two languages of a bilingual child seems to be largely similar in terms of the amount of references to internal states, and in terms of the particular states that children tend to refer to. It might be that tendency per se to focus on the internal states of story protagonists is generally language independent and reflects the child’s mentalizing abilities, or their tendency to spontaneously make use of these abilities. This issue brings us to the next question posed in the present thesis

The present analysis also set out to find the best predictors of ISL production in the Polish and English narratives. The potential predictors included the children’s age, vocabulary size and receptive grammar knowledge in each language, theory of mind performance, the story structure, and the total number of words produced in a story. Interestingly, the results from correlations between the number of internal state terms produced in the narratives in each language and the mentioned predictors yielded different patterns for Polish and English. In Polish, the overall number of internal state terms correlated significantly (and positively) only

with the story structure score, and the total number of words in a story both in bilinguals and monolinguals. This suggests that the large number of internal state terms in the Polish story was associated with a well-developed ability to tell a coherent story that includes the protagonists' goals, their attempts at the goals, and the outcomes of their actions, and child's verbosity in a story, i.e. great amount of words in a story. Accordingly, story structure and the total number of words constituted the best model predicting the overall number of all internal state terms in a Polish story.

The pattern was slightly different with the English stories of bilingual children. Here, the number of internal state terms correlated positively with age, receptive vocabulary size, the intensity of reflection in the theory of mind task, and the total number of words in the English story. Of these, the significant predictors of the number of internal state terms in an English narrative were the children's age and the total number of words in a story. Thus, the larger number of internal state terms in the English narratives was related to the bilinguals' age and the larger number of words in a story. Though research has generally demonstrated a steady increase over time in the number of internal state terms used by children, the evidence came either from children younger than those studied here (Becker et al., 2017; Rudek & Haden, 2005) or from monolingual preschoolers (Booth & Hall, 1995; Astington & Pelletier, 2003). Here of interest is the influence of age observed in the second or majority language of bilinguals (English), but absent in their first or home language (Polish). A possible explanation is that age is an important predictor of the use of internal state terms in early childhood, but its impact on the tendency to refer to internal states in a story washes out in pre- and early school age (hence no significant correlations in the Polish monolingual population). In bilinguals, this pattern is valid in their L1, but not the L2 – where age is still an important indicator of the number of references to internal states in a narrative. This might be related to the accumulated exposure to English in bilingual children. The majority of children studied here had both Polish parents and

spoke Polish at home. Moreover, they were acquiring Polish from birth, and thus their length of exposure to Polish was equal their age. Due to largely Polish-speaking environment at home, the bilinguals' Polish Internal State Lexicon might be similar to that of their monolinguals peers. However, the bilingual children may receive relatively less exposure to ISL in English (a point raised by Altman et al., 2016). First of all, their average onset of acquisition of English was around their first birthday. Thus, their accumulated exposure to English (in general) was smaller than their exposure to Polish. Additionally, the ISL input in English for these children might come from book-reading and story-telling at preschools and via other regular activities in the majority language (English), which for the younger of the participants may not be yet taking the larger part of a day. If that were correct, we might expect to see the bilinguals' English ISL grow with time spent in the English schooling system, here obscured as their age. Also, it must be noted that the relations between the internal state terms and the predicting variables were in fact bidirectional. Therefore, the number of internal state terms and the total number of words in a story and the story structure are in fact interrelated, and the present data do not yield conclusive results as to the directionality of the relationship. Nevertheless, the established strong links between the use of ISL and the story structure score are not to be underestimated. In fact, telling a good story requires the narrator to consider the listener's perspective and present him/her with the important story elements, such as the setting, the logical sequence of events (e.g. goals, attempts and outcomes). And it seems that the better storyteller a child is, the more they refer to internal states of their story protagonists.

It is also worth noting that the present analysis failed to find any direct link between the use of IST and the vocabulary knowledge (receptive or productive) or the mentalizing abilities to reflect on thinking. Specifically, though receptive vocabulary and intensity of reflection on thinking did correlate moderately with the number of internal state terms produced by bilinguals in English, neither of these variables proved to be important predictors of the use of ISL.

Moreover, no such relations were observed for Polish. This might corroborate previous results from Meins et al. (2006) and Charman and Shmueli-Goetz (1998) who did not find any significant correlations between the children's use of internal state terms (in storytelling or when describing a friend) and their receptive vocabulary size or theory of mind performance. Meins and colleagues (2006) suggest that in fact, some individual differences may be at play, specifically some factors that do not reflect language abilities or ToM development, e.g. motivational and personality differences. Charman Shmueli-Goetz (1998) also suggest that in an experimental setting, children are not involved or motivated enough to make genuine attributions of internal states. In the long run, it may be that acquiring theory of mind does not strictly entail that the child will refer to internal states during an on-line task of narrating a story. Nevertheless, it would be interesting to explore the relation between the child's vocabulary, theory of mind and the ISL production in a slightly different context of spontaneous interaction, e.g. play time with parent, peers or siblings.

Another explanation for no relation between the use of IST and the vocabulary knowledge (receptive or productive) or the mentalizing abilities is that these factors might be influential for the ISL acquisition at the earlier stages of child development. We have seen that Nielsen and Dissanayake (2000) found a significant correlation between false-belief tasks and the use of ISL in preschool children, but Charman and Shmueli-Goetz (1998) who studied 7 year olds, did not find any significant relation. It might be possible that the role of language and theory of mind factors on the ISL acquisition decreases with time and actually gives way to some other factors, e.g. social factors such as the presence of siblings or peer-to-peer interaction (e.g. Brown & Dunn, 1992).

One more surprising result was a moderate correlation between receptive grammar abilities and the number of internal state terms found in bilinguals in Polish. This should be mentioned together with a strong correlation between receptive grammar abilities and theory of

mind performance (both appropriateness and intensity of reflection on thinking). It confirms the argument upheld mainly by De Villiers that passing a false-belief task requires the understanding of complex grammatical structures (Astington & Jenkins, 1999; de Villiers, 2000; de Villiers, 2007; Hale & Tager-Flusberg, 2003, see also Section 2.5). Another significant cross-task correlation is that between the receptive grammar and the story structure in the produced narratives (see Table 14 and Table 15). Together, these cross-task correlations might show that grammar abilities are a reflection of more general communicative ability, and thus are to some extent indicative of mentalizing abilities as well.

In addition, the present analysis aimed to assess the narrative context as a potential source of the ISL and a medium of the ISL enhancement. Generally, giving children a model story and asking them comprehension questions about the internal states of story protagonists sensitized them to their knowledge, desires, beliefs which resulted in more internal state references in the retellings and answers to the comprehension questions. All children referred more to internal states of story protagonists in the stories retold after an adult model (compared to their spontaneously told stories). Thus, a model story told by an adult draws the children's attention to the internal states of the characters and encourages them to use more internal state terms. This completes the picture presented by previous studies showing that when retelling stories after a model, children use more elaborate vocabulary and address more goals of the protagonists (Isbell et al. 2004; Peterson and McCabe 1992). It also suggests that retelling as a narrative training strategy might not only improve children's understanding of the internal states of the story protagonists, as shown by Lewis and collaborators (1994), but also increase children's production of internal state terms. Moreover, these effects were independent of group (bilingual, monolingual) and the language (Polish, English) in which the bilingual stories were told. Thus, the retelling effect seems to be global, i.e. was observed in both bilinguals and monolinguals, and across bilingual's languages.

The retold stories showed an increase of the number of references to mental and perceptual terms (relative to the told stories) in both groups and in both languages. On the other hand, there was no significant increase between the modes in the area of emotional terms (either in bilingual or monolinguals, or in bilinguals' Polish or English stories). This might confirm that pre- and early school children focus less on the protagonists' emotions and more on mental and perceptual states that allow them to comment on the protagonists' knowledge and explain their actions. However, we must note that the model stories told by the experimenters contained more references to mental states (6 out of 14 internal state terms), and perceptual states (5 out of 14 internal state terms) than emotional terms (3 out of 14 internal state terms). Thus, no increase in the emotional references in the retellings might actually mirror low proportion of these terms in the model.

Both bilinguals and monolinguals used more internal state terms when answering the comprehension questions about their internal states, as compared when they spontaneously told a story. That this effect of context was independent of group and appeared in both languages of the bilinguals advances the previous findings on monolingual children. These showed that children use more internal state terms when interacting with peers and adults (Brown et al., 1996; Hughes et al., 2007; Pinto et al., 2016), than in a non-interactive context, e.g. when telling a story (Meins et al., 2006). The present analysis also corroborates Mein's claim that developing theory of mind does not necessarily translate to spontaneously using the ToM abilities to describe, explain and interpret the behaviour of others.

Interestingly, there were also some cross-language differences: bilinguals used more internal state terms when explicitly asked about the protagonists internal states in English (as compared to Polish). This difference was largely fed by the emotional states, i.e. bilingual children used more emotional terms in their English interviews than in the Polish ones. An explanation comes from a qualitative look at the Polish and English interviews of the children:

based on the excerpt of a narrative told by a bilingual girl in Polish in English (see Subsection 5.2.5.), when answering comprehension questions in English (“how does X feel?”), the girl referred to the protagonist’s emotions (“sad”, “happy”, “scared”), but when answering the same questions in Polish, she used very general expressions (“feels bad”, “feels very bad”, “feels good”), which were not counted as instances of internal state terms. Moreover, the way this girl answered the Polish questions was in fact similar to way Polish monolingual children answered as well, as evinced by an interview from a monolingual girl of the same age. At the beginning of this chapter it was pointed out that the use of internal state terms seems to be largely similar across the two languages of a bilingual, both in terms of the number of references to internal states, and in terms of the particular states that children tend to refer to. It was suggested that the tendency to refer the internal states of story protagonists is generally language independent. However, while this can be principally true, it is possible that at the same time there might be slight cross-linguistic differences in the way such tendency is executed in the particular languages. The Polish expression “czuć się dobrze/źle” (to feel good/bad) is in fact a common answer to the question about a person’s disposition and might well explain the low proportion of emotional terms in the Polish interviews (compared to the English interviews). Thus, it is possible that in some cases, e.g. in some of the subclasses of internal state terms, the coding and scoring of internal state terms is liable to cross-linguistic differences, caused for example by the differences in common and typical responses to the same questions. This finding stresses the importance of using transparent and comprehensive classifications of terms when studying the acquisition of internal state terms.

Study limitations

The present study has certain limitations. First, the task used to elicit internal state terms, i.e. Multilingual Assessment Instrument for Narratives (MAIN, Gagarina et al., 2012) was designed with a general purpose of investigating narratives, not internal state terms per se. Although the review of different methods used for studying the ISL (see section 1.3.) revealed that structured narratives based on pictures may be a useful method with regards to elicitation of internal state terms, there is still a gap in finding an optimal method. In particular, it is possible that the picture stories used in the present study might not have been filled enough with events which rely on the understanding of the story characters' internal states, although this issue was considered when the tool was constructed (Gagarina et al., 2015). That is directly linked with another limitation of this study, namely, a low amount of internal state terms produced in the narratives. Though the ISL generally constitutes a small proportion of children's language production, and the low amount of IST was thus expected, it nevertheless forced the use of non-parametric tests in the statistical analyses. These are characterized by comparatively weaker power and might have obscured some of the results otherwise visible.

There are a few possible solutions to the problem that could be undertaken in the future but could not be done with the present data. First, a potential tool used to elicit internal state terms should require the child to refer more to the mental representations of story characters. For example, Bokus (1991; 2013) created stories specifically for the purpose of investigating mind-reading in children's narratives. For example, her pictures included instances of misrepresentation not present in the pictures used here. Also, the stories included in that studies did not obviously state what action should the protagonist undertake and some of the events had to seem extraordinary from the narrator's view (e.g. boys flying). Similarly, in a series of studies, Symons et al. (2005) used a story that included a mistaken identity revealed at the end of the book, or specially drawn pictures that presented some emotional scene (e.g. a man coming

down the stairs into a dark, cellar carrying a gift-wrapped box under his arm, and a boy standing behind rubbing his eyes while tears ran down his cheeks). Interestingly, Symons et al. (2005) have found a significant correlation between the use of ISL in these tasks and the performance on theory of mind test, a link not established here. Thus, using pictures or stories filled with events which rely on understanding the story characters' internal states might elicit more internal state terms which would allow for more valid analyses. Similarly, one could design a study where more or longer samples of child language are elicited. Such a solution could also yield more credible results, for example by eliciting narratives in various moments throughout the child's day.

On the other hand, one might decide to measure children's spontaneous tendency to refer to internal states of self and others. This can be done via a more interactive context, e.g. long-term observations of children during unstructured play-time with their parents, siblings or peers. This might be especially important in the view of the present results, suggesting that children's production of internal state terms might be in fact independent of the children's performance of theory of mind tasks. The present thesis cannot provide a clear answer as to what other factors – instead of theory of mind – might contribute to the amount of internal state terms production in children. However, the results indicate that as to the production of ISL in a narrative context, the child's ability to tell a coherent story is of obvious importance. In fact, telling a good story requires the narrator to consider the listener's perspective and present him/her with the important story elements.

Overall, although the MAIN was not primarily designed to elicit the ISL and this might add to the low amount of produced IST, its use in the Bi-SLI-Poland project was justified by the possibility of wide cross-linguistic comparisons. Besides, the overall internal state terms ratio to total words in a story amounting to 3% (in bilinguals and monolinguals, in Polish and English) is comparable to the ratios reported by other researchers (Bartsch & Wellman, 1995;

Pascual et al., 2008; Tardif & Wellman, 2000). Likewise, the tool proved to be good enough for analyzing three main categories of internal state terms (emotional, mental, perceptual) when comparing with other existing methods also having their limitations. Thanks to the possibility of cross-linguistic comparisons it also opens the opportunity to study ISL across various languages which might be one of the future directions discussed in subsequent section.

Further directions

There are few important strands that future research on the ISL use in bilingual children could explore. First, it might focus on the development of the Internal State Lexicon in children over time. This has been – for a long time – the focus of the ISL research in monolingual children. The present results juxtaposed with the available evidence suggest this might be an important issue in the case of bilinguals as well. Shiro et al. (2017, in press) found differences between bilingual and monolingual children in regards to the use of specific subclasses of IST. The present study has found none. However, as pointed out in the Discussion, the sample studied here was considerably older (children aged 4.5-7) than the one in the study by Shiro and colleagues (children aged 2.5). Thus, the results from the two studies still leave out a relevant stage of child development not accounted for and call for an investigation of developmental trajectories in monolinguals and bilinguals across different ages. It would be particularly interesting to study the acquisition of ISL in bilinguals longitudinally, and that in both of their languages.

