DE GRUYTER

Leda Berio

TALKING ABOUT THINKING

EPISTEMIC STUDIES
PHILOSOPHY OF SCIENCE, COGNITION AND MIND

Leda Berio

Talking About Thinking

Epistemic Studies

Philosophy of Science, Cognition and Mind

Volume 49

Leda Berio

Talking About Thinking

Language, Thought, and Mentalizing

DE GRUYTER

ISBN 978-3-11-074843-7
e-ISBN (PDF) 978-3-11-074847-5
e-ISBN (EPUB) 978-3-11-074855-0
ISSN 2512-5168

Library of Congress Control Number: 2021938024

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

Printing and binding: CPI books GmbH, Leck

www.degruyter.com

To every migrant.

Acknowledgements

As this book was built on the foundation of the research made during my PhD, I would like to express my deepest appreciation to my supervisor, Gottfried Vosgerau, for the ongoing academic support during my PhD, the insightful feedback, and for a lot of what I have learned while at HHU. I am also grateful to Albert Newen, my second advisor, for the incredibly helpful feedback and advice. I would also like to extend my gratitude to David Hommen and Katja Biermann-Ruben for the patient feedback given during advisory meeting, and for their ongoing support, as well as to Christoph Kann and Hana Filip for their valuable work as members of my examination committee. A sincere thank you should also go to Richard Moore, who has patiently discussed with me several ideas in this work, and to the editorial team at De Gruyter for their careful and precise support.

The research for my PhD thesis, which is partially what this book is based on, was funded by DFG Collaborative Research Centre 991: The Structure of Representations in Language, Cognition, and Science, Project D02. My most sincere thanks also go to the entirety of the SfB 991 team, including and especially my fellow PhD students, who never failed to provide insightful comments and encouragement. The same holds for the Philosophy department at HHU, including but not limited to Nicolas Lindner, Arne Weber and René Baston, for their amicable presence, their valuable input, and their inspiring research. My most sincere thanks also go to Eva Kellner and Eva Gentes, for their infinite patience in explaining every little bit of German bureaucracy to me.

All of the conferences attended in the last five years, and all of the questions received during every talk, have helped me getting to this point. A special thank you should also go to the audience of SSPP 2019, which was particularly reactive and engaged, and gave me meaningful input for developing some of these ideas.

I have had the luck and privilege of growing up in a family that never failed to express encouragement when it came to my somewhat unrestricted passion for learning in general, and for Philosophy in particular. It would not have been possible for me to arrive at this point without my parents' unshakable trust in me, their love, and their material and emotional support. Grazie, mamma e papà.

I am also indebted, as usual, to my brother. He has been my first confident, my first friend, and my certainty across my entire life. The process that led me to produce this book was no exception: his quiet support has been the reason why, even living more than eight hundred kilometers away from home, I have never felt truly alone.

https://doi.org/10.1515/9783110748475-001

Being an Italian in Germany would have been much harder without having a community to go back to when I needed to feel at home again, and without new friends making the journey easier. I'm extremely grateful to Alessandra, Camilla, ed Elisa, for being my brightest example of sisterhood; to Andrea, Francesco e Marta, for always making me feel like I have never ceased to belong at their side; to Giulio and Luca, for sharing, among countless other things, some of my (dis)adventures as a friend, fellow expatriate, and researcher; to Alex, for being my first friend here, as well as an amazing, inspiring person; to Ben and Esther, for sharing joy, frustration, and climbing evenings that made the hard work all the more enjoyable.

I am grateful to my extended families, both the Italian and the American one, for being a positive and bright presence in my life. The visits, the trips, and even the video calls in this last, hard year have continuously made my days richer, and for that I will always be thankful.

I would also like to extend my deepest gratitude to every woman before me that fought for her right to study, learn, and research. I would be nowhere if it was not for the effort of those before me.

Finally, none of this would have been possible without my partner, soon husband, Kurt. From the moment I met him, my life has been filled with more joy, laughter, and love than I could ever have hoped. His insightful intelligence has always helped me, in and out of academic contexts, and his compassion is my daily inspiration. It is also thanks to him if I am a better person, and a better researcher, then when I started exploring the topic of this book.

Contents

List of figures and tables

Figures

Tables

Part I: **Preliminaries**

Chapter 1: Introduction

1.1 Talking about thinking and thinking about others

The question of what makes humans able to communicate, has fascinated researchers since the very first philosophical and scientific inquiries. This is not surprising, since other animals' communication skills, while impressive, do not match natural human language.

Not only does natural human language seem to be a unique communicative tool, it also appears to come with a set of fascinating cognitive skills. Being able to communicate verbally entails an ability to use words, and also an ability to understand how our message will be received, guess how the interlocutor will understand it, and, in general, relate to and interact with other people in a meaningful way. In a sense, communicating is the enterprise of making two minds, two internal worlds, interact. If that was not enough, language allows us to discuss, evaluate, and predict other people's behavior in a verbal form, to build narratives around other agent's behavior, and to discuss them with our peers. This book is, in a way, about how communicating verbally and understanding others are deeply intertwined human abilities.

The ability to understand other's behavior, and to do so in the light of one's mental states, is often called *mindreading* or *mentalizing* in philosophical literature. It entails, or so the traditional story goes, the ability to attribute beliefs, desires, and thoughts to other agents. This ability has long been thought to be related to language: on the received view, language develops as a consequence of the fact that humans have mentalizing skills. In a nutshell, since we are able to understand what goes on in other people's minds, we have the skills to develop a language that allows us to communicate about it. This has been strongly influenced by the idea that mentalizing skills are possibly innate and modular, i.e. they are mostly present at birth, and only marginally dependent on experience.