Another reason for exploring the ISL longitudinally is that different factors might be influential for the ISL acquisition at different stages of child development. We have already seen that in the course of child development, there is a shift from the importance of the quantity to the importance of quality of parental state talk (Slaughter et al., 2007; see Section 2.4.).

It is possible that as there is a shift of specific social factors important to the acquisition of ISL, similarly the child-individual factors change their role. Specifically, it may be that language abilities or theory of mind performance are more important to the ISL at the early stages of child development, but give ground to other factors later on. This might be why Nielsen and Dissanayake (2000) found a significant correlation between false-belief tasks and the use of ISL in preschool children, but Charman and Shmueli-Goetz (1998) failed to establish such a link in 7 year old children. For now, we do not know when precisely such a shift could take place and whether it needs to be preceded by reaching a specific level of language or theory of mind performance. It is also unclear what other factors (apart from social ones, such as peer-to-peer interaction) might be influential for the ISL in school-age children.

It would be also extremely important to connect the ISL production to the ISL input found in the children's immediate environment: e.g. family members, peers, teachers, and during regular activities, e.g. story-times at home or in other available contexts (e.g. library). Thus, future research might examine how bilingual children understand and use internal state terms in both of their languages, and investigate to what degree this is influenced e.g. by the parental use of internal state terms in child-directed storytelling. Indeed, research on monolinguals has found that maternal storytelling rich in references to mental states significantly advances the child's competence in the area of mental state use and understanding (e.g. Ruffman et al., 2002).

Notably, in the case of bilinguals, input containing internal state terms in the L1 may serve as the underlying foundation in determining the meaning of internal states later acquired in the L2 as well. The data analysed here showed that bilinguals use the ISL similarly in both of their languages. However, it would be interesting to explore whether the competence to discern other minds, developed in the first language, enhances the growth of the same competence in the second language. In other words, does the prior acquisition of a concept help

in the succeeding acquisition of a new word in another language? If that were the case, when the bilingual children first acquire internal state concepts in their home language, this conceptual knowledge may serve as a scaffolding for the acquisition of the corresponding terms in their second language. Thus, further research could investigate closely the possibility of a cross-language conceptual transfer of the knowledge of internal state terms. This question also opens the door to exploring cross-cultural differences in the ISL acquisition. Indeed, studies comparing the ToM performance between notably distant cultures (e.g. American vs. Chinese, American vs. Japanese), have indicated some cross-cultural differences e.g. in the order of acquisition of specific ToM concepts. If the cultural factors, such as history, tradition, parenting styles can shape ToM (Afeke & Gut, 2018; Gut, 2016), it is also possible they exert influence on the acquisition of internal state terms. This influence may be visible in the content of the parental ISL use or in the order of the acquisition of ISL subclasses in children. In such studies, the MAIN (Gagarina et al., 2012), which was designed to elicit narratives from children coming from various cultures, should prove a convenient tool to elicit language samples from children speaking different languages.

Last but not least, future research on the acquisition of ISL might include both measures of production and comprehension of internal state terms. A task measuring internal state terms comprehension might be an important enrichment and give a more complete picture of the ISL acquisition. Just like the acquisition of general vocabulary, the acquisition of full meanings of the Internal State Lexicon is a gradual process that can be tapped into. Thus, the ISL acquisition in children can be investigated by measuring the children's understanding of small discrepancies between the full meanings of internal state terms, e.g. the difference between "know", "think" and "guess".

General conclusions and practical implications

The present study informed us that while bilinguals and monolinguals may differ significantly in relation to their language and theory of mind performance, they do not differ in the amount of internal state terms produced when telling a story. Also, bilinguals use the ISL similarly in their two languages. These two findings broaden our scope of knowledge on the bilingual language acquisition in children. They also show that the relatively poorer language abilities of bilingual children do not impede their production of internal state terms. Next, it was found that giving children a model story and asking them comprehension questions about the internal states of story protagonists sensitized them to their knowledge, desires, beliefs which resulted in more internal state references in the retellings and answers to the comprehension questions. This has important implications for the future investigations of the ISL production in narratives and theory of mind development. Specifically, it shows that developing theory of mind does not necessarily translate to spontaneously using the ToM abilities to describe, explain and interpret the behaviour of others. This should caution us against overestimating the narrative production as a measure of theory of mind. Also, it shows that children may require an additional incentive to make use of their mentalizing abilities. This in turn has important implications for parents and practitioners working with children. It shows that the use of internal state terms may be enhanced and shaped by this particular behaviour of sharing story-telling with children and engaging them in conversations about the protagonist's internal states: their knowledge, belief and emotions. Though the present study does not provide evidence of long-lasting effects of such modelling practices, it should nevertheless sensitize parents and teachers who engage in story telling or book reading to their role as narrator models. Thus, they may be encouraged to include numerous references to internal states in their interactions with children, to explain cause and effects of different internal states and to draw the children's attention to those states by asking them questions about the protagonists'

behaviour. Notably, such scaffolding behaviour in parents is not to be assumed as self-evident, as there are possible cultural differences in the parental beliefs about storytelling (see e.g. Zevenbergen, Haman, & Olszańska, 2012; Zevenbergen et al., 2016). Finally, the efficacy of such modeling could be the aim of future studies, designed with a specific goal of investigating the effect of modelling in natural interactions.

Also, in the context of bilingual upbringing, parents who use a minority language at home (i.e. language other than that spoken by the majority of the society) may be inclined to concentrate on boosting their children's proficiency in the majority language, in order to ease their admittance into the surrounding society. In extreme cases, this may lead them to decrease the use of the home language at home, in favor of the community language, especially if the home language is of lower social prestige than the majority language (Gathercole and Thomas, 2009). However, the present research suggests that bilingual children who speak a minority language at home are perfectly capable of acquiring the ISL in their majority language as well, and moreover, they can use the ISL comparably in both of their languages and at a level similar to monolingual peers, even when their general abilities in the L1 and the L2 (e.g. vocabulary size) is relatively lower when compared to monolinguals. Thus, families that use a minority language (L1) at home may be encouraged to use internal state language towards children, even in one language only, as it may still influence the child's use and understanding of internal state terms in both languages.

The present thesis points also to some methodological issues that should be considered. Specifically, as was shown by the review of lists of words coded as internal state terms (see Appendix A), studies of the ISL differ in regard to the choice of the investigated subclasses and specific terms. The review list included twelve studies that appended at least a substantial list of terms coded as the ISL, investigated typically developing child populations and derived their terms from children's speech. Out of 441 items in the list, 48 items were categorized differently

in at least 2 out of 12 studies, and 9 items were categorized differently in 3 out of 12 studies. This means that altogether 13% of the items were categorized into different subclasses across studies. Thus, divergent categorization of items is not that uncommon and should be a matter of researchers' careful examination. Moreover, researchers should be attentive in excluding from their analyses those instances of the ISL use as idiomatic or conversational device, i.e. without indicating the actual internal states. A natural solution would be for researchers studying the ISL to include in their publications full lists of words coded. Then it would be possible to compare the similarities and differences in the study results in the view of their ISL classifications. Such a practice could help us disentangle genuine differences in the results from those caused by diverging classifications of the ISL.

References

Adrián, J. E., Clemente, R. A., & Villanueva, L. (2007). Mothers' Use of Cognitive State Verbs in Picture-Book Reading and the Development of Children's Understanding of Mind: A Longitudinal Study. *Child Development*, 78(4), 1052–1067.

Adrian, J. E., Clemente, R. A., Villanueva, L., & Rieffe, C. (2005). Parent–child picture-book reading, mothers' mental state language and children's theory of mind. *Journal of Child Language*, 32(3), 673–686. <https://doi.org/10.1017/S0305000905006963>

Afek, J., & Gut, A. (2018). The Chinese Model of Early Education from the Perspective of Cross-Cultural Psychological Research. *Annals of Arts*, 65(9), 5–28. <https://doi.org/10.18290/rh.2017.65.9-1>

Altman, C., Armon-Lotem, S., Fichman, S., & Walters, J. (2016). Macrostructure, microstructure, and mental state terms in the narratives of English–Hebrew bilingual preschool children with and without specific language impairment. *Applied Psycholinguistics*, 37(1), 165–193. <https://doi.org/10.1017/S0142716415000466>

Antonietti, A., Liverta-Sempio, O., Marchetti, A., & Astington, J. W. (2006). Mental Language and Understanding of Epistemic and Emotional Mental States. In A. Antonietti, O. Liverta-Sempio, & A. Marchetti (Eds.), *Theory of Mind and Language in Developmental Contexts* (pp. 1–30). Springer US. Retrieved from http://link.springer.com/chapter/10.1007/0-387-24997-4_1

Astington, J. W., & Baird, J. A. (2005). *Why Language Matters for Theory of Mind*. Oxford, UK: Oxford University Press.

Astington, J. W., & Jenkins, J. M. (1995). Theory of mind development and social understanding. *Cognition and Emotion*, 9(2–3), 151–165. <https://doi.org/10.1080/02699939508409006>

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology*, *35*(5), 1311–1320. <https://doi.org/10.1037/0012-1649.35.5.1311>

Astington, J. W., & Pelletier, J. (2004). *Theory of mind and metacognitive vocabulary development in first and second languages*. Unpublished manuscript, University of Toronto.

Au, T. K., & Glusman, M. (1990). The Principle of Mutual Exclusivity in Word Learning: To Honor or Not to Honor? *Child Development*, *61*(5), 1474–1490. <https://doi.org/10.2307/1130757>

Baron-Cohen, S. (1991). Precursors to a theory of mind: Understanding attention in others. In A. Whiten (Ed.), *Natural theories of mind: Evolution, development and simulation of everyday mindreading* (pp. 233–251). Oxford: Blackwell.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a ‘theory of mind’? *Cognition*, *21*(1), 37–46. [https://doi.org/10.1016/0010-0277\(85\)90022-8](https://doi.org/10.1016/0010-0277(85)90022-8)

Bartsch, K., & Wellman, H. (1989). Young Children’s Attribution of Action to Beliefs and Desires. *Child Development*, *60*(4), 946. <https://doi.org/10.2307/1131035>

Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind* (Vol. viii). New York, NY, US: Oxford University Press.

Becker Razuri, E., Hiles Howard, A. R., Purvis, K. B., & Cross, D. R. (2017). Mental State Language Development: The Longitudinal Roles of Attachment and Maternal Language. *Infant Mental Health Journal*, *38*(3), 329–342. <https://doi.org/10.1002/imhj.21638>

Berman, R. A. (1988). On the ability to relate events in narrative. *Discourse Processes*, *11*(4), 469–497. <https://doi.org/10.1080/01638538809544714>

Berman, R. A., & Slobin, D. I. (1995). *Relating Events in Narrative: A Crosslinguistic Developmental Study*. Psychology Press.

Białecka-Pikul, M. (2004). Nabywanie wiedzy społecznej z perspektywy badań nad dziecięcymi teoriami umysłu. *Przegląd Psychologiczny*, 47(3), 305–317.

Białecka-Pikul, M. (2012). *Narodziny i rozwój refleksji nad myśleniem*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.

Bialystok, E. (1987). Influences of bilingualism on metalinguistic development. *Second Language Research*, 3(2), 154–166. <https://doi.org/10.1177/026765838700300205>

Bialystok, E., Craik, F. I. M., Green, D. W., & Gollan, T. H. (2009). Bilingual Minds. *Psychological Science in the Public Interest*, 10(3), 89–129. <https://doi.org/10.1177/1529100610387084>

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, Aging, and Cognitive Control: Evidence From the Simon Task. *Psychology and Aging*, 19(2), 290–303. <https://doi.org/10.1037/0882-7974.19.2.290>

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism (Cambridge, England)*, 13(4), 525–531. <https://doi.org/10.1017/S1366728909990423>

Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: evidence from the dimensional change card sort task. *Developmental Science*, 7(3), 325–339. <https://doi.org/10.1111/j.1467-7687.2004.00351.x>

Bishop, D. (2003). Test for Reception of Grammar (TROG-2) | Pearson Assessment. Retrieved from <http://www.pearsonclinical.co.uk/Psychology/ChildCognitionNeuropsychologyandLanguage/>

ChildLanguage/TestforReceptionofGrammar(TROG-2)/TestforReceptionofGrammar(TROG-2).aspx

Bohnacker, U., Lindgren, J., & Öztekin, B. (2016). Turkish-and German-speaking bilingual 4-to-6-year-olds living in Sweden: Effects of age, SES and home language input on vocabulary production. Retrieved from <https://ir.canterbury.ac.nz/handle/10092/12907>

Bokus, B. (1978). Effect of adult-shared vs. nonshared perception of a picture on its description by the three-year-old. *Polish Psychological Bulletin*, 9(4), 239–243.

Bokus, B. (2000). *Światy fabuły w narracji dziecięcej*. Warszawa: Energia.

Bokus, B. (2004). Inter-mind phenomena in child narrative discourse. *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)*, 14(4), 391–408. <https://doi.org/10.1075/prag.14.4.01bok>

Bokus, B. (2013). Pejzaż akcji i pejzaż świadomości w opowiadaniach dziecięcych (z badań nad kompetencją komunikacyjną sześciolatków). In B. Bokus & Z. Kloch (Eds.), *Teksty i interpretacje* (Vol. 17, pp. 298–308). Piaseczno: Studio LEXEM.

Bonifacci, P., Barbieri, M., Tomassini, M., & Roch, M. (2017). In few words: linguistic gap but adequate narrative structure in preschool bilingual children. *Journal of Child Language*, 1–28. <https://doi.org/10.1017/S0305000917000149>

Booth, J. R., & Hall, W. S. (1995). Development of the understanding of the polysemous meanings of the mental-state verb know. *Cognitive Development*, 10(4), 529–549. [https://doi.org/10.1016/0885-2014\(95\)90025-X](https://doi.org/10.1016/0885-2014(95)90025-X)

Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology*, 18(6), 906–921. <https://doi.org/10.1037/0012-1649.18.6.906>

Bretherton, I., McNew, S., & Beeghly-Smith, M. (1981). Early person knowledge as expressed in gestural and verbal communication: When do infants acquire a ‘theory of mind’. In M. E. Lamb & L. R. Sherrod (Eds.), *Infant Social Cognition: Empirical and Theoretical Considerations* (pp. 333–373). Psychology Press.

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, 8(6), 535–543. <https://doi.org/10.1111/j.1467-7687.2005.00445.x>

Brown, J. R., Donelan-McCall, N., & Dunn, J. (1996). Why Talk about Mental States? The Significance of Children’s Conversations with Friends, Siblings, and Mothers. *Child Development*, 67(3), 836–849. <https://doi.org/10.1111/j.1467-8624.1996.tb01767.x>

Brown, J. R., & Dunn, J. (1992). Talk with Your Mother or Your Sibling? Developmental Changes in Early Family Conversations about Feelings. *Child Development*, 63(2), 336–349. <https://doi.org/10.1111/j.1467-8624.1992.tb01631.x>

Bruner, J. (1986). *Actual Minds, Possible Worlds*. Cambridge: Harvard University Press.

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, 112(2), 337–342. <https://doi.org/10.1016/j.cognition.2009.05.006>

Buttelmann, D., Over, H., Carpenter, M., & Tomasello, M. (2014). Eighteen-month-olds understand false beliefs in an unexpected-contents task. *Journal of Experimental Child Psychology*, 119, 120–126. <https://doi.org/10.1016/j.jecp.2013.10.002>

Callaghan, T., Rochat, P., Lillard, A., Claux, M. L., Odden, H., Itakura, S., ... Singh, S. (2005). Synchrony in the Onset of Mental-State Reasoning Evidence From Five Cultures. *Psychological Science*, 16(5), 378–384. <https://doi.org/10.1111/j.0956-7976.2005.01544.x>

Capps, L., Losh, M., & Thurber, C. (2000). 'The Frog Ate the Bug and Made his Mouth Sad': Narrative Competence in Children with Autism. *Journal of Abnormal Child Psychology*, 28(2), 193–204. <https://doi.org/10.1023/A:1005126915631>

Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science*, 11(2), 282–298. <https://doi.org/10.1111/j.1467-7687.2008.00675.x>

Carlson, S. M., & Moses, L. J. (2001). Individual Differences in Inhibitory Control and Children's Theory of Mind. *Child Development*, 72(4), 1032–1053. <https://doi.org/10.1111/1467-8624.00333>

Carlson, S. M., Moses, L. J., & Hix, H. R. (1998). The Role of Inhibitory Processes in Young Children's Difficulties with Deception and False Belief. *Child Development*, 69(3), 672–691. <https://doi.org/10.2307/1132197>

Carpendale, J. I. M., & Lewis, C. (2004). Constructing an understanding of mind: The development of children's social understanding within social interaction. *Behavioral and Brain Sciences*, 27(1). <https://doi.org/10.1017/S0140525X04000032>

Carpendale, J., & Lewis, C. (2006). *How Children Develop Social Understanding*. Wiley.

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social Cognition, Joint Attention, and Communicative Competence from 9 to 15 Months of Age. *Monographs of the Society for Research in Child Development*, 63(4), i. <https://doi.org/10.2307/1166214>

Chapman, M. (1991). The epistemic triangle: Operative and communicative components of cognitive competence. In M. Chandler & M. Chapman (Eds.), *Criteria for*

Competence: Controversies in the Conceptualization and Assessment of Children's Abilities (pp. 209–228). Hillsdale, NJ: Lawrence Erlbaum Associates.

Charman, T., & Shmueli-Goetz, Y. (1998). The relationship between theory of mind, language and narrative discourse: An experimental study. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, 17(2), 245–271.

Clark, E. V. (1993). *The Lexicon in Acquisition*. Cambridge: Cambridge University Press. Retrieved from <http://ebooks.cambridge.org/ref/id/CBO9780511554377>

Clark, E. V. (2016). *First language acquisition* (Third edition). Cambridge, United Kingdom: Cambridge University Press.

Clark, E. V. (2017). Lexical Acquisition and the Structure of the Mental Lexicon. *Oxford Research Encyclopedia of Linguistics*. <https://doi.org/10.1093/acrefore/9780199384655.013.96>

Clarke-Stewart, K. A., & Beck, R. J. (1999). Maternal Scaffolding and Children's Narrative Retelling of a Movie Story. *Early Childhood Research Quarterly*, 14(3), 409–434.

Cummins, J. (1979). Cognitive/Academic Language Proficiency, Linguistic Interdependence, the Optimum Age Question and Some Other Matters. *Working Papers on Bilingualism Toronto*, (19), 197–202.

De Houwer, A. (2009). *Bilingual first language acquisition*. Bristol, UK ; Buffalo, NY: Multilingual Matters.

De Houwer, A. (2014). The absolute frequency of maternal input to bilingual and monolingual children: A first comparison. In T. Grüter & J. Paradis (Eds.), *Trends in Language Acquisition Research* (Vol. 13, pp. 37–58). Amsterdam: John Benjamins Publishing Company. Retrieved from <https://benjamins.com/catalog/tilar.13.03deh>

De Houwer, A., Bornstein, M. H., & Putnick, D. L. (2014). A bilingual–monolingual comparison of young children’s vocabulary size: Evidence from comprehension and production. *Applied Psycholinguistics*, 35(6), 1189–1211. <https://doi.org/10.1017/S0142716412000744>

de Villiers, J. (2000). Language and theory of mind: what are the developmental relationships? . In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other Minds. Perspectives from Developmental Cognitive Neuroscience* (pp. 83–123). Oxford: Blackwell. Retrieved from <http://home.edo.uni-dortmund.de/hoffmann/Biblios/Erwerb.html>

de Villiers, J. (2007). The Interface of Language and Theory of Mind. *Lingua. International Review of General Linguistics. Revue Internationale de Linguistique Generale*, 117(11), 1858–1878. <https://doi.org/10.1016/j.lingua.2006.11.006>

Doherty, M. J. (2000). Children’s understanding of homonymy: metalinguistic awareness and false belief. *Journal of Child Language*, 27(2), 367–392.

Doherty, M., & Perner, J. (1998). Metalinguistic awareness and theory of mind: Just two words for the same thing? *Cognitive Development*, 13(3), 279–305. [https://doi.org/10.1016/S0885-2014\(98\)90012-0](https://doi.org/10.1016/S0885-2014(98)90012-0)

Dunn, J., Bretherton, I., & Munn, P. (1987). Conversations about feeling states between mothers and their young children. *Developmental Psychology*, 23(1), 132–139. <https://doi.org/10.1037/0012-1649.23.1.132>

Dunn, J., & Brophy, M. (2005). Communication, Relationships, and Individual Differences in Children’s Understanding of Mind. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 50–69). New York, NY, US: Oxford University Press.

Dunn, J., Brown, J., & Beardsall, L. (1991). Family talk about feeling states and children's later understanding of others' emotions. *Developmental Psychology*, 27(3), 448–455. <https://doi.org/10.1037/0012-1649.27.3.448>

Dunn, L. M., Dunn, J., & Styles, B. (2009). British Picture Vocabulary Scale (BPVS3). Retrieved from <https://www.g1-assessment.co.uk/products/british-picture-vocabulary-scale-bpvs3/>

Dyer, J. R., Shatz, M., & Wellman, H. M. (2000). Young children's storybooks as a source of mental state information. *Cognitive Development*, 15(1), 17–37.

Dyer-Seymour, J., Shatz, M., Wellman, H., & Saito, M. (2004). Mental state expressions in US and Japanese children's books. *International Journal of Behavioral Development*, 28(6), 546–552. <https://doi.org/10.1080/01650250444000261>

Esposito, A. G., Baker-Ward, L., & Mueller, S. (2013). Interference Suppression vs. Response Inhibition: An Explanation for the Absence of a Bilingual Advantage in Preschoolers' Stroop Task Performance. *Cognitive Development*, 28(4), 354–363. <https://doi.org/10.1016/j.cogdev.2013.09.002>

Fan, S. P., Liberman, Z., Keysar, B., & Kinzler, K. D. (2015). The Exposure Advantage: Early Exposure to a Multilingual Environment Promotes Effective Communication. *Psychological Science*, 26(7), 1090–1097.

Farhadian, M., Abdullah, R., Mansor, M., Redzuan, M., Gazanizadand, N., & Kumar, V. (2010). Theory of mind in bilingual and monolingual preschool children. *Journal of Psychology*, 1(1), 39–46.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., ... Stiles, J. (1994). Variability in Early Communicative Development. *Monographs of the Society for Research in Child Development*, 59(5), i-185. <https://doi.org/10.2307/1166093>

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D. J., Bates, E., Hartung, J. P., ... Reilly, J. S. (1993). *The MacArthur Communicative Development Inventories: User's Guide and Technical Manual*. Baltimore: Paul H. Brookes Publishing Co.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories: User's guide and technical manual (2nd ed.)*. Baltimore, MD: Brookes Publishing.

Feys, J. (2016). Nonparametric tests for the interaction in two-way factorial designs using R. *The R Journal*, 8(1), 367–378.

Fiestas, C. E., & Peña, E. D. (2004). Narrative Discourse in Bilingual Children: Language and Task Effects. *Language, Speech, and Hearing Services in Schools*, 35(2), 155–168. [https://doi.org/10.1044/0161-1461\(2004/016\)](https://doi.org/10.1044/0161-1461(2004/016))

Flavell, J. H., Green, F. L., Flavell, E. R., Watson, M. W., & Campione, J. C. (1986). Development of Knowledge about the Appearance-Reality Distinction. *Monographs of the Society for Research in Child Development*, 51(1), i-87. <https://doi.org/10.2307/1165866>

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, 44(3), 677–694.

Frank, R. E., & Hall, W. S. (1991). Polysemy and the acquisition of the cognitive internal state lexicon. *Journal of Psycholinguistic Research*, 20(4), 283–304.

Friesen, D. C., & Bialystok, E. (2012). Metalinguistic Ability in Bilingual Children: The Role of Executive Control. *Rivista Di Psicologica Applicata*, 12(3), 47–56.

Gagarina, N. (2016). Narratives of Russian–German preschool and primary school bilinguals: Rasskaz and Erzaehlung. *Applied Psycholinguistics*, 37(1), 91–122. <https://doi.org/10.1017/S0142716415000430>

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiūnienė, I., ...
Walters, J. (2012). *MAIN: Multilingual assessment instrument for narratives*. Zentrum für
Allgemeine Sprachwiss.

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balciuniene, I., ...
Walters, J. (2015). Assessment of narrative abilities in bilingual children. In S. Armon-Lotem,
J. de Jong, & N. Meir (Eds.), *Assessing Multilingual Children: Disentangling Bilingualism from
Language Impairment* (pp. 241–274). Amsterdam: Multilingual Matters.

Gagarina, N., Klop, D., Tsimpli, I. M., & Walters, J. (2016). Narrative abilities in
bilingual children. *Applied Psycholinguistics*, 37(1), 11–17.
<https://doi.org/10.1017/S0142716415000399>

Gagarina, N. V., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balciuniene, I., ...
Walters, J. (2012). MAIN: Multilingual Assessment Instrument for Narratives. Retrieved from
<http://publikationen.ub.uni-frankfurt.de/frontdoor/index/index/docId/34782>

Gleitman, L. (1990). The Structural Sources of Verb Meanings. *Language Acquisition*,
1(1), 3–55.

Goetz, P. J. (2003). The effects of bilingualism on theory of mind development.
Bilingualism: Language and Cognition, 6(1), 1–15.
<https://doi.org/10.1017/S1366728903001007>

Gopnik, A., & Astington, J. W. (1988). Children's Understanding of Representational
Change and Its Relation to the Understanding of False Belief and the Appearance-Reality
Distinction. *Child Development*, 59(1), 26–37. <https://doi.org/10.2307/1130386>

Gopnik, A., & Rosati, A. (2001). Duck or rabbit? Reversing ambiguous figures and
understanding ambiguous representations. *Developmental Science*, 4(2), 175–183.
<https://doi.org/10.1111/1467-7687.00163>

Gordon, K. R. (2016). High proficiency across two languages is related to better mental state reasoning for bilingual children. *Journal of Child Language*, 43(2), 407–424. <https://doi.org/10.1017/S0305000915000276>

Grazzani, I., & Ornaghi, V. (2012). How do use and comprehension of mental-state language relate to theory of mind in middle childhood? *Cognitive Development*, 27(2), 99–111. <https://doi.org/10.1016/j.cogdev.2012.03.002>

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1(2), 67–81. <https://doi.org/10.1017/S1366728998000133>

Grosjean, F. (1984). *Life with Two Languages: An Introduction to Bilingualism*. Cambridge, Mass.: Harvard University Press.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, 36(1), 3–15. [https://doi.org/10.1016/0093-934X\(89\)90048-5](https://doi.org/10.1016/0093-934X(89)90048-5)

Grosjean, F. (2010). *Bilingual: life and reality*. Cambridge, Mass: Harvard University Press.

Gut, A. (2016). Językowe i kulturowe uwarunkowania myślenia o umyśle, relacji umysł-ciało i atrybucji stanów mentalnych. In A. Brożek, A. Chybińska, M. Grygianiec, & M. Tkaczyk (Eds.), *Myśli o języku, nauce i wartościach* (pp. 380–394). Warszawa: Wydawnictwo Naukowe Semper.

Hale, C. M., & Tager-Flusberg, H. (2003). The Influence of Language on Theory of Mind: A Training Study. *Developmental Science*, 6(3), 346–359. <https://doi.org/10.1111/1467-7687.00289>

Hall, W. S., & Nagy, W. E. (1979). Theoretical issues in the investigation of words of internal report. *Center for the Study of Reading Technical Report; No. 146*. Retrieved from

http://www.ideals.illinois.edu/bitstream/handle/2142/17639/ctrstreadtechrepv01979i00146_opt.pdf?sequence=1

Haman, E., & Fronczyk, K. (2012). Obrazkowy Test Słownikowy – Rozumienie (OTSR). *Gdańsk: Pracownia Testów Psychologicznych I Pedagogicznych*.

Haman, E., & Smoczyńska, M. (2010). Zadanie Nazywania Obrazków. *Material Niepublikowany, Warszawa: Wydział Psychologii, Uniwersytet Warszawski*.