More recent theories and empirical discoveries, however, seem to cast some doubt on this picture. On the one hand, apes seem to be capable of some (albeit limited) mindreading (Call and Tomasello, 2008; Suddendorf and Whiten, 2003). On the other hand, growing evidence suggests that language might be what aids the development of mentalizing (Diessel and Tomasello, 2001; de Villiers, 2005; Hale and Tager-Flusberg, 2003; Wellman et al., 2001). This seems to turn the traditional picture upside-down: perhaps it is not that our ability to communicate stems from our ability to understand others, but that our ability to understand others originates from our ability to communicate.

https://doi.org/10.1515/9783110748475-002

This is connected with another important issue in philosophy of cognition, which is that of the role of language in general: what is the role that natural language occupies in our cognitive architecture? Does acquiring language shape and rewire our brain and cognitive skills dramatically, or should we, rather, think of language acquisition as a domain-specific step in our cognitive developmental path, something that is only used for the specific purpose of communicating and barely entails cognitive changes in the rest of the system?

This book offers a partial reply to these questions – one relative to the development of mindreading skills and their relationship with language. I take an empirically motivated and philosophically argued stance against the traditional picture: I argue that the interaction between pre-verbal mentalizing abilities, language acquisition, and more sophisticated forms of mindreading is a dynamical and complex interaction. It is not the case that our mentalizing ability precedes and enables language *tout-court*; on the contrary (and this is the second stance I take), acquiring language provides a substantial boost to our abilities to understand others.

Arguing for this position, I simultaneously argue for a specific relation between language acquisition and our ability to understand others, and for a picture of cognition that sees language as a powerful cognitive tool. Far from being just a communicative device, the role of language in this picture is that of enhancing pre-existing abilities and potentially enabling new ones.

1.1.1 A brief background

Research on social cognition has been animated for a very long time by a crucial study (Wimmer and Perner, 1983), in which, for the first time, the false belief task paradigm was used: experimenters realized that 3-year-olds, but not 4-year-olds, reliably failed at a task in which they had to attribute a false belief to an agent in order to pass. The underlying idea was to test whether children were able to understand that a character leaving the room before the object of interest was moved was not supposed to be able to know, upon return, that the object had changed location. The first interpretation of this finding, which was replicated several times, denies to preschoolers under four the ability to correctly understand that agents can have false beliefs. The study generated a long, heated debate between, initially, Simulation Theory and Theory Theory of mentalizing (Gopnik (2001); Perner (1991); Gordon (1986) and others, summarized in chapter 4), followed by a long series of empirical studies aimed at replicating or reconsidering the initial conclusions (Onishi and Baillargeon (2005); Baron-Cohen et al. (1985); Zaitchik (1990); Scott et al. (2011); Scott and Baillargeon (2009) and

many others, summarized in chapter 4), and then an ongoing attempt to achieve clarity in the current literature and to formulate accounts that take into consideration all of the available data (Goldman (2006); Perner et al. (2015); Apperly and Butterfill (2009); Apperly (2011) and many others (chapter 5). On the one hand, there is a debate around whether or not children under the age of four are endowed with the representational and conceptual abilities necessary to attribute false belief; on the other hand, another fundamental question is "What exactly are the representational abilities involved?"

The debate on the role of language in cognition also has deep roots. Since the proposal of the Whorf-Sapyr hypothesis (Whorf, 1956), the research landscape has changed several times, going from a complete rejection of linguistic relativity to a more open-minded attitude towards the claim that language can shape our conceptual world. As I explore in the first chapter, the question has been discarded from the very start by so-called *communicative* views of language, according to which natural language is naught but a tool to communicate thoughts that are in another format (Fodor, 2008). This kind of view sees propositional thought as completely independent of language. On the other side of the spectrum, several theories attribute to language a pivotal role in "rewiring cognition" (Karmiloff-Smith, 1994), or providing the possibility for propositional thought per se (Tillas, 2015a). This debate is entrenched within several questions: how does a specific language shape the conceptual categories humans have (Lupyan, 2009, 2012), how are the concepts acquired in the first place (Carey, 2009), what is the underlying role of cultural practices (Gopnik, 2001), and what is the correct way to consider cognitive architecture (Carruthers, 2012)?

Finally, another domain of philosophical and psychological analysis is concerned with the more specific problem of how much of our social cognitive abilities is related to the development of language skills, ranging from theories that consider language as constitutive of some of the abilities involved in false belief reasoning (de Villiers, 2005), to theories that stress the cultural role of social narratives (Nelson, 2005; Hutto, 2008b), to theories that lean towards a middle ground between the two positions and argue for a role for both social development and linguistic acquisition (Garfield et al., 2001), and to many other positions besides (Van Cleave and Gauker, 2013; Gordon, 2007; Bermúdez, 2009; Montgomery, 2005). This debate has extended to linguistics (de Villiers, 2005), as it involves the analysis of how language-specific structures influence on the development of cognitive abilities.

This work's leading question is concerned with the extent to which language acquisition influences our cognitive abilities, with a focus on false belief reasoning in relation to the acquisition of mental state verbs. This might appear to be a very specific question; however, as the reader will see, there are many inter-

twined issues and problems which come together to form an answer to this question, and it is on the literature listed above that the work in this book relies.

The set-up of this book is funnel-like. It starts with a rather general set of issues (theories of language and cognition) and works towards a very specific one one; this not only reflects my personal path of inquiry, but also avoids the risk of obtaining an account of the influence of language on false belief reasoning that is completely disconnected from a general view of cognition. I will come back to this topic more than once in the upcoming chapters. This does not mean that I give a final answer to the question of what the exact relation between linguistic and non-linguistic cognition is in general; it means however that I give an account of how mental state terms help the child developing abilities to pass the false belief task in a way that does not disregard research concerning language and cognition, but rather is compatible with it. At the same time, a fundamental concern in this book is to give an account of the role of language in social cognition that also fits with the developmental trajectory of the empirical literature that focuses on pre-verbal skills; this is important, because, as will be seen in chapter 5, many rival accounts fall short of this objective. In this sense, while the scope of the book might appear narrow, there is a conscious effort to make my account fit with the bigger picture in two ways: in relation to the development of other skills that are central for mentalizing, on the one hand, and in relation to the more general problem of how linguistic coding and non-linguistic coding interact in human cognition, on the other.