Haman, E., Wodniecka, Z., Marecka, M., Szewczyk, J., Białecka-Pikul, M., Otwinowska, A., ... Foryś-Nogała, M. (2017). How Does L1 and L2 Exposure Impact L1 Performance in Bilingual Children? Evidence from Polish-English Migrants to the United Kingdom. *Frontiers in Psychology, 8*. <https://doi.org/10.3389/fpsyg.2017.01444>

Hansen, P., Łuniewska, M., Simonsen, H. G., Haman, E., Mieszkowska, K., Kołak, J., & Wodniecka, Z. (2017). Picture-based vocabulary assessment versus parental questionnaires: A cross-linguistic study of bilingual assessment methods. *International Journal of Bilingualism, 1367006917733067*.

Happé, F. G. E. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders, 24*(2), 129–154. <https://doi.org/10.1007/BF02172093>

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children* (Vol. xxiii). Baltimore, MD, US: Paul H Brookes Publishing.

Hart, B., & Risley, T. R. (2003). The Early Catastrophe. The 30 Million Word Gap. *American Educator, 27*(1), 4–9.

Hoff, E. (2017). How bilingual development is the same as and different from monolingual development. *OLBI Working Papers*, 8. Retrieved from <https://uottawa.scholarsportal.info/ojs/index.php/ILOB-OLBI/article/view/2114>

Hoff, E., & Core, C. (2013). Input and language development in bilingually developing children. *Seminars in Speech and Language*, 34(4), 215–226. <https://doi.org/10.1055/s-0033-1353448>

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, 39(1), 1–27. <https://doi.org/10.1017/S0305000910000759>

Hoff, E., Rumiche, R., Burridge, A., Ribot, K. M., & Welsh, S. N. (2014). Expressive Vocabulary Development in Children from Bilingual and Monolingual Homes: A Longitudinal Study from Two to Four Years. *Early Childhood Research Quarterly*, 29(4), 433–444. <https://doi.org/10.1016/j.ecresq.2014.04.012>

Hoff, E., Welsh, S., Place, S., & Ribot, K. M. (2014). Properties of dual language input that shape bilingual development and properties of environments that shape dual language input. In T. Grüter & J. Paradis (Eds.), *Trends in Language Acquisition Research* (Vol. 13, pp. 119–140). Amsterdam: John Benjamins Publishing Company. Retrieved from <https://benjamins.com/catalog/tilar.13.07hof>

Howard, A. A., Mayeux, L., & Naigles, L. R. (2008). Conversational correlates of children's acquisition of mental verbs and a theory of mind. *First Language*, 28(4), 375–402. <https://doi.org/10.1177/0142723708091044>

Hughes, C., & Dunn, J. (1997). 'Pretend you didn't know': Preschoolers' talk about mental states in pretend play. *Cognitive Development*, 12(4), 477–497. [https://doi.org/10.1016/S0885-2014\(97\)90019-8](https://doi.org/10.1016/S0885-2014(97)90019-8)

Hughes, C., Lecce, S., & Wilson, C. (2007). 'Do you know what I want?' Preschoolers' talk about desires, thoughts and feelings in their conversations with sibs and friends. *Cognition & Emotion*, 21(2), 330–350. <https://doi.org/10.1080/02699930600551691>

Iluz-Cohen, P., & Walters, J. (2011). Telling stories in two languages: Narratives of bilingual preschool children with typical and impaired language. *Bilingualism: Language and Cognition*, 15(1), 58–74. <https://doi.org/10.1017/S1366728911000538>

Jenkins, J. M., & Astington, J. W. (1996). Cognitive factors and family structure associated with theory of mind development in young children. *Developmental Psychology*, 32(1), 70–78. <https://doi.org/10.1037/0012-1649.32.1.70>

Jenkins, J. M., Turrell, S. L., Kogushi, Y., Lollis, S., & Ross, H. S. (2003). A Longitudinal Investigation of the Dynamics of Mental State Talk in Families. *Child Development*, 74(3), 905–920. <https://doi.org/10.1111/1467-8624.00575>

Johnson, C. N., & Wellman, H. M. (1980). Children's Developing Understanding of Mental Verbs: Remember, Know, and Guess. *Child Development*, 51(4), 1095–1102. <https://doi.org/10.2307/1129549>

Kay-Raining Bird, E., Cleave, P. L., Curia, J., & Dunleavy, M. (2008). Parental Talk About Internal States to Their Child With Autism. *Focus on Autism and Other Developmental Disabilities*, 23(3), 166–175. <https://doi.org/10.1177/1088357608319530>

Kiebzak-Mandera, D., Otwinowska, A., Białecka-Pikul, M. (2012). Zadania narracyjne: narracja i renarracja. *Materiał Niepublikowany*.

Klassert, A., Gagarina, N., & Kauschke, C. (2014). Object and action naming in Russian- and German-speaking monolingual and bilingual children. *Bilingualism: Language and Cognition*, 17(1), 73–88. <https://doi.org/10.1017/S1366728913000096>

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>

Kovács, Á. M. (2009). Early bilingualism enhances mechanisms of false-belief reasoning. *Developmental Science*, *12*(1), 48–54. <https://doi.org/10.1111/j.1467-7687.2008.00742.x>

Kristen, S., Chiarella, S., Sodian, B., Aureli, T., Genco, M., & Poulin-Dubois, D. (2014). Crosslinguistic Developmental Consistency in the Composition of Toddlers' Internal State Vocabulary: Evidence from Four Languages. *Child Development Research*, *2014*, 1–8. <https://doi.org/10.1155/2014/575142>

Kroll, J. F., Bogulski, C. A., & McClain, R. (2012). Psycholinguistic perspectives on second language learning and bilingualism: The course and consequence of cross-language competition. *Linguistic Approaches to Bilingualism*, *2*(1), 1–24. <https://doi.org/10.1075/lab.2.1.01kro>

Kunnari, S., Välimaa, T., & Laukkanen-Nevala, P. (2016). Macrostructure in the narratives of monolingual Finnish and bilingual Finnish–Swedish children. *Applied Psycholinguistics*, *37*(1), 123–144. <https://doi.org/10.1017/S0142716415000442>

Kurcz, I. (2007). *Psychologiczne aspekty dwujęzyczności*. Gdańsk: Gdańskie Wydawnictwo Psychologiczne.

LaBounty, J., Wellman, H. M., Olson, S., Lagattuta, K., & Liu, D. (2008). Mothers' and Fathers' Use of Internal State Talk with their Young Children. *Social Development*, *17*(4), 757–775. <https://doi.org/10.1111/j.1467-9507.2007.00450.x>

Lanza, E. (1992). Can bilingual two-year-olds code-switch? *Journal of Child Language*, *19*(3), 633–658. <https://doi.org/10.1017/S0305000900011600>

Leseman, P. P. M. (2000). Bilingual Vocabulary Development of Turkish Preschoolers in the Netherlands. *Journal of Multilingual and Multicultural Development*, 21(2), 93–112. <https://doi.org/10.1080/01434630008666396>

Leslie, A. M. (1987). Pretense and representation: The origins of ‘ theory of mind.’ *Psychological Review*, 94(4), 412.

Leslie, A. M., German, T. P., & Polizzi, P. (2005). Belief-desire reasoning as a process of selection. *Cognitive Psychology*, 50(1), 45–85. <https://doi.org/10.1016/j.cogpsych.2004.06.002>

Levy, E., & Nelson, K. (1994). Words in discourse: a dialectical approach to the acquisition of meaning and use. *Journal of Child Language*, 21(2), 367–389. <https://doi.org/10.1017/S0305000900009314>

Lewis, C., Freeman, N. H., Hagestadt, C., & Douglas, H. (1994). Narrative access and production in preschoolers’ false belief reasoning. *Cognitive Development*, 9(4), 397–424. [https://doi.org/10.1016/0885-2014\(94\)90013-2](https://doi.org/10.1016/0885-2014(94)90013-2)

Lewis, C., & Osborne, A. (1990). Three-Year-Olds’ Problems with False Belief: Conceptual Deficit or Linguistic Artifact? *Child Development*, 61(5), 1514–1519. <https://doi.org/10.1111/j.1467-8624.1990.tb02879.x>

Liu, D., Wellman, H. M., Tardif, T., & Sabbagh, M. A. (2008). Theory of mind development in Chinese children: A meta-analysis of false-belief understanding across cultures and languages. *Developmental Psychology*, 44(2), 523–531. <https://doi.org/10.1037/0012-1649.44.2.523>

Lohmann, H., & Tomasello, M. (2003). The Role of Language in the Development of False Belief Understanding: A Training Study. *Child Development*, 74(4), 1130–1144.

Longobardi, E., Spataro, P., & Renna, M. (2014). Relationship between False Belief, Mental State Language, Metalinguistic Awareness and Social Abilities in School-age Children. *Procedia - Social and Behavioral Sciences*, *114*, 365–371. <https://doi.org/10.1016/j.sbspro.2013.12.713>

Losh, M., & Capps, L. (2003). Narrative Ability in High-Functioning Children with Autism or Asperger's Syndrome. *Journal of Autism and Developmental Disorders*, *33*(3), 239–251. <https://doi.org/10.1023/A:1024446215446>

Lumey, T., & Miller, A. (2004). *Leaps: Regression Subset Selection. R Package Version, 2*.

Mair, P., & Schoenbrodt, F. (2017). *WRS2: Wilcox robust estimation and testing. R package version (Version 0.9-2)*.

Mair, P., & Wilcox, R. (2016). *Robust Statistical Methods in R Using the WRS2 Package*. Technical report, Harvard University.

Martin-Rhee, M. M., & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: Language and Cognition*, *11*(1). <https://doi.org/10.1017/S1366728907003227>

Martucci, K. (2016). Shared storybook reading in the preschool setting and considerations for young children's theory of mind development. *Journal of Early Childhood Research*, *14*(1), 55–68. <https://doi.org/10.1177/1476718X14523750>

Maviş, I., Tunçer, M., & Gagarina, N. (2016). Macrostructure components in narrations of Turkish–German bilingual children. *Applied Psycholinguistics*, *37*(1), 69–89. <https://doi.org/10.1017/S0142716415000429>

Meins, E. (2013). Sensitive attunement to infants' internal states: operationalizing the construct of mind-mindedness. *Attachment & Human Development, 15*(5–6), 524–544. <https://doi.org/10.1080/14616734.2013.830388>

Meins, E., & Fernyhough, C. (1999). Linguistic acquisitional style and mentalising development: The role of maternal mind-mindedness. *Cognitive Development, 14*(3), 363–380.

Meins, E., Fernyhough, C., Arnott, B., Leekam, S. R., & de Rosnay, M. (2013). Mind-Mindedness and Theory of Mind: Mediating Roles of Language and Perspectival Symbolic Play. *Child Development, 84*(5), 1777–1790. <https://doi.org/10.1111/cdev.12061>

Meins, E., Fernyhough, C., Johnson, F., & Lidstone, J. (2006). Mind-mindedness in children: Individual differences in internal-state talk in middle childhood. *British Journal of Developmental Psychology, 24*(1), 181–196. <https://doi.org/10.1348/026151005X80174>

Meltzoff, A. N. (1999). Origins of theory of mind, cognition and communication. *Journal of Communication Disorders, 32*(4), 251–269.

Miękisz, A. (2016). Zasób słownictwa we wczesnej dwujęzyczności u dzieci przyswajających język polski na emigracji. Retrieved from <https://depotuw.ceon.pl/handle/item/2214>

Miękisz, A., Haman, E., Łuniewska, M., Kuś, K., O'Toole, C., & Katsos, N. (2017). The impact of a first-generation immigrant environment on the heritage language: productive vocabularies of Polish toddlers living in the UK and Ireland. *International Journal of Bilingual Education and Bilingualism, 20*(2), 183–200. <https://doi.org/10.1080/13670050.2016.1179259>

Mieszkowska, K., Haman, E., Białecka-Pikul, M., & Otwinowska, A. (2016). *METVOC-Polish: Polish adaptation of Metacognitive Vocabulary Test*. Unpublished material. Faculty of Psychology, University of Warsaw. This tool is an adaptation of a tool developed by

Astington, J. W., & Pelletier, J. (2004). Theory of mind and metacognitive vocabulary development in first and second languages. Unpublished manuscript, University of Toronto.

Mieszkowska, K., Łuniewska, M., Kołak, J., Kacprzak, A., Wodniecka, Z., & Haman, E. (2017). Home Language Will Not Take Care of Itself: Vocabulary Knowledge in Trilingual Children in the United Kingdom. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.01358>

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, 78(2), 622–646.

Miranda, A., Baixauli, I., & Colomer, C. (2013). Narrative writing competence and internal state terms of young adults clinically diagnosed with childhood attention deficit hyperactivity disorder. *Research in Developmental Disabilities*, 34(6), 1938–1950. <https://doi.org/10.1016/j.ridd.2013.03.014>

Montrul, S. A. (2008). *Incomplete Acquisition in Bilingualism: Re-examining the Age Factor* (Vol. 39). Amsterdam: John Benjamins Publishing Company. Retrieved from <http://www.jbe-platform.com/content/books/9789027290410>

Moore, C., Bryant, D., & Furrow, D. (1989). Mental Terms and the Development of Certainty. *Child Development*, 60(1), 167. <https://doi.org/10.2307/1131082>

Naigles, L. R. (2000). Manipulating the input: Studies in mental verb acquisition. In B. Landau, J. Sabini, J. Jonides, & E. L. Newport (Eds.), *Perception, cognition, and language: Essays in honor of Henry and Lila Gleitman* (pp. 245–274). Cambridge, MA, US: The MIT Press.

Nelson, K. (1998). *Language in Cognitive Development: The Emergence of the Mediated Mind*. Cambridge University Press.

Nelson, K. (2007). *Young Minds in Social Worlds: Experience, Meaning, and Memory*. Cambridge, Mass.: Harvard University Press.

Nelson, K., Plesa Skwerer, D., Goldman, S., Henseler, S., Presler, N., & Fried Walkenfeld, F. (2003). Entering a Community of Minds: An Experiential Approach to 'Theory of Mind'. *Human Development*, 46(1), 24–46. <https://doi.org/10.1159/000067779>

Nicolopoulou, A. (2002). Peer-group Culture and Narrative Development. In S. Blum-Kulka & C. E. Snow (Eds.), *Talking to Adults: The Contribution of Multiparty Discourse to Language Acquisition* (1 edition). Mahwah, N.J: Psychology Press.

Nicolopoulou, A., Cates, C. B., de Sá, A., & Ilgaz, H. (2014). Narrative performance, peer group culture, and narrative development in a preschool classroom. In A. Cekaite, S. Blum-Kulka, V. Grover, & E. Teubal (Eds.), *Children's Peer Talk: Learning from Each Other* (pp. 42–62). New York: Cambridge University Press. Retrieved from https://www.academia.edu/7623903/Narrative_Performance_Peer_Group_Culture_and_Narrative_Development_in_a_Preschool_Classroom_2014_

Nicolopoulou, A., & Richner, E. S. (2007). From Actors to Agents to Persons: The Development of Character Representation in Young Children's Narratives. *Child Development*, 78(2), 412–429. <https://doi.org/10.1111/j.1467-8624.2007.01006.x>

Nielsen, M., & Dissanayake, C. (2000). An investigation of pretend play, mental state terms and false belief understanding: In search of a metarepresentational link. *British Journal of Developmental Psychology*, 18(4), 609–624.

Norbury, C. F., & Bishop, D. V. M. (2003). Narrative skills of children with communication impairments. *International Journal of Language & Communication Disorders*, 38(3), 287–313. <https://doi.org/10.1080/136820310000108133>

Oller, D. K., Zurer Pearson, B., & Cobo-Lewis, A. B. (2007). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics*, 28(2).
<https://doi.org/10.1017/S0142716407070117>

Onishi, K. H., & Baillargeon, R. (2005). Do 15-Month-Old Infants Understand False Beliefs? *Science (New York, N.y.)*, 308(5719), 255–258.
<https://doi.org/10.1126/science.1107621>

Otwinowska, A., Mieszkowska, K., Białeczka-Pikul, M., Opacki, M., & Haman, E. (2018). Retelling a model story improves the narratives of Polish-English bilingual children. *International Journal of Bilingual Education and Bilingualism*, 1–25.
<https://doi.org/10.1080/13670050.2018.1434124>

Paradis, J., Genesee, F., & Crago, M. B. (Eds.). (2011). *Dual language development and disorders: a handbook on bilingualism and second language learning* (2nd ed). Baltimore, Md: Paul H. Brookes Pub. Co.

Pascual, B., Aguado, G., Sotillo, M., & Masdeu, J. C. (2008). Acquisition of mental state language in Spanish children: a longitudinal study of the relationship between the production of mental verbs and linguistic development. *Developmental Science*, 11(4), 454–466. <https://doi.org/10.1111/j.1467-7687.2008.00691.x>

Pearson, B. Z. (2002). Language and mind in the stories of bilingual children. In L. Verhoeven & S. Strömquist (Eds.), *Narrative development in a multilingual context* (pp. 135–174). Amsterdam: John Benjamins.

Pearson, B. Z., Fernandez, S. C., Lewedeg, V., & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*, 18(1), 41–58.
<https://doi.org/10.1017/S0142716400009863>

Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1993). Lexical Development in Bilingual Infants and Toddlers: Comparison to Monolingual Norms. *Language Learning*, 43(1), 93–120. <https://doi.org/10.1111/j.1467-1770.1993.tb00174.x>

Piaget, J. (1923). *The Language and Thought of the Child* (1 edition). London: Routledge & Kegan Paul.

Pinto, G., Tarchi, C., & Bigozzi, L. (2016). Peer Interaction Does Not Always Improve Children's Mental State Talk Production in Oral Narratives. A Study in 6- to 10-Year-Old Italian Children. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.01669>

Raven, J. C. (2003). *Test Matryc Ravena – wersja kolorowa. (T. Szustrowa i A. Jaworowska, Trans.)*. Warszawa: Pracownia Testów Psychologicznych PTP.

Reese, E., Suggate, S., Long, J., & Schaughency, E. (2010). Children's oral narrative and reading skills in the first 3 years of reading instruction. *Reading and Writing*, 23(6), 627–644. <https://doi.org/10.1007/s11145-009-9175-9>

Roy, B. C., Frank, M. C., DeCamp, P., Miller, M., & Roy, D. (2015). Predicting the birth of a spoken word. *Proceedings of the National Academy of Sciences*, 112(41), 12663–12668. <https://doi.org/10.1073/pnas.1419773112>

Rudek, D. J., & Haden, C. A. (2005). Mothers' and Preschoolers' Mental State Language During Reminiscing Over Time. *Merrill-Palmer Quarterly*, 51(4), 523–549. <https://doi.org/10.1353/mpq.2005.0026>

Ruffman, T., Perner, J., Naito, M., Parkin, L., & Clements, W. A. (1998). Older (but not younger) siblings facilitate false belief understanding. *Developmental Psychology*, 34(1), 161–174. <https://doi.org/10.1037//0012-1649.34.1.161>

Ruffman, T., Slade, L., & Crowe, E. (2002). The Relation between Children's and Mothers' Mental State Language and Theory-of-Mind Understanding. *Child Development*, 73(3), 734–751. <https://doi.org/10.1111/1467-8624.00435>

Rumpf, A.-L., Kamp-Becker, I., Becker, K., & Kauschke, C. (2012). Narrative competence and internal state language of children with Asperger Syndrome and ADHD. *Research in Developmental Disabilities*, 33(5), 1395–1407. <https://doi.org/10.1016/j.ridd.2012.03.007>

Sabbagh, M. A., & Callanan, M. A. (1998). Metarepresentation in action: 3-, 4-, and 5-year-olds' developing theories of mind in parent-child conversations. *Developmental Psychology*, 34(3), 491.

Shatz, M., Wellman, H. M., & Silber, S. (1983a). The acquisition of mental verbs: A systematic investigation of the first reference to mental state. *Cognition*, 14(3), 301–321. [https://doi.org/10.1016/0010-0277\(83\)90008-2](https://doi.org/10.1016/0010-0277(83)90008-2)

Shatz, M., Wellman, H. M., & Silber, S. (1983b). The acquisition of mental verbs: A systematic investigation of the first reference to mental state. *Cognition*, 14(3), 301–321. [https://doi.org/10.1016/0010-0277\(83\)90008-2](https://doi.org/10.1016/0010-0277(83)90008-2)

Shiro, M., Hoff, E., Ribot, K. M., & Shanks, K. (2017). *Internal state lexicon in 30-month-old bilingual spontaneous speech*. Presented at the 14th International Congress for the Study of Child Language, Lyon, France.

Shiro, M., Hoff, E., & Shanks, K. (in press). El uso de palabras que remiten a emociones, cognición y volición en niños bilingües de 30 meses. In C. Rojas & V. Oropeza (Eds.), *Diferencias individuales en la adquisición del lenguaje. Factores lingüísticos, cognitivos, socioambientales*. Mexico: UNAM.

Slaughter, V., Peterson, C. C., & Carpenter, M. (2009). Maternal mental state talk and infants' early gestural communication. *Journal of Child Language*, 36(5), 1053–1074. <https://doi.org/10.1017/S0305000908009306>

Slaughter, V., Peterson, C. C., & Mackintosh, E. (2007). Mind What Mother Says: Narrative Input and Theory of Mind in Typical Children and Those on the Autism Spectrum. *Child Development*, 78(3), 839–858. <https://doi.org/10.1111/j.1467-8624.2007.01036.x>

Smoczyńska, M. (2008). *The Test for the Reception of Grammar – Polish translation*. Jagiellonian University, Cracow, Poland.

Spencer, T. D., & Slocum, T. A. (2010). The Effect of a Narrative Intervention on Story Retelling and Personal Story Generation Skills of Preschoolers With Risk Factors and Narrative Language Delays. *Journal of Early Intervention*, 32(3), 178–199. <https://doi.org/10.1177/1053815110379124>

Stein, N. L., & Glenn, C. G. (1975). An Analysis of Story Comprehension in Elementary School Children: A Test of a Schema. Retrieved from <http://eric.ed.gov/?id=ED121474>

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of Beliefs by 13-Month-Old Infants. *Psychological Science*, 18(7), 580–586. <https://doi.org/10.1111/j.1467-9280.2007.01943.x>

Symons, D. K. (2004). Mental state discourse, theory of mind, and the internalization of self–other understanding. *Developmental Review*, 24(2), 159–188. <https://doi.org/10.1016/j.dr.2004.03.001>

Symons, D. K., Fossum, K.-L. M., & Collins, T. B. K. (2006). A Longitudinal Study of Belief and Desire State Discourse During Mother–Child Play and Later False Belief Understanding. *Social Development*, 15(4), 676–692. <https://doi.org/10.1111/j.1467-9507.2006.00364.x>

Symons, D. K., Peterson, C. C., Slaughter, V., Roche, J., & Doyle, E. (2005). Theory of mind and mental state discourse during book reading and story-telling tasks. *British Journal of Developmental Psychology*, 23(1), 81–102. <https://doi.org/10.1348/026151004X21080>

Tager-Flusberg, H. (1992). Autistic children's talk about psychological state: Deficits in the early acquisition of a theory... *Child Development*, 63(1), 161. <https://doi.org/10.1111/1467-8624.ep9203091734>

Tardif, T., & Wellman, H. M. (2000). Acquisition of mental state language in Mandarin- and Cantonese-speaking children. *Developmental Psychology*, 36(1), 25–43. <https://doi.org/10.1037/0012-1649.36.1.25>

Taumoepeau, M., & Reese, E. (2013). Maternal reminiscing, elaborative talk, and children's theory of mind: An intervention study. *First Language*, 33(4), 388–410. <https://doi.org/10.1177/0142723713493347>

Taumoepeau, M., & Ruffman, T. (2006). Mother and infant talk about mental states relates to desire language and emotion understanding. *Child Development*, 77(2), 465–481.

Taumoepeau, M., & Ruffman, T. (2008). Stepping Stones to Others' Minds: Maternal Talk Relates to Child Mental State Language and Emotion Understanding at 15, 24, and 33 Months. *Child Development*, 79(2), 284–302. <https://doi.org/10.1111/j.1467-8624.2007.01126.x>

Thompson, R. A. (2006). Conversation and Developing Understanding: Introduction to the Special Issue. *Merrill - Palmer Quarterly; Detroit*, 52(1), 1–16.

Thordardottir, Elin (2011). The relationship between bilingual exposure and vocabulary development. *International Journal of Bilingualism*, 15(4), 426–445. <https://doi.org/10.1177/1367006911403202>

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. Dunham (Eds.), *Joint Attention: Its Origins and Role in Development* (pp. 85–101). Hillsdale, NJ: LEA.

Uccelli, P., & Páez, M. M. (2007). Narrative and Vocabulary Development of Bilingual Children From Kindergarten to First Grade: Developmental Changes and Associations Among English and Spanish Skills. *Language, Speech, and Hearing Services in Schools*, 38(3), 225–236. [https://doi.org/10.1044/0161-1461\(2007/024\)](https://doi.org/10.1044/0161-1461(2007/024))

Wellman, H. M., Fang, F., Liu, D., Zhu, L., & Liu, G. (2006). Scaling of theory-of-mind understandings in Chinese children. *Psychological Science*, 17(12), 1075–1081.

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, 75(2), 523–541.

Williams, K. (2007). Expressive Vocabulary Test, Second Edition. Retrieved from <http://www.pearsonclinical.com/language/products/100000416/expressive-vocabulary-test-second-edition-evt-2.html>

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103–128.

Wodniecka, Z., Mieszkowska, K., Durlak, J., & Haman, E. (2018). Kiedy $1 + 1 \neq 2$, czyli jak dwujęzyczni przyswajają i przetwarzają język(i). In E. Czaplewska (Ed.), *Logopedia międzykulturowa* (pp. 91–131). Gdańsk: Harmonia Universalis.

Zevenbergen, A. A., Haman, E., & Olszańska, A. A. (2012). Middle-Class Polish and American Mothers' Beliefs Regarding Parent-Child Narratives. *Journal of Cross-Cultural Psychology*, 43(6), 979–998. <https://doi.org/10.1177/0022022111416005>

Zevenbergen, A. A., Haman, E., & Zevenbergen, J. (n.d.). 'Do You Remember Going to the Beach?': References to Internal States in Polish and American Mother-Preschooler Shared Narratives. *In Review*.

Zevenbergen, A. A., Holmes, A., Haman, E., Whiteford, N., & Thielges, S. (2016). Variability in mothers' support for preschoolers' contributions to co-constructed narratives as a function of child age. *First Language*, 36(6), 601–616.
<https://doi.org/10.1177/0142723716673955>

Appendix A:

Below is the list of 46 studies published online and accessible to the author with keywords that included one of the following: mental state language, mental state talk, internal state language, metacognitive terms, mental terms, mental verbs, emotion state language, feeling states.

Authors	Year of publication	Included in the review list?	Reason for exclusion	ISL generated by:	Investigated population	ISL gathered via:
Hall & Nagy	1979	included	-	-	-	theoretical review
Bretherton, McNew, & Beeghly-Smith	1981	included	-	children	TD	internal state CDI
Shatz, Wellman, Silber	1983	included	-	children	TD	speech transcripts
Dunn, Bretherton, & Munn	1987	included	-	mother-children triads	TD	speech transcripts
Dyer, Shatz & Wellman	2000	included	-	books	-	book transcripts
Nielsen and Dissanayake	2000	included	-	parent-child dyad	TD	speech transcripts
Ruffman, Slade & Crowe	2002	included	-	mother-child dyads	TD	speech transcripts
Jenkins, Turrell, Kogushi, Lollis & Ross	2003	included	-	parent-child dyad	TD	speech transcripts
Dyer, Shatz, Wellman & Saito	2004	included	-	books	-	book transcripts
LaBounty, Wellman, Olson, Lagattuta, & Liu	2008	included	-	parent-child dyad (both mothers and fathers)	TD	speech transcripts
Pascual, Aguado, Sotillo, & Masdeu	2008	included	-	children	TD	speech transcripts
Zevenbergen, Haman, Zevenbergen	in review	included	-	mother-child dyads	TD	speech transcripts
Johnson & Wellman	1980	<u>excluded</u>	only few terms investigated: remember, know, guess	children	TD	speech transcripts
Moore, Bryant & Furrow	1989	<u>excluded</u>	comprehension test with a forced choice format	children	TD	comprehension of ISL