For these reasons, I start with an overview of the issues underlying the interaction between language and thinking in chapter 2. I argue in this chapter that the most fruitful perspective is *pluralism*, i. e. a perspective according to which the problem of the relation between language and thought is tractable only once we define with more clarity the domains of our inquiry. On the one hand, we have to choose the cognitive domain to explore; on the other hand, it is useful to think about which kind of linguistic information is most likely to be relevant to that domain. This is the underlying reason why the rest of the book focuses on the very specific question, whether acquiring mental state terms (described in chapter 3) has an impact on how children pass the famous false belief task in its explicit variation. The answer to this question is mostly given in chapter 4, and it is one that builds upon the empirical data and the theoretical contributions, which are further analyzed in chapter 5. However, this book is set up to answer a related, but different question as well: not only if mental state terms play a role, but also *how*. The answer to this question is given in chapter 6, which contains my theoretical proposal. In chapter 7, I indicate a path for further research and draw more general conclusions.

1.2 Methodological note

What makes this work a philosophical enterprise is the fact that I constantly relate the issues at hand to the underlying theoretical questions about human cognition and mind; what makes this work an empirically-informed philosophical enterprise is that I give constant priority to experimental data, taking seriously the importance of experimental paradigms, design, and analysis choices. While this is not meant to be a psychological inquiry, and a very modest amount of attention is dedicated to statistical analysis of data, I do put empirical reviews at the center of this work. This is not only true for psychological data, but also for language: the reader will realize, while reading this book, that a lot of importance is given to linguistic analysis. This is the case for two fundamental reasons: on the one hand, it is rare that philosophy, psychology, and linguistics tend towards the same research themes in such an evident way as in language and mentalizing research. Studying the influence of language learning on mentalizing tasks necessarily means considering language acquisition literature as well as cognitive psychological data. On the other hand, an issue that is of fundamental importance for me and my work is that of cross-cultural and cross-linguistic analysis. While the cross-linguistic data available is limited, I make an effort through this work to specify which findings, heavily supported by research with English-speaking participants, can indeed be extended to other populations, and which of the linguistic features are shared by other languages. As I further elaborate in chapter 3, this is of fundamental importance in order to make today's psychology and philosophy as impartial as possible. Hence, when achievable, I report on cross-linguistic data available in the literature: this sometimes includes reporting on technical linguistic phenomena, but I do provide the theoretical tools that are necessary for understanding the underlying issues. While not being a linguistics book, this work is meant to be informative for philosophers, linguists, and psychologists alike, in providing conceptual tools that lie at the interface of these disciplines.

For these reasons, the reader will find that two chapters (chapter 3 and chapter 4) are completely dedicated to empirical reviews. As will be clear, this includes critical review and analysis of the theoretical implications of the study. In the same spirit, chapter 2 contains a literature review, albeit shorter, of the data confirming that the possibility that language influences non-linguistic cognitive tasks is worth exploring.

In my theoretical literature analysis chapters, 2 and 5, it will be clear that I bridge two different bodies of literature, interconnected but separated: on the one hand, the theoretical proposals concerning the relation between language and thought, and on the other hand those concerning the specific relation be-

tween mentalizing abilities and mentalizing. As explained at length in chapter 5, I give more attention to theories that openly address the issue of mentalizing and language, despite the fact that these are not always the most discussed theories in the literature. This constitutes a novel contribution of this work too, since I put some less discussed theories, from both linguistics and philosophy, in dialogue with psychological theories and philosophical theories on the interface between thinking and speaking. I consider the angle taken in this book a novel one, and one that makes this book a valuable contribution to the debate. While I do set aside some important theoretical contributions, I do specify my reasons for doing so in the relevant chapters.

1.3 Terminological choices

When dealing with social cognition, it is very hard to avoid loaded terms. As the reader will see, I have chosen the term "mentalizing" to describe the ability to explain other people's behavior on the basis of their mental states. I have avoided adopting the term "Theory of Mind" because of the long tradition of theories (Gopnik (2001); Gopnik and Astington (1988); Leslie (1987); Perner (1991); Perner et al. (1987); Wimmer et al. (1988) among many others) that consider mentalizing mostly theory-based; an explanation of this is given in chapter 4 and then reconsidered briefly in chapter 5. I have also avoided the term *mindreading* for similar reasons; skeptics like Hutto (2008b) have expressed their doubts about the idea that we should conceptualize mentalizing activity as the ability to "read" somebody's mind, even when this only means attributing something like a mental state. I do believe that we engage in attributing activities: however, I have chosen the term "mentalizing" (and not "mind-minding" as Hutto might prefer) as a relatively neutral label for the activities in social cognition that are of central interest in this work. As will become clear to the reader very soon, focusing on the role of language also means focusing on the onset of language abilities, which is also the reason why most of my attention goes to the so-called "explicit false belief task". A clear explanation of this theoretical choice is found in chapter 4.

Another term I have used parsimoniously in this work is "*concept*"; avoiding it is only partially possible, and the reader will see that I do mention my commitment to concepts as theoretical entities in the final chapters. The reason why I avoid referring to concepts of mental states for most of the book is the same as that advanced by Apperly (2011): the term is loaded, and the debate around what exactly constitutes a mental representation or a context is a very hard, very heated one. I do suggest in my own proposal that children do not need to

have well-formed concepts of mental states when they pass the false belief task, and that it is reasonable to assume that these only come in handy after a considerable amount of cognitive and linguistic development has already occurred, and mentalizing abilities that are considerably more sophisticated than those belonging to a 4-year-old come into play. However, I do focus on earlier stages of development and on how preschoolers deal with tasks that appear to get much easier around their fourth birthday. As will be seen, this entails dealing with debates that cross-cut disciplinary boundaries.

1.4 A map of this book

In chapter 2, the reader can find an overview of the most general questions related to language and thought. The chapter includes an empirical overview focused on color cognition, where I analyze the available evidence in favour of thinking that linguistic labels influence non-linguistic cognitive processes. I then proceed to include data outside of the color cognition debate showing that setting up an exploration of the specific role of specific linguistic input on cognitive abilities is worth pursuing. In light of the results of the empirical literature review, I present an overview of some prominent theories regarding the interaction of language and thought, concluding that only some of them manage to account for the data presented in the empirical overview.

In chapter 3, I discuss specific data about the acquisition of mental state verbs. I delineate some of the most important questions in the debate and connect the data with views on how the meaning of verbs is acquired in the first place. While a lot of the data presented is linguistic in nature, the chapter underlines the connections with the philosophical and psychological debate on acquisition of language and sets up the foundations for the following empirical review. The chapter also takes into consideration cross-linguistic data and cross-cultural reports on how mental state terms differ with respect to when they are acquired and the structures in which they appear across some different cultures and languages. Readers that are especially well-informed on the literature and on the fundamental issues regarding mental state verbs can safely skip this chapter.

In chapter 4, I first explain the problem of mentalizing and false belief tasks, specifying which ones are the most used paradigms and explaining why so much of the focus in this book is on the explicit false belief task. After setting up the debate, the chapter proceeds by analyzing the findings related to the interaction between false belief reasoning abilities and language acquisition. The chapter takes into consideration data about semantic, syntactic, and pragmatic skills,

along with data regarding narrative practices, storytelling, and clinical data on populations with conditions like ASD, SLI, schizophrenia, deafness and William's Syndrome. The overview makes it clear that language definitely has an influence on false belief reasoning, but that the nature of this influence is complex and multi-faceted. This chapter too can be overlooked by those who are familiar with the empirical data related to explicit false belief reasoning. While I do dedicate time to scrutinizing some of the claims and methods of the studies in this chapter, the reader who follows the debate closely might not need to revise the available evidence.

In chapter 5, I analyze the existent accounts of this influence and argue that none of them are completely satisfying in terms of theoretical coherence nor empirical validity. I point out that the weakness of many accounts lies in the fact that they do not take into consideration the bigger picture, as they do not explain how linguistic skills interact with pre-existent abilities in preschoolers, nor address the larger problem of how to position their view in a larger picture of cognition. I point out that approaches that rely on linguistic input *per se* especially do not take into consideration more general questions about the relation between language and thought, and that approaches that stress the socio-cultural components are often underspecified.

In chapter 6, I propose my own account of how language acquisition has an impact on mentalizing skills. The main idea, already sketched in Berio (2020a) but here expanded upon, is to take seriously both the empirical data suggesting that specific language features have a direct impact on children's ability to pass the false belief task, and the idea that social-cultural development is playing a major role. I name my proposal LALAS (Language Associations Labels And Schemata) and I argue that linguistic input helps the children to individuate reliable patterns in explanations of other's behavior, providing them with the ability to form abstract schemata that can be used in folk-psychology narratives to better memorize, store, and retrieve relations between agents and states of affairs. I relate my account to two different views on mentalizing (Apperly, 2011; De Bruin and Newen, 2014) and I also place it in a more general view of language and cognition, providing two viable alternatives: we can situate it within, first, Tillas' (2015) view, or second, within a more general approach according to which language is a powerful cognitive tool.

In chapter 7, I summarize the findings of this work and I provide some suggestions for further research. I also discuss how LALAS is, at the moment, supported by the available empirical data and how it could gain further support, and I indicate the kind of empirical study is needed to clarify the matter further.

Chapter 2:
Language and thought: in favor of a role for linguistic information in cognition

This chapter argues in favor of the view that linguistic processes have an influence on cognitive development. The scope of this chapter is more general than the rest of the book, since it will not focus on the specific domain of mental state terms and social cognition. On the contrary, the theme of the acquisition of mental state terms and its influence on non-linguistic cognitive processes will be left aside, since it will be further developed in subsequent chapters. However, this chapter is essential for the aim of this book, since it is meant to prove that an inquiry about the role that the learning of specific language structures has in non-linguistic cognitive development is valid in the first place. In the first part, I will delineate some central assumptions and problems in the research regarding the interface between language and thought. In the second part, I will discuss some relevant data available in the literature that supports the idea that language learning and linguistic development have a significant impact on non-linguistic cognition. I will firstly focus on color cognition as a specific case, and then give an overview of the data in other semantic domains. In the third part, I will provide a review of some of the prominent theories that *are not* compatible with the presented data and a longer overview of theories that fit the presented data better, and analyze their implications, assumptions, and predictions, in favor of the claim that linguistic processes influence cognitive development.

An important point that has to be kept in mind is that in this chapter it will not be argued that linguistic processes are necessarily pervasive in cognition, nor will the data about color serve the purpose of providing a definite proof of linguistic relativity. The aim of the chapter, on the contrary, is to argue that it is worth pursuing an investigation into the effects of language, the domains in which these effects are present, and which theoretical tools are best to analyze them.

2.1 Language relativity: which kind of influence? Preliminary assumptions

For a long time, the debate regarding the influence of language on cognitive processes, and in particular on color cognition, has been influenced by a strong

https://doi.org/10.1515/9783110748475-003

bias against the idea of language relativity: one of its earliest formulations, namely the Sapir-Whorf hypothesis (so labeled by Carroll in Whorf (1956)), suggests a particularly strong and simplistic influence of language on thought. As has been argued subsequently (Gopnik, 2001; Lupyan et al., 2007; Wolff and Holmes, 2010), the hypothesis can be read in different ways, and not all of them are as radical as they might appear. What makes the Sapir-Whorf (Whorf, 1956) hypothesis naturally difficult to accept is the idea of a *parallel structure* between language and thought; the claim, which is rather bold, is often accompanied by the stronger prediction that language categories basically *rewire* perceptual and conceptual capabilities. As we will see, this bears some similarity to a more accepted view often labeled the *rewiring hypothesis* (Bermúdez, 2005); however, language relativity in its original whorfian formulation has a stronger, more radical standpoint.