Appendix A

Authors	Year of publication	Included in the review list?	Reason for exclusion	ISL generated by:	Investigated population	ISL gathered via:
Dunn, Brown & Beardsall	1991	<u>excluded</u>	investigated themes (e.g. positive emotion), not terms; used an overall number of themes (not terms)	parent-child interaction	TD	speech transcripts
Booth, Hall, Robison & Kim	1997	<u>excluded</u>	only one term investigated: know	children	TD	speech transcripts
Ligeza	1998	<u>excluded</u>	did not differentiate between ISL subclasses, used an overall number of terms	children	TD	speech transcripts
Baumgartner, Biagini & Devescovi	1998	<u>excluded</u>	no full list, only examples given	children	TD	speech transcripts
Hughes & Dunn	1998	<u>excluded</u>	no full list, only examples given	children	TD	speech transcripts
Tardif & Wellman	2000	<u>excluded</u>	only 5 Chinese terms investigated: want, want/think, know how, is able, know that	children	TD Mandarin and Cantonese children	speech transcripts
Capps, Losh, & Thurber	2000	<u>excluded</u>	no full list, only examples given	children	ASD (vs. developmental delays vs. TD)	speech transcripts
Symons, Peterson, Slaughter, Roche & Doyle	2005	<u>excluded</u>	did not differentiate between ISL subclasses, used an overall number of terms	children	TD	speech transcripts
Adrian, Clemente, Villanueva & Rieffe	2005	<u>excluded</u>	no full list, only examples given	parent-child interaction	TD	speech transcripts
Pelletier	2006	<u>excluded</u>	comprehension test with a forced choice format	children	TD L1 and L2 learners	comprehension of ISL
Antonietti, Liverta-Sempio, Marchetti & Astington	2006	<u>excluded</u>	comprehension test with a forced choice format	children	TD	comprehension of ISL

Appendix A

Authors	Year of publication	Included in the review list?	Reason for exclusion	ISL generated by:	Investigated population	ISL gathered via:
Taumoepeau & Ruffman	2006	<u>excluded</u>	used a list already in the review (Bretherton et al. 1981), as a CDI addition	children	TD	CDI
Meins, Fernyhough, Johnson, & Lidstone	2006	<u>excluded</u>	no full list, only examples given	children	TD	speech transcripts
Howard, Mayeux, & Naigles	2008	<u>excluded</u>	think, know, remember, guess and forget, were extracted	mothers	-	speech transcripts
Kay-Raining Bird, Cleave, Curia & Dunleavy	2008	<u>excluded</u>	ASD population	parent-child interaction	ASD	speech transcripts
Taumoepeau & Ruffman	2008	<u>excluded</u>	used a list already in the review (Bretherton et al. 1981), as a CDI addition	children	TD	CDI
Wang, Doan & Song	2010	<u>excluded</u>	no full list, only examples given	children	TD	speech transcripts
Rumpf, Kamp-Becker, Becker, & Kauschke	2012	<u>excluded</u>	no full list, only examples given	children	ASD vs. ADHD vs. TD	speech transcripts
Meins, Fernyhough, de Rosnay, Arnott, Leekam, & Turner	2012	<u>excluded</u>	no full list, only examples given	mothers	TD	speech transcripts
Osório, Meins, Martins, Martins & Soares	2012	<u>excluded</u>	no full list, only examples given	mother-child dyads	TD	speech transcripts
Grazzani & Ornaghi	2012	<u>excluded</u>	no full list, only examples given	children	TD	speech transcripts
Meins, Fernyhough, Arnott, Leekam & de Rosnay	2013	<u>excluded</u>	used a list already in the review (Bretherton et al. 1981), as a CDI addition	children	TD	CDI
Miranda, Baixauli, & Colomer	2013	<u>excluded</u>	no full list, only examples given	adults	ADHD (vs. non-ADHD)	written transcripts
Ornaghi & Grazzani	2013	<u>excluded</u>	no full list, only examples given	children	TD	speech transcripts

Appendix A

Authors	Year of publication	Included in the review list?	Reason for exclusion	ISL generated by:	Investigated population	ISL gathered via:
Kristen, Chiarella, Sodian, Aureli, Genco & Poulin-Dubois	2014	<u>excluded</u>	used a list already in the review (Bretherton et al. 1981), as a CDI addition	children	TD	CDI
Siller, Swanson, Serlin, & George	2014	<u>excluded</u>	no full list, only examples given	children	ASD	speech transcripts
Morgan, Meristo, Mann, Hjelmquist, Surian & Siegal	2014	<u>excluded</u>	no full list, only examples given	parents	deaf children	speech transcripts
Gamannossi & Pinto	2014	<u>excluded</u>	no full list, only examples given	children	TD	speech transcripts
Pinto, Tarchi & Bigozzi	2016	<u>excluded</u>	no full list, only examples given	children	TD	speech transcripts
Altman, Armon-Lotem, Fichman & Walters	2016	<u>excluded</u>	no full list, only examples given	children	TD bilinguals vs. SLI bilinguals	speech transcripts
Becker Razuri, Hiles Howard, Purvis & Cross	2017	<u>excluded</u>	only few terms investigated: want, think, and know	children	TD	speech transcripts
Ebert, Peterson, Slaughter & Weinert	2017	<u>excluded</u>	no full list, only examples given	parent-child interaction	TD	speech transcripts

Appendix A

Below is the list of words coded as internal state terms in the 12 studies included in the review.

Note: dv: desire/volition, mo: moral/obligation, p: perceptual, ph: physiological, t: traits, e: emotional, c: cognitive

	Hall & Nagy	Bretherton, McNew, & Beeghly-Smith	Shatz, Wellman, Silber	Dunn, Bretherton, & Munn	Dyer, Shatz & Wellman	Nielsen and Dissanayake	Ruffman, Slade & Crowe	Jenkins, Turrell, Kogushi, Lollis & Ross	Dyer, Shatz, Wellman & Saito	LaBounty, Wellman, Olson, Lagattuta, & Liu	Pascual, Aguado, Sotillo, & Masdeu	Zevenbergen, Haman, Zevenbergen	Mieszkowska
year of publication	1979	1981	1983	1987	2000	2000	2002	2003	2004	2008	2008	in review	current thesis
ISL generated by	-	children	children	mother-child triads	books	parent-child dyad	mother-child dyads	parent-child dyad	books	parent-child dyad	children	mother-child dyads	children
ISL generated in	-	speech: spontaneous	speech: spontaneous	speech: spontaneous	-	speech: free play	speech: describing a picture	speech: spontaneous	-	speech: picture-book reading	speech: structured play	speech: personal narratives	speech: narratives based on pictures
ISL gathered via	theoretical review	internal state CDI	speech transcripts	speech transcripts	book transcripts		speech transcripts	speech transcripts	book transcripts	speech transcripts	speech transcripts	speech transcripts	speech transcripts
population	-	TD	TD	TD	-	TD	TD	TD	-	TD	TD	TD	TD, bilingual, monolingual
participant age (in yrs)	-	1-3	2-4	1-2	-	3-4.5	3(T1), 3.5(T2), 4(T3)	2 (T1), 4 (T2)	-	3.5(T1), 5(T2)	3 (T1), 3.5 (T2), 4 (T3), 4.5 (T4), 5 (T5)	3-5	4.5-7
participant sample	-	N=30	N=30	N=43 families	-	N=40	N=70	N=40	-	N=106	N=25	N=64	N=150
Subclasses considered in total	4	6	1	3	4	1	4	3	6	3	2	2	4
emotion	1	1		1	1		1	1	1	1		1	1
cognitive	1	1	1	1	1	1	1	1	1	1	1	1	1
cognitive emotion (where emotion has a crucial cognitive component)									1				
desire/volition/modality	1	1			1		1	1	1	1	1		1
moral/obligation/evaluation		1			1				1				
perceptual	1	1											1
physiological		1		1			1						
traits									1				

Appendix A

			Hall & Nagy	Bretherton, McNew, & Beeghly-Smith	Shatz, Wellman, Silber	Dunn, Bretherton, & Munn	Dyer, Shatz & Wellman	Nielsen and Dissanayake	Ruffman, Slade & Crowe	Jenkins, Turrell, Kogushi, Lollis & Ross	Dyer, Shatz, Wellman & Saito	LaBounty, Wellman, Olson, Lagattuta, & Liu	Pascual, Aguado, Sotillo, & Masdeu	Zevenberg en, Haman, Zevenberg en	Mieszkowska
year of publication			1979	1981	1983	1987	2000	2000	2002	2003	2004	2008	2008	in review	current thesis
Most frequently coded as ↓			full list included	full list included	full list included	examples given	examples given	full list included	examples given	full list included	examples given	full list included	full list included	full list included	full list included
<i>terms in total</i>			297	74	19	61	81	16	51	41	35	50	21	99	54
<i>emotion terms total</i>			150	22	0	30	37	0	20	25	10	30	0	44	29
<i>mental terms total</i>			85	11	19	11	32	16	13	12	10	16	17	55	17
<i>perception terms total</i>			40	13	0	0	0	0	0	0	0	0	0	0	8
aback	adj	e	e												
absent-minded	adj	c												c	
accept	verb	c	c												
ache	n/v	p	p		ph										
admit	verb	c				c									
adore	verb	e												e	
afraid	adj	e	e		e			e	e	e	e	e	e	e	e
agitate	verb	e	e												
agree	verb	c	c											c	
aim (to)	verb	dv	dv												
alarm	n/v	e	e												
all right	adverb	e		e											
amaze	verb	e	e			e						e			
amusing	adj	e	e												
anger	noun	e	e											e	
angry	adj	e	e	e	e			c	e	e	e	e			e
annoy, -ed	verb	e	e		e			e						e	e
anticipate	verb	c	c												
anticipation	noun	c	c												
anxious	adj	e	e			e									
appall-ed/-ing	adj	e	e												e
appear	verb	p	p												
appetite	noun	p	p												
applause	noun	e				e									

Appendix A

			Hall & Nagy	Bretherton, McNew, & Beeghly-Smith	Shatz, Wellman, Silber	Dunn, Bretherton, & Munn	Dyer, Shatz & Wellman	Nielsen and Dissanayake	Ruffman, Slade & Crowe	Jenkins, Turrell, Kogushi, Lollis & Ross	Dyer, Shatz, Wellman & Saito	LaBounty, Wellman, Olson, Lagattuta, & Liu	Pascual, Aguado, Sotillo, & Masdeu	Zevenberg en, Haman, Zevenberg en	Mieszkowska
year of publication			1979	1981	1983	1987	2000	2000	2002	2003	2004	2008	2008	in review	current thesis
appreciate	verb	e	e												
approve	verb	e	e												
ashamed	adj	e	e											e	
asleep	adj	ph		ph		c									
assume	verb	c	c												
astonish-ed/-ing	adj	e	e												
attitude	noun	e	e												
awake	adj	p	p	ph		c									
aware	adj	c	c												
bad (feeling)	adj	e		e/mo							mo			e	e
bear (I can't bear)	verb	e	e												
beat (tired)	adj	p	p												
belief	noun	c	c												
believe	verb	c	c		c			c		c			c	c	
beloved	adj	e												e	
best	adj	t									t				
bet	verb	c	c		c			c						c	
better	adverb	e		e		ph					mo				
blank (draw a blank)	noun	c	c												
blue	adj	e	e												
blush	verb	e					e								
bor-ed/-ing	adj	c	c			c	e		e					c	
bother	verb	e	e												
brave	adj	e									t			e	e
broken-hearted	adj	e					e								
burns (burns me up)	verb	e	e												
buy (I can't buy that)	verb	c	c												
calm	adj/noun	e										e		e	

Appendix A

			Hall & Nagy	Bretherton, McNew, & Beeghly-Smith	Shatz, Wellman, Silber	Dunn, Bretherton, & Munn	Dyer, Shatz & Wellman	Nielsen and Dissanayake	Ruffman, Slade & Crowe	Jenkins, Turrell, Kogushi, Lollis & Ross	Dyer, Shatz, Wellman & Saito	LaBounty, Wellman, Olson, Lagattuta, & Liu	Pascual, Aguado, Sotillo, & Masdeu	Zevenbergen, Haman, Zevenbergen	Mieszkowska
year of publication			1979	1981	1983	1987	2000	2000	2002	2003	2004	2008	2008	in review	current thesis
		dv/mo		dv/mo											
can	verb	dv/mo		dv/mo											c
care	n/v	e	e		e					dv		c		e	
careful	adj	c	c												
catch (= understand, perceive)	verb	c	c												
certain	adj	c	c												
change one's mind	verb	dv	dv												
cheat	verb	c												c	
cheer	verb	e	e		e						e				
cheerful	adj	e	e		e										e
choose	verb	dv	dv											c	
clever	adj	c				c									c
click	verb	c	c												
clucked her tongue	verb	e					e								
clutching her heart	verb	e					e								
cold (feeling cold)	adj	p	p	p		ph									p
come up with	verb	c												c	
comfort	verb	p	p												
comfortable	adj	p	p			ph									
comfy	adj	p	p												
conceivable	adj	c	c												
concentrate	verb	c	c												
concentration	noun	c	c												
concept	noun	c	c												
concern	noun	e	e												
concerned	adj	e	e									e			
conclusion	noun	c	c												
confidence	noun	c	c												

Appendix A

			Hall & Nagy	Bretherton, McNew, & Beeghly-Smith	Shatz, Wellman, Silber	Dunn, Bretherton, & Munn	Dyer, Shatz & Wellman	Nielsen and Dissanayake	Ruffman, Slade & Crowe	Jenkins, Turrell, Kogushi, Lollis & Ross	Dyer, Shatz, Wellman & Saito	LaBounty, Wellman, Olson, Lagattuta, & Liu	Pascual, Aguado, Sotillo, & Masdeu	Zevenberg en, Haman, Zevenberg en	Mieszkowska
year of publication			1979	1981	1983	1987	2000	2000	2002	2003	2004	2008	2008	in review	current thesis
confuse	verb	c	c										c		
conscious	adj	c	c												
consider	verb	c	c						c						
consideration	noun	c	c												
convince	verb	c	c				c								
could tell	verb	c					c								
count	verb	c												c	
crabby	adj	e										e			
cross	adj	e	e			e			e			e			e
cry	n/v	e		e		e			ph			e			e
curious	adj	c	c				c								
dawn (on someone)	verb	c	c												
dazed	adj	e	e												
dead	adj	c				c									
deceive	verb	c	c											c	c
decide	verb	c	dv										c	c	
dejected	adj	e	e												
deliberately	adverb	dv	dv												
delight	n/v	e	e												
delighted	adj	e	e											e	
delirious	adj	e	e												
depress-ed/-ing	adj	e	e												
desire	noun	dv	dv												
desperate	adj	e	e												
desperately	adverb	e	e												
desperation	noun	e	e												
determined	adj	dv	dv												
difficult	adj	e					e								
dirty	adj	e		e											