The debate around language relativity has been at the center of further research mainly because more recent studies have investigated the role for language in non-linguistic tasks, while re-adjusting the strong assumptions characteristic of the original Whorfian proposal. This is especially true of the color debate, as we will see in 2.2, in part because what was initially taken as the final word on the relativist approach to color cognition (namely the study by Berlin and Kay (1969)) has been scaled down to an important but not decisive piece of the evidence. As will be further explained in 2.2, the color debate is concerned with a particular aspect of language, which is labeling: most studies in color cognition focus on whether or not color terms in one language have any influence on performance as far as color recognition, classification, or memory are concerned. Another important clarification is that the debate is concerned with influence on perception and categorization tasks. Color cognition is often considered the privileged, if not exclusive, ground for deciding about the whole debate concerning Whorfianism and linguistic relativity, which is one of the reasons why it will here receive particular attention: however, it is beneficial to keep in mind that it is only concerned with a specific aspect, which is how *lexical entries* influence *perception and attention mechanisms*. This order of specifications does not only apply to the color cognition debate: more generally, talk about "language" and "thought" is quite clearly a misleading simplification, as argued in Berio (2020b) and Beaulac (2014). The issue is often treated in terms of a relation between "language" as a whole and "cognition", in the best cases substituted by the still vague notion of "perceptual" or "cognitive" mechanisms. However, to understand whether linguistic processes interact with other cognitive processes, it is necessary to formulate the question in more precise terms, asking *which* linguistic-related process (labeling, vocabulary development, syntax, or semantics)

can interact with *which* possible cognitive process, for instance long term memory, working memory, recognition abilities, perspective shifting skills, and so on.

This combines well with a recent proposal by Beaulac (2014), who names it the "pluralistic view"; in other words, the idea that language can be thought of as having several (and quite different) roles in cognition. This obviously implies not only that it is possible to talk about different sources of influence and interaction between language and thought, but also that there are several *ways* in which this can happen. To this extent, Beaulauc suggests a logical distinction between processes that would be present even without language and are not influenced by it (1), processes that are enhanced by the interaction with language faculties (2), processes that are modified by it (implying possible empowerment/improvement in performance but also a "negative" or "neutral" influence on the process) (3), and processes that require language to exist (4). Note that, while I think it's legitimate to assume that influence does not directly imply improvement (hence the distinction between (2) and (3)), at the same time I think that (3) clearly stands out as a more general case of (2) and therefore the two cases might as well be unified. This reduces the cases in which processes do have some relation to language to two macrocategories, for which it would be sufficient to say whether or not language faculty is a necessary condition for a certain task. The distinction between this kind of influence surely helps to clarify the matter, as it focuses on individual singular processes; however, integrating this approach to Wolff and Holmes's typology of positions (2010), the picture becomes even clearer. In their instructive review, Wolff and Holmes cite a great variety of studies that fit into the idea that language and thought, while differing fundamentally at the level of structure, still interact in significant ways, drawing a boundary between this case and language-of-thought and language determinism.

Language as spotlight, as they define it, refers to the possibility of language highlighting specific properties and drawing attention on them, and refers to phenomena in which encoding specific features linguistically can make attention linger on them. A quoted example is Boroditsky and colleagues' study (Boroditsky et al., 2003), in which the ability to memorize associations between proper and common nouns was disrupted for the participants when the biological gender of the proper noun's referent and the common noun gender did not match[1]. This is clearly a case in which language influences attention mechanisms and so might be regarded as a case of (3). *Thinking for speaking* refers to processes that are modified by the fact that language will be needed: participants in a study by

1 Though, see Mickan et al. (2014) for discussion.

Casasanto et al. (2004), for instance, were influenced by the fact that they had to provide verbal descriptions of videos after watching them, to the extent that they focused on characteristics of movement that are encoded in their languages (manner for English, path for Greek). In this case there's an instance of (2), where a process is modified and influenced by linguistic coding, and arguably attention on certain aspects is enhanced and modified by it. The same applies to the expression *language as a meddler*, that clearly refers to the idea of linguistic coding interacting with other kinds of representation. Finally, *language as an augmenter* refers to the possibility of combining linguistic representations with non-linguistic ones in order to perform tasks that are unlikely to be fulfilled otherwise, and pertains somehow to (3); *language as inducer* refers to the process of language priming particular modes of processing, which may remain engaged when language is not induced anymore.

Let us have a look at the classification that results from combining the Wolff and Holmes (2010) and Beaulac (2014) proposals:

1. Modified processes:
   - Language as meddler
   - Thinking for speaking
   - Language as augmenter
2. Enhanced processes:
   - Language as augmenter
   - Language as spotlight
3. New processes:
   - Language = Thought
   - Inherently linguistic processes
   - Language as inducer

All these distinctions considered, there is focus on:

- The possible different processes involved;
- The different language aspects involved;
- The relation between the linguistic coding and the cognitive processes;
- The kind of role that the linguistic skills can play in the interaction.

Once all of these factors are taken into consideration, it is easier to overcome the hypersimplification that stems from naive formulations of the linguistic relativity issue – and therefore to scare away the ghost of a relativist perspective that was so strongly opposed in the past. Once these specifications are made and the nature of the inquiry is better defined, it is consequently also easier to analyze the problem further and to consider whether some conclusion can be drawn. In the next section, I will make some remarks about the problem of linguistic relativity

itself and its interpretation in the literature, before presenting data that sheds some light on the question "What kind of influence between linguistic and non-linguistic information is possible?"

2.1.1 Universalism, "deep" and "shallow" Whorfianism; interwined issues

As noted by Lalumera (Lalumera, 2014) and Berio (2020b) the evidence available in the literature cross-cuts the distinction between Whorfianism and universalism. Some influence of linguistic labels on perception and categorization mechanisms, as it will be seen, is present in the literature, but it does not necessarily support the stronger claim that language always rewires conceptual categories.

In this spirit, Lalumera deems it more useful to use the distinction between "deep" and "shallow" Whorfianism, marking a difference between the constant, pervasive, and stable influence of linguistic labels and those cases in which it is "only" a flexible, context dependent, task dependent influence of some sort. This brings the focus to a more fine-grained question: for the universalist, Whorfianism threatens the idea of concepts as independent from the specific language of the speaker, and as a consequence the idea that humans have a universal conceptual repertoire. Whorfianism, on the other hand, is partially motivated by the fact that universalism seems to exclude any interference of language with mental representations, discarding the possibility that a learned language has consequences for conceptual structures. Talking about deep and shallow Whorfianism, on the other hand, brings the focus to a pragmatic issue, i.e. "How and when do linguistic processing and linguistic labeling interfere with non-linguistic processes and how, if indeed at all, does acquiring a language influence the processes themselves?" In this sense, the relevant question concerns when the influence is relevant and how stable and pervasive it is[2].

2 As noted in Berio (2020b), this issue is connected to that of cognitive penetrability, which is defined as the property of perceptual experience to be influenced by what happens at the "higher" cognitive level; we speak of cognitive penetration when perceptual experience is influenced by beliefs, desires, intentions, and concepts (Newen, 2011; Newen and Vetter, 2017). On the one hand, the issue of penetrability is parallel to the issue of the interface between language and thought processes: admitting that linguistic information has an influence on non-linguistic (and specifically perceptual) processing means admitting permeability of perceptual experience. The problem of permeability, on the other hand, is of a broader nature, as it comprises considerations regarding modularity and specialization of brain areas; in other terms, the debate regarding permeability brings us to a broader scale of issues regarding cognition, while at the same time focusing on the specific problem of defining perception. Admitting permeability

As will be seen in the upcoming empirical review, there is not much evidence for a pervasive, stable, task-independent influence of language on many cognitive processes. However, a lot of evidence supports the idea behind "shallow whorfianism", at least if defined as a thesis about an influence of language that is only detectable in specific tasks. From a universalism-Whorfianism point of view, this kind of influence is irrelevant, because the question at issue is whether speaking a different language irreversibly shapes the conceptual repertoire in a deep, pervasive way, which is not a prediction made by "shallow Whorfianism". However, as Lalumera points out:

> [...] some Whorfian effects show themselves to be task dependent and temporary. A question on this point is worth raising here. Is that enough to deem such effects as uninteresting, qua task dependent and temporary? The answer is that it would be enough, but at the price of committing to the view that only stable and context-free representations are employed in perception and cognition. (Lalumera, 2014, p.7)

In other words, as argued in Berio (2020b), if one starts from the idea that mental representations are not necessarily amodal, invariant, and context-independently recruited, even so-called "shallow" effects acquire new importance, especially given the fact that, even if fleeting, task-dependent and temporary, linguistic interference can be relevant for a complete description of mental processes. Lalumera argues that fleeting, transient influence of language on cognitive tasks is worth exploring to shed new light on language and thought interaction issues, among other things. A very similar point is made by Lupyan (2012), who proposes a mechanism to deal with this transient interaction, which will be explored in section 2.4.3.2.

2.2 Color cognition and data

2.2.1 Why color?

In the second part of this chapter, I will discuss two main bodies of literature. Crucially, I will review studies that have successfully shown that linguistic information like labels can indeed affect non-linguistic processes. My first area of focus will be that of color cognition; this is because, among the many conceptual domains, color terms have been particularly essential, historically speaking, to

---

means admitting that experience changes depending on (among other things) linguistic processes. This is obviously related to the topic of this chapter, but does not overlap with it.

initiating the debate regarding the interaction between language and thought, to the point that they have been considered the most important battleground. Their importance, on the other hand, is not exclusively historical; the increasingly large body of literature on color cognition and color labeling continues to inform the debate, and has important consequences for the issues raised by Cognitive Penetration theories as well. In what follows, I will show how empirical studies give good reasons to argue for a role for language information in non-linguistic tasks; at the same time, I will highlight the limits of such a claim, and its consequence for the general issue of the relation between language and thought. Subsequently, I will briefly discuss how data in other semantic domains seem to support the case for a more general interaction between language and thought, expanding the cognitive processes considered beyond categorization and visual recognition, which are central for the color debate. This second empirical review will also show that, despite the encouraging results that point to the influence of language on several cognitive processes, there are good reasons to think that this influence is somehow limited, especially in the case of higher-order cognitive processes.

2.2.2 Categories and color terms: evidence

In this section, I will focus on studies on color terms. As will be seen, most of these deal with perception tasks and with the relation between linguistic labels, categorization, and perceptual information. For this reason, these studies have been the battleground for research concerning cognitive penetration as well; the results do not only cast light on the issue of the role of linguistic labels, but also on the more general question regarding how linguistic coding interacts with non-linguistic processes. This, in turn, is considered a privileged ground for verifying the more general question, whether or not perception is *penetrable* with respect to higher cognition.