Appendix A

			Hall & Nagy	Bretherton, McNew, & Beeghly-Smith	Shatz, Wellman, Silber	Dunn, Bretherton, & Munn	Dyer, Shatz & Wellman	Nielsen and Dissanayake	Ruffman, Slade & Crowe	Jenkins, Turrell, Kogushi, Lollis & Ross	Dyer, Shatz, Wellman & Saito	LaBounty, Wellman, Olson, Lagattuta, & Liu	Pascual, Aguado, Sotillo, & Masdeu	Zevenbergen, Haman, Zevenbergen	Mieszkowska
year of publication			1979	1981	1983	1987	2000	2000	2002	2003	2004	2008	2008	in review	current thesis
disagree [i.e. not believe similarly]	verb	c												c	
disappoint	verb	e	e												
disappointed	adj	e			e				e			e			e
discouraged	adj	e												e	
discover [i.e. learn]	verb	c												c	
disgust	n/v	e	e							e					
disgusting	adj	e	e							e					
disillusioning	adj	e	e												
dismal	adj	e	e												
displease	verb	e	e												
distracting	adj	c	c												
distress	n/v	e												e	
disturb	verb	e	e		c										
disturbing	adj	e	e												
dizzy	adj	p	p		ph										
doubt	n/v	c	c												
down	adj	e	e											e	
dream	n/v	c	c	c	c		c	dv			c	c		c	
dull	adj	e				e									
embarrass	verb	e	e												
embarrass-ed/-ing	adj	e										e			
encourage	verb	e	e												
encouragement	noun	e	e												
enjoy	verb	e	e		e			e	e		e			e	
enjoyment	noun	e	e						e						
enthused	adj	e	e												
enthusiasm	noun	e	e												
envy	n/v	e	e											e	

Appendix A

			Hall & Nagy	Bretherton, McNew, & Beeghly-Smith	Shatz, Wellman, Silber	Dunn, Bretherton, & Munn	Dyer, Shatz & Wellman	Nielsen and Dissanayake	Ruffman, Slade & Crowe	Jenkins, Turrell, Kogushi, Lollis & Ross	Dyer, Shatz, Wellman & Saito	LaBounty, Wellman, Olson, Lagattuta, & Liu	Pascual, Aguado, Sotillo, & Masdeu	Zevenberg en, Haman, Zevenberg en	Mieszkowska
year of publication			1979	1981	1983	1987	2000	2000	2002	2003	2004	2008	2008	in review	current thesis
exasperat-ed/-ing	adj	e	e							e					
excite	verb	e												e	
excit-ed/-ing	adj	e	e						e	e		e			e
exclaim	verb	e					e								
exhaust-ed/-ing	adj	p	p												
expect	verb	c	c				c		c	c		c			
experience	n/v	c	c												
fake	adj/v	c												c	
familiar	adj	c					c								
familiarize	verb	c												c	
favourite	adj	e	e												
fear	n/v	e	e							e	e			e	
feel	verb	e	e	e			e		e		e	e	dv	e	p
feel (I feel it would be best)	verb	c	c												
feel (soft, warm)	verb	p		p		ph									
figure (out)	verb	c	c												
figure [i.e. believe, (I figured it would happen)]	verb	c	c		c			c						c	
find out [i.e. learn about]	verb	c	c				c							c	
fit (throw a fit)	verb	e	e												
focus [i.e. attention]	verb	c												c	
follow (I don't follow you = don't understand)	verb	c	c												
fond	adj	e	e												

Appendix A

			Hall & Nagy	Bretherton, McNew, & Beeghly-Smith	Shatz, Wellman, Silber	Dunn, Bretherton, & Munn	Dyer, Shatz & Wellman	Nielsen and Dissanayake	Ruffman, Slade & Crowe	Jenkins, Turrell, Kogushi, Lollis & Ross	Dyer, Shatz, Wellman & Saito	LaBounty, Wellman, Olson, Lagattuta, & Liu	Pascual, Aguado, Sotillo, & Masdeu	Zevenberg en, Haman, Zevenberg en	Mieszkowska
year of publication			1979	1981	1983	1987	2000	2000	2002	2003	2004	2008	2008	in review	current thesis
fool	n/v	c	c												
forget	verb	c	c	c	c			c	c	c	c	c	c	c	c
freezing	adj	p		p		ph									
friendly	adverb	e					e								e
fright	noun	e	e												
frighten	verb	e	e			e									
frightened	adj	e	e			e						e		e	e
frown	n/v	e					e								
frustrated	adj	e							e						
frustrat-ed/-ing	adj	e	e												
fumes	noun	e	e												
fun	noun	e				e			e	e	e	e			
funny	adj	e		e						e	t				
furious	adj	e	e											e	
fuss	n/v	e	e			e									
get (=understand)	verb	c	c												
get used to	verb	c	c												
giggle	verb	e					e		ph						
give up	verb	c					c								
glad	adj	e	e			e				e		e		e	e
gnash teeth	verb	e					e								
go with [i.e. decide]	verb	c												c	
good [mood]	adj	e/mo		e/mo										e	
greedy	adj	ph				ph									
grief	noun	e	e											e	
groan	verb	e					e								
grumpy	adj	e							e						e
guess	verb	c	c	c	c			c		c	c	c	c	c	
happy	adj	e	e	e		e			e	e	e	e	e	e	e
hard	adj	dv		dv											

Appendix A

			Hall & Nagy	Bretherton, McNew, & Beeghly-Smith	Shatz, Wellman, Silber	Dunn, Bretherton, & Munn	Dyer, Shatz & Wellman	Nielsen and Dissanayake	Ruffman, Slade & Crowe	Jenkins, Turrell, Kogushi, Lollis & Ross	Dyer, Shatz, Wellman & Saito	LaBounty, Wellman, Olson, Lagattuta, & Liu	Pascual, Aguado, Sotillo, & Masdeu	Zevenberg en, Haman, Zevenberg en	Mieszkowska
year of publication			1979	1981	1983	1987	2000	2000	2002	2003	2004	2008	2008	in review	current thesis
hate	verb	e	e			e						e			
hateful	adj	e	e												
have (half) a mind	verb	dv	dv												
have fun	verb	e		e											
have to	verb	dv/mo		dv/mo							mo				
having a great/good time	verb	e		e		e									
head (can't get it out of my head)	noun	c	c												
hear	verb	p	p	p											
heart (set one's heart on)	verb	dv	dv												
helpful	adj	t									t				
hilarious	adj	e	e												
homesick	adj	e				e									
hope	n/v	dv	c	dv	c	dv	c	dv	dv	dv	dv	dv	dv	c	
horrible	adj	mo				mo									
hot	adj	p		p		ph									
hug	n/v	eb		eb											
hungry	adj	ph	p	ph		ph			ph						p
hurt	adj	e	e	p		ph			ph	e					
hysterical	adj	e	e												
idea	noun	c	c		c				c					c	c
ignore	verb	c	c				c								
ill	adj	ph				ph									
imaginary	adj	c	c												
imagination	noun	c	c				c								
imagine	verb	c	c						c				c		

Appendix A

			Hall & Nagy	Bretherton, McNew, & Beeghly-Smith	Shatz, Wellman, Silber	Dunn, Bretherton, & Munn	Dyer, Shatz & Wellman	Nielsen and Dissanayake	Ruffman, Slade & Crowe	Jenkins, Turrell, Kogushi, Lollis & Ross	Dyer, Shatz, Wellman & Saito	LaBounty, Wellman, Olson, Lagattuta, & Liu	Pascual, Aguado, Sotillo, & Masdeu	Zevenbergen, Haman, Zevenbergen	Mieszkowska
year of publication			1979	1981	1983	1987	2000	2000	2002	2003	2004	2008	2008	in review	current thesis
impression (get the impression, under the impression)	noun	c	c												
infuriate	verb	e	e												
insist	verb	c				c									
inspire	verb	c				c									
intend	verb	dv	dv												
interest	verb	e	e												
interest-ed/-ing	adj	e	e					e						c	
ing	adj	e	e												
invent	verb	c	c												
irritate-ed/-ing	adj	e	e												
jealous	adj	e	e											e	e
joke	n/v	c				c									
joy	noun	e	e												
jump	n/v	e				e									
keen on	verb	dv							dv						
kick	n/v	e	e												
kind	adj	e									e				e
kiss	verb	e		e		ph					e				
know	verb	c	c	c	c		c	c	c	c		c	c	c	c
knowledge	noun	c	c												
laugh	verb	eb		eb		e			ph						
learn	verb	c	c				c						c	c	c
let	verb	mo		mo							mo				
lie	verb	e								e				c	
like	verb	e	e	e		e			dv			dv	c	e	e
listen	verb	p		p											
lonely	adj	e										e			e
lonesome	adj	e	e												
look	verb	p	p	p											p

Appendix A

			Hall & Nagy	Bretherton, McNew, & Beeghly-Smith	Shatz, Wellman, Silber	Dunn, Bretherton, & Munn	Dyer, Shatz & Wellman	Nielsen and Dissanayake	Ruffman, Slade & Crowe	Jenkins, Turrell, Kogushi, Lollis & Ross	Dyer, Shatz, Wellman & Saito	LaBounty, Wellman, Olson, Lagattuta, & Liu	Pascual, Aguado, Sotillo, & Masdeu	Zevenberg en, Haman, Zevenberg en	Mieszkowska
year of publication			1979	1981	1983	1987	2000	2000	2002	2003	2004	2008	2008	in review	current thesis
look forward [i.e. anticipate]	verb	e	e											c	
love	verb	e	e	e		e			dv	e	e	e		e	e
lucky	adj	e				e									
mad	adj	e	e	e		c				e				e	e
maddening	adj	e	e												
magic	adj/noun	c				c									
make fun	verb	c				c									
make sense	verb	c				c									
make up [i.e. create in one's mind]	verb	c	c			c								c	
make up one's mind	verb	dv	dv												
make-believe	noun	c	c												
may	verb	c/mo		c/mo							mo				
maybe	adverb	c		c											
mean (behavior)	adj	e										e		e	
mean (I mean)	verb	c	c	c	c	c	c	c	c	c	c	c	c	c	c
memorial	noun	c												c	
memory	noun	c				c								c	
merry	adj	e												e	
messy	adj	e		e											
might	verb	c									c				
mind (crossed my mind, come to mind)	noun	c	c												
mischief	noun	c				c									
miserable	adj	e	e		e										
miss	verb	e	e												

Appendix A

			Hall & Nagy	Bretherton, McNew, & Beeghly-Smith	Shatz, Wellman, Silber	Dunn, Bretherton, & Munn	Dyer, Shatz & Wellman	Nielsen and Dissanayake	Ruffman, Slade & Crowe	Jenkins, Turrell, Kogushi, Lollis & Ross	Dyer, Shatz, Wellman & Saito	LaBounty, Wellman, Olson, Lagattuta, & Liu	Pascual, Aguado, Sotillo, & Masdeu	Zevenberg en, Haman, Zevenberg en	Mieszkowska
year of publication			1979	1981	1983	1987	2000	2000	2002	2003	2004	2008	2008	in review	current thesis
mistake [i.e. believe erroneously]	noun	c												c	
misunderstand	verb	c	c												
mixed up	verb	e	e												
mood	noun	e	e												
moody	adj	e	e												
must	verb	mo		mo							mo				c
naughty	adj	mo		mo											e
nauseous	adj	p	p												c
need	verb	dv		dv		dv					dv				
nervous	adj	e	e			e									e
nice (behavior)	adj	e		e		e			e			e			e
notice	verb	p	p			c									p
nuisance	noun	e	e		e										
observation	noun	p	p												
observe	verb	p	p												
of course	adverb	c									c				
ok	adverb	e		e											
ought	verb	mo				mo									
passion	noun	e	e												
patient	adj	e				e									
pay attention	verb	c	c			c								c	
peek	verb	p	p												
peep	verb	p	p												
pick [i.e. choose]	verb	dv	dv											c	
pick up [i.e. learn]	verb	c	c											c	
piss(ed) (off)	adj	e	e												
pity	noun	e	e												
plan	n/v	dv	dv			c									

Appendix A

			Hall & Nagy	Bretherton, McNew, & Beeghly-Smith	Shatz, Wellman, Silber	Dunn, Bretherton, & Munn	Dyer, Shatz & Wellman	Nielsen and Dissanayake	Ruffman, Slade & Crowe	Jenkins, Turrell, Kogushi, Lollis & Ross	Dyer, Shatz, Wellman & Saito	LaBounty, Wellman, Olson, Lagattuta, & Liu	Pascual, Aguado, Sotillo, & Masdeu	Zevenberg en, Haman, Zevenberg en	Mieszkowska	
year of publication			1979	1981	1983	1987	2000	2000	2002	2003	2004	2008	2008	in review	current thesis	
play	verb	c					c									
plead	verb	e					e									
please	verb	e	e													
pleased	adj	e							e							
pleasure	noun	e	e													
ponder	verb	c	c													
positive	adj	c	c													
possessed	adj	e	e													
prefer	verb	e	e						dv							
preferences	noun	e	e													
pretend	verb	c	c	c	c			c		c		c	c	c		
pretend(adj)	adj	c			c					c						
probably	adverb	c					c									
proud	adj	e	e	e											e	
prove	verb	c					c									
purpose	noun	dv	dv													
put on [i.e. trick]	verb	c													c	
raging	adj	e	e													
ravenous	adj	p	p													
raving	adj	e	e													
reacting	adj	e	e													
reaction	noun	e	e													
real	adj	c		c												
realize	verb	c							dv				c			c
really	adverb	c									c					
reason (v)	n/v	c	c													
recall	verb	c	c												c	
recognize	verb	c	c				c					c				p
regret	verb	e	e													c
rejoice	verb	e	e													
relaxed	adj	p	p			c									e	