Winawer et al. (2006) is one of the most cited studies in this field of research. Russian has an obligatory distinction between *light blue* and *dark blue* (*goluboy* and *siniy*) like Greek and Italian. In the study, subjects (divided into a Russian speaking group and an English-speaking group) were shown three color squares arranged in a triad; the task consisted in saying which one of the bottom squares was identical to the one on top, while reaction times were measured. In "within category" trials, the square was of the same color category as the match, whereas in "cross-category" trials the distractor and the match belonged to different categories of the Russian color categorisation system. The question was whether the absence or presence of a linguistically encoded color boundary would affect per-

formance and, more specifically, whether Russian speakers would make faster cross-category discriminations than within-category ones. The results provided an affirmative answer, showing a difference between the performance of Russian and English speakers. Crucially, the effect disappeared if the subjects also had to perform a verbal interference task (mentally rehearsing strings of digits) at the same time: blocking language resources with task-irrelevant processing was as a consequence preventing the effect. As will be pointed out, the role of interference tasks of this kind can be crucial for interpreting the results of these sorts of studies, and it is hence a topic thoroughly explored in the literature. Winawer et al. (2006) also found through statistical analysis that the difference between cross-category and within-category trial performance for Russian speakers increased the more difficult the discrimination was. The results were interpreted by the authors as supporting a decisive influence of color labels on the perceptual task. The fact that the facilitation disappears when a linguistic interference is added suggests that the effect on perception is temporary and tied to the specificity of the task. At the same time, the results of the verbal interference condition suggest that language labels are likely the cause of the effect, because linguistic coding seems to be involved. This is arguably a case of "language as a meddler" (Wolff and Holmes, 2010), where there is an *online* interference that takes place during a certain task and that is heavily dependent on the context and conditions of the task itself. In this case, moreover, labels *change* the performance in a color discrimination task. The fact that the effect was larger for more difficult discriminations, finally, suggests that labels might be a facilitator in the task. This is in line with the idea that language can act as an *enhancer* as well, i. e. that it can actually boost already existing skills. Note that such a facilitating effect was confirmed in other studies: for example, Roberson et al. (2004) presents evidence of a similar kind, related to memory. Better recognition memory performance was observed for Berimno speakers, who only have 5 basic color terms, for colors for which they possess a specific linguistic label.

Cases like Winawer's seem to suggest that conceptual representations of colors, and consequently their labels, can be used and activated during a perceptual task; one of the possible interpretations of the results is that, while English speakers operate by comparing different perceptual inputs without activating linguistically coded representations, Russian speakers use a different strategy, namely they employ color concepts and their labels; at least that is what seems to be suggested by the difference in performance. Crucially, however, this kind of strategy seems to be replaced by the same strategy English speakers employ, in case of linguistic interference: somehow, then, performing another linguistic task "blocks" or inhibits the label-influenced strategy. Given the fact that the task is still possible for English speakers, this is clearly not something

that prevents them from performing, regardless of the presence of color labels. What this study seems to suggest is that recruiting or not recruiting linguistic information can depend on the type of task: in this sense, the choice of strategy is flexible. The effect of verbal interference has been explored in several studies. In Roberson and Davidoff (2000), subjects had to perform a visual task consisting in choosing between two color patches, one of which was similar to the target that was previously shown, with within-category and cross-category trials, in three conditions: verbal interference, visual interference, and control. In the verbal interference conditions, subjects had to read non-basic color words aloud; in the visual interference condition, they had to track a line of dots with their eyes. Results showed that performance for cross-category judgments was better in every condition except for the verbal interference condition. While within-categories judgments were not affected by verbal interference, visual interference caused a reduction in performance in both cross-category and within-category trials. The finding about verbal interference was confirmed in the second experiment; in a third experiment, they compared the verbal interference task with color words with one employing words unrelated to the color domain, finding that the same effect was confirmed regardless of the kind of words used. The authors concluded that there is an effect of verbal coding on categorical perception.

Brain-imaging research also brought some results to the table. A famous study by Tan et al. (2008) showed that Wernicke's area, which is dedicated mostly to language processing and not known to be a visual area, had a stronger activation during a same-different visual recognition task when the target colors where "easy-to-name" as compared to "hard-to-name". The authors argue for a strong involvement of language categories in perceptual activity, which is only partially sustained by this correlational data: however, it is surely interesting that a linguistic area seems to be automatically activated in a non-verbal task, and that it is activated in a selective way (i.e. the activation varies depending on the stimuli). Interestingly, the distinction between blue and light blue has been investigated in an ERPs study on Greek speakers: in Thierry et al. (2009), Greek and English participants had to individuate a target, a square shape, in a stream of stimuli which was mostly composed of circles. Circles were either light green, dark green, dark blue or light blue. The two greens were matched with the two blues used in terms of luminance distance, and green and blue stimuli were equidistant from the background in terms of saturation and luminance. Pre-tests determined that the instances of dark and light blue were identified by Greek native speakers as good examples of *bleu* and *ghalazio*, i.e. light blue and dark blue. The study investigated visual mismatch negativity (vMMN), i.e. an early component that detects automatic and pre-attentive detection of change in perceptual stimulus. Results showed that the vMMN effect was signif-

icantly greater for blue compared to green deviants in Greek participants, but not in English participants. This prompted the experimenters to focus on other factors, like P1, i.e. the positive peak elicited by visual stimuli in the parieto-occipital regions. This component is documented to be sensitive to color boundaries (Fonteneau and Davidoff, 2007). While P1 was only significantly modulated by luminance and not by color, the same was not true for Greek speakers, for whom the interaction between luminance and color significantly modulated P1. P1 latencies were the same for greens and blues for English speakers, but the same was not true for Greek participants. The study is particularly relevant because, as the authors underline, it shows an effect at *early stages* of visual perception, where automatic and pre-attentive processes take place. Moreover, similar results where achieved in a study by Mo et al. (2011) investigating visual mismatch negativity, thus confirming an early and automatic involvement of label-like information.