Appendix A

			Hall & Nagy	Bretherton, McNew, & Beeghly-Smith	Shatz, Wellman, Silber	Dunn, Bretherton, & Munn	Dyer, Shatz & Wellman	Nielsen and Dissanayake	Ruffman, Slade & Crowe	Jenkins, Turrell, Kogushi, Lollis & Ross	Dyer, Shatz, Wellman & Saito	LaBounty, Wellman, Olson, Lagattuta, & Liu	Pascual, Aguado, Sotillo, & Masdeu	Zevenberg en, Haman, Zevenberg en	Mieszkowska
year of publication			1979	1981	1983	1987	2000	2000	2002	2003	2004	2008	2008	in review	current thesis
reluctant	adj	c					c								
reluctantly	adverb	e	e												
remember	verb	c	c	c	c			c	c	c		c	c	c	
remind	verb	c							c					c	
reproachfully	adverb	e					e								
resent	verb	e	e												
resolution	noun	dv	dv												
resolve	verb	dv	dv												
respect	n/v	e	e												
restless	adj	p	p												
roll eyes	verb	e					e								
rotten	adj	mo					mo								
sad	adj	e	e	e		e			e	e		e		e	e
sadness	noun	e	e							e					
satisfied	adj	e	e			ph									
scare	verb	e	e			e									
scar-ed/-y	adj	e	e	e		e			e	e		e		e	
scream	n/v	e					e			e					
see	verb	p	p	p											p
see (= understand, find out)	verb	c	c											c	
seem	verb	c	c										c	c	
serious	adj	e	e												
seriously	adverb	e	e												
shame	n/v	e	e											e	
shock	n/v	e	e												
shock-ed/-ing	adj	e	e												c
shook	verb	e	e												
should	verb	mo		mo											
shy	adj	e										e		e	
sick	adj	e	e	ph		ph	e								

Appendix A

			Hall & Nagy	Bretherton, McNew, & Beeghly-Smith	Shatz, Wellman, Silber	Dunn, Bretherton, & Munn	Dyer, Shatz & Wellman	Nielsen and Dissanayake	Ruffman, Slade & Crowe	Jenkins, Turrell, Kogushi, Lollis & Ross	Dyer, Shatz, Wellman & Saito	LaBounty, Wellman, Olson, Lagattuta, & Liu	Pascual, Aguado, Sotillo, & Masdeu	Zevenberg en, Haman, Zevenberg en	Mieszkowska
year of publication			1979	1981	1983	1987	2000	2000	2002	2003	2004	2008	2008	in review	current thesis
sickening	adj	e	e												
sight	noun	p	p												
silly	adj	c									c				
sleep	n/v	ph		ph		c									
sleepy	adj	ph	p	ph		c			ph						
mind)	verb	c	c												
smart [i.e. relating intelligence]	adj	c												c	
smell	verb	p	p	p											
smile	verb	eb		eb					ph			e			
snapped	verb	e				e									
solemn	adj	e				e									
sore	adj	p	p		ph										
sorrows	noun	e	e												
sorry	adj	e	e											e	
soul	noun	e	e												
sound	adj	p	p												
spirit	noun	e	e												
stamp feet	verb	e				e									
stand (can't stand)	verb	e	e												
startle	verb	e				e									
startled	adj	e										e			
starv-ed/-ing	adj	p	p	ph											
stress	n/v	e												e	
stuffed	verb	p	p												
stun	verb	e	e												
suck breath	verb	e				e									
suffer	verb	e	e												
suppose	verb	c	c		c							c		c	
supposed to	adj	mo		mo		mo									

Appendix A

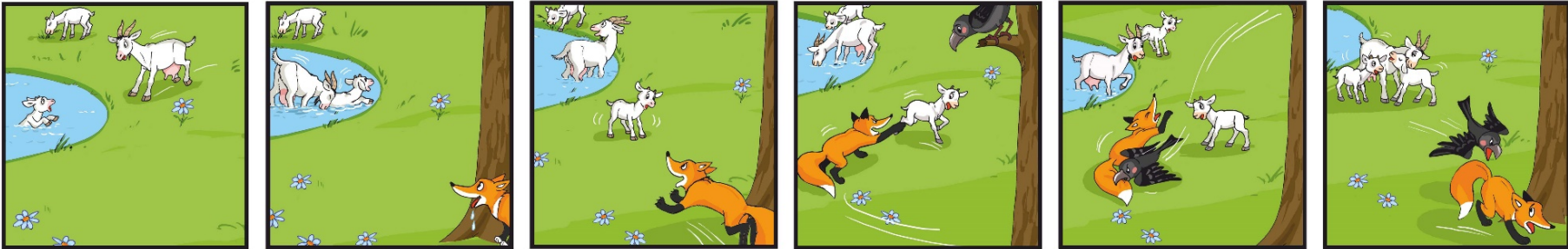
			Hall & Nagy	Bretherton, McNew, & Beeghly-Smith	Shatz, Wellman, Silber	Dunn, Bretherton, & Munn	Dyer, Shatz & Wellman	Nielsen and Dissanayake	Ruffman, Slade & Crowe	Jenkins, Turrell, Kogushi, Lollis & Ross	Dyer, Shatz, Wellman & Saito	LaBounty, Wellman, Olson, Lagattuta, & Liu	Pascual, Aguado, Sotillo, & Masdeu	Zevenbergen, Haman, Zevenbergen	Mieszkowska
year of publication			1979	1981	1983	1987	2000	2000	2002	2003	2004	2008	2008	in review	current thesis
sure	adj	c	c												c
surprise	n/v	e	e		e					e				e	
surpris-ed/-ing	adj	e	e	e	e				e	e	c	e		c	c
suspect	verb	c	c											c	
suspicion	noun	c	c												
sweet	adj	e													e
swallow (I can't swallow that)	verb	c	c												
sympathetic	adj	e	e												
sympathy	noun	e	e												
tantrum	noun	e	e												
taste	verb	p	p	p											
teach	verb	c	c				c							c	
tear	verb	e					e								
tempted	adj	e	e												
tense	adj	e	e												
terrible	adj	mo					mo								
terrified	adj	e												e	
terror	noun	e	e												
thankful	adj	e	e												
the bad	noun	mo					mo								
the good	noun	mo					mo								
think	verb	c	c	c	c			c	c	c		c	c	c	c
thirsty	adj	ph	p	ph		ph			ph						
thought	noun	c	c							c		c			
threaten	verb	e	e												
threatening	adj	e	e												
thrilled	adj	e	e												
tired	adj	ph	p	ph		c			ph						
tolerate	verb	e												e	
touch	n/v	p	p												

Appendix A

			Hall & Nagy	Bretherton, McNew, & Beeghly-Smith	Shatz, Wellman, Silber	Dunn, Bretherton, & Munn	Dyer, Shatz & Wellman	Nielsen and Dissanayake	Ruffman, Slade & Crowe	Jenkins, Turrell, Kogushi, Lollis & Ross	Dyer, Shatz, Wellman & Saito	LaBounty, Wellman, Olson, Lagattuta, & Liu	Pascual, Aguado, Sotillo, & Masdeu	Zevenberg en, Haman, Zevenberg en	Mieszkowska	
year of publication			1979	1981	1983	1987	2000	2000	2002	2003	2004	2008	2008	in review	current thesis	
track (keep track of, lose track of)	verb	c	c													
trick	n/v	c			c			c						c	c	
troubling	adj	e					e									
trust	n/v	e	e													
uncomfortable	adj	p	p							c						
understand	verb	c	c	c	c			c	c		c	c	c	c		
unhappy	adj	e	e			e			e			e				
unlearn	verb	c												c		
upset	adj	e	e			e			e	e		e		e	e	
view	n/v	p	p													
volunteer	verb	dv	dv													
wail	verb	e					e									
wake up	verb	ph		ph		c										
want	verb	dv	dv	dv			dv		dv	dv	dv	dv	dv	dv	c	c
warm	adj	p	p	p		ph										
watch	verb	p	p	p											p	
willing	adj	dv	dv													
willingly	adverb	dv	dv													
wish (v)	n/v	dv	dv	dv	c		dv	c	dv	dv	dv	dv	dv			
wonder	verb	c	c		c		c	c	c	c		c		c		
worried	adj	e	e			e			e						e	
worry	noun	e	e									c	c	e		
would like to	verb	dv	dv				dv				dv		dv			
yucky	adj	e		e												
yucky	adj	ph				ph										
yuk	adj	ph				ph										
zonked	adj	p	p													

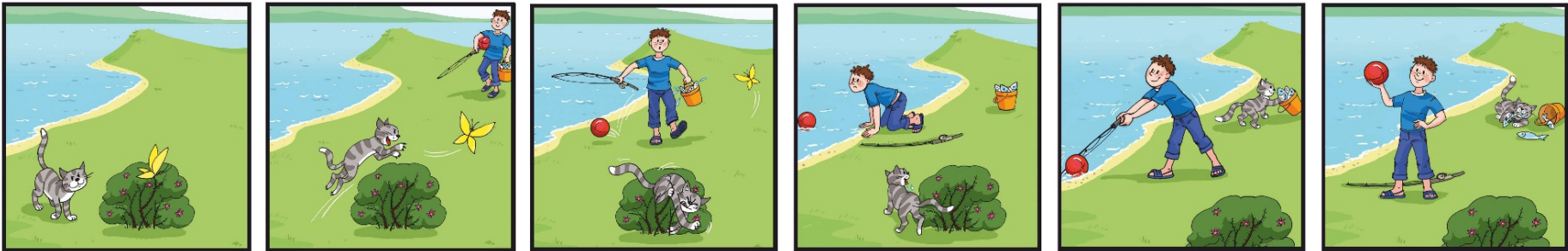
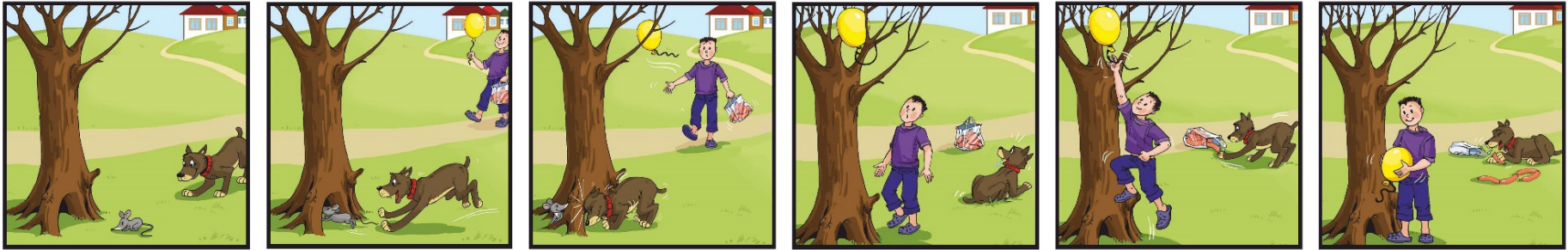
Appendix B:

Picture stories used for the Telling mode in the Multilingual Assessment Instrument for Narratives (MAIN: Birds, Goats)



Appendix B

Picture stories used for the Retelling mode in the Multilingual Assessment Instrument for Narratives (MAIN: Dog, Cat)



Appendix B

Story scripts for retelling mode in the MAIN narratives.

Cat

1.

One day there was a playful cat who saw a butterfly sitting on a bush.

2.

She jumped up because she wanted to catch it. Meanwhile, a cheerful boy was coming back from fishing with a bucket and a ball in his hands. He saw that the cat was chasing something.

3.

The boy shouted : “Cat, Cat, what are you chasing?” But the butterfly flew away and the cat fell into the bush and hurt herself. The boy was so surprised, that the ball fell out of his hand and into the water.

4.

The boy decided to get his ball back. Meanwhile, the cat noticed the boy’s bucket and thought: “What did the boy leave in the bucket?” She saw the fish and wanted to get it.

5.

But the boy began pulling his ball out of the water. At the same time, the cat reached for the fish that the boy left and thought, “That’s going to be delicious.”

6.

The boy was glad he got his ball again. But he did not notice that the cat was eating the delicious fish.

Dog

1.

One day there was a playful dog who saw a mouse sitting near a tree.

2.

He jumped up because he wanted to catch it. Meanwhile, a cheerful boy was coming back from shopping with a shopping bag and a balloon in his hands. He saw that the dog was chasing something.

3.

The boy shouted: “Dog, Dog, what are you chasing?” But the mouse ran away and the dog ran into the tree and hurt himself. The boy was so surprised, that the balloon flew out of his hand and onto the tree.

4.

The boy decided to get his balloon back. Meanwhile, the dog noticed the boy’s shopping bag and thought: “What did the boy leave in the bag?” He saw the sausages and wanted to get them.

5.

But the boy began pulling his balloon down from the tree. At the same time, the dog was reaching for sausage that the boy left and thought, “That’s going to be delicious!”.

6.

The boy was glad he got his balloon again. But he did not notice that the dog was eating the tasty sausage.

Appendix C:

The final lists of internal state terms produced by Polish monolingual and Polish-English bilingual children during the MAIN narrative task. The gray cells mark the items that were similar in meaning across the Polish and English lists. The cells that are merged (indicate items that have more than one synonym in one of the languages, e.g. “see” means the same as Polish “widzieć” and “zobaczyć”).

MENTAL TERMS:	
PL	ENG
	dissapointed
	forbid
	going to
	realize
	recognize
	seem
	trick
	learn
móc	can
móc	could
sprytny	clever
zdecydować	decide
zapomnieć	forget
mieć (dobry) pomysł	have the idea
wiedzieć	know
musieć	must
potrzebować	need
trzeba	
zdziwić się	shocked
zaskoczony	surprised
myśleć	think
próbować	try
chcieć	want
być pewnym	sure
być pewnym	confident
być ciekawym	
dać radę	
dowiedzieć się	
mieć ochotę	
postanowić	
pozwalać	
przygotować	
twierdzić	
udać się	
udawać	
umieć	
zastanawiać się	
znaleźć sposób	

PERCEPTION TERMS:	
PL	ENG
	recognize
zimno	cold
usłyszeć	hear
głodny	hungry
patrzeć	look
spojrzeć	
zauważyć	notice
widzieć	see
zobaczyć	
ogłądać	watch
podglądać	
zajrzeć	

Appendix C

EMOTION TERMS:	
PL	ENG
	brave
	dissappointed
	excited
	feel better
	feel well
	friendly
	great
	grumpy
	jealous
	joyful
	kind
	playful
	sweet
	worried
bać się	afraid
samotny	alone
	lonely
wkurzony	annoyed
przerażony	appalled
zły	bad
	angry
radosny	cheerful
obrażony	cross
płakać	cry
beczeć	
wy/przestarszony	scared
	frightened
zadowolony	glad
szczęśliwy	happy
lubić	like
kochać	love
wściekać się/wściekły	mad
(nie)grzeczny	naughty
zdenerwowany	nervous
miły	nice
smutny	sad
martwić się	upset
bezpieczny	
cierpieć	
cieszyć się	
dziwnie	

fajnie	
gniewać się	
groźnie	
groźny	
nienawidzieć	
pocieszony	
rozpacz	
strasznie	
straszny	
śmiać się	
wesoły	
zawstydzony	
złośliwy	
zmęczony	
zrobić przykrość	
żałować	