Interestingly, another study investigated similar effects for sign language (Xia et al., 2019). Signers show different brain activities when communicating compared to verbal speakers, with different patterns of activation in the right hemisphere indicating a bilateralization of the language processing areas: this is likely to be due to the spatial nature of sign language, being that the right hemisphere is fundamentally dominant in spatial processing. In the first experiment, participants (deaf and hearing) had to decide as quickly as possible whether the target (the different color patch) was on the right or left part of the screen. Easy-to-sign/name colors and hard-to-sign/name colors were used in the trials, and different lateralization effects were found. For the hearing group there were shorter reaction times for easy colors, but only when the target was in the right visual field; the same was not true for the deaf group, where easy colors were recognized faster in both hemispheres.

In the second experiment, a spatial interference task was introduced for a similar experiment with non-hearing only participants. While easy-to-sign colors were still recognized faster when presented to the left visual field, the effect disappeared when the stimuli was presented to the right visual field. In a third experiment, nMMN was measured for non-hearing participants, verifying that deviance in color generated less "surprise" if the deviant stimuli occurred in the same lexical category as the previous color patch, even when the cross-category color patch was perceptually equidistant. This confirmed an effect of linguistic information in a pre-attentive stage of perception. The study is particularly relevant because it puts the "Whorfian effect" found in other contexts in relation to sign language, pointing at how the spatial nature of signs modifies the patterns found in hearing subjects.

The "flexibility" of the lateralization effect was also tested for *novel* color categories in a study by Zhou et al. (2010), where participants were trained with new categories (learning two different names for different shades of blue, for example), and had to perform a visual search task. Reaction times for (newly) cross-category stimuli were faster in the right visual field then the left visual field, as registered in other studies with "normal categories".

Other results regarding cross-linguistic data comes from Roberson et al. (2008), who explored differences between English and Korean speakers, and once again brought into focus the issue of lateralization. Korean has fifteen basic color terms, whereas English only has eleven. Since it is often argued that language centers are located in the left hemisphere, and categorisation functions belong to clusters in the right hemisphere, the study investigated the categories of *yeoundu* and *chorok*, (respectively *yellowgreen* and *green* in Korean) on the basis of this distinction. In the task, participants were presented with an array of color patches, one of which was different from the others. The patches all belonged to the category *green* for English speakers, whereas for Korean speakers the "odd ball" patch could belong either to the category *yeoundu* or *chorok*. Participants had to identify the odd ball, saying whether it was on the right or left of the screen: this was thought to cause the stimulus to be processed either in the right or the left hemisphere. The results showed difference in cross-category and within-category discrimination: Korean speakers made faster cross-category judgments compared to within-category ones, and there was an effect regardless of the visual field. A further comparison between fast responders and slow responders revealed that fast responders were only facilitated when the stimulus was presented in the right visual field, whereas the effect was present for slow responders even for stimuli presented in the left visual field. This, the authors argue, supports the idea that the effect is due to linguistic labels: in the case of slower responses, time allowed the information to be transmitted via the *corpus callosum*. As in other cases presented above, the study revealed that the influence of labels, though reliable, was however task-dependent. Lateralization is central in a different way in the ERP study by Liu et al. (2009), where the N2 posterior central component elicited in a visual-search task was investigated and showed to be larger for cross-category and within-category trials only in the left hemisphere, which suggested the involvement of the language processing areas in distinguishing the colors belonging to different linguistic categories. Evidence very similar to Roberson et al. (2008) was collected by Gilbert et al. (2007) for English. Similarly to Winawer et al. (2006), there was a cross-category effect, but it was connected to lateralization. English speakers were faster in locating a target, the color of which was different from that of the other objects presented as stimuli, when it was of a different color category (so, for example, green and

not blue like the other objects) than when it was not (for example, a shade of green among other shades of green), but only when the target was presented in the right visual field, which suggested the task was influenced by the hemisphere.

Evidence comes from lesion studies as well. In a study by Davidoff and Roberson (Roberson et al., 1999), a patient with lesions in the Wernicke's area (which is one of the centers of language) was examined. The patient, whose perceptual abilities seemed intact, was unable to perform free-sorting tasks for color patches according to hue and lightness in a way that reflected non-lesioned subjects; while remembering some color labels for basic color terms (like red, yellow, and brown), he was unable to associate them with the right color groups once a smaller range of patches was given, despite being able to match pairs of chips with similar color. According to the experimenters, his impairments showed that the language deficit had consequences for categorization in a deep way; despite the fact that implicit information for categorization was available to the patient, making explicit judgments in a categorical task was almost impossible. Lesion studies are especially valuable to neuroscience because they allow neuroscientists to investigate how very specific conditions and impairments affect specific tasks. However, we should not underestimate the complexity of brain lesions, and the limited extent to which we are able to attribute a specific impairment to a functional region of the brain. The importance of these studies remains, as they provide good reasons to explore a functional link between brain areas and task performance.

2.2.3 Colors: the importance of flexible tasks

An important point emerges from this short literature review: the effect of language on category judgements and perceptual tasks like visual searching is somehow dependent on the task at hand. As said, I do not believe this to be a good reason for dismissing the results; on the contrary, I think the evidence is very clear in pointing in the direction of an important role for labels in color cognition. The fact that this effect varies depending on the task at hand, as underlined by Lalumera (2014), Lupyan (2012) and Berio (2020b), does not imply that these effects are not relevant. It only implies, strictly speaking, that our categorization skills are not *solely* based on our mastery of language. This is in line with other results emerging from the literature, which will be treated in section 2.3.2, and confirms one of the assumptions spelled out at the beginning of this chapter: it is of fundamental importance to restrict one's analysis to the question of which *specific processes* are affected by which *specific language information*. It is

also a good indication of something that will be argued for in more detail later, i. e. that neither views that attribute to language a constitutional role in thinking nor views that completely dismiss its influence on cognition mechanisms can be correct. In the previous section, it has been argued that there is good evidence that labels for colors have an impact on how we perform in some simple color cognition tasks. In what follows, I will describe further evidence in favor of a relevant role of language in cognition.

2.3 Beyond the color debate

2.3.1 Not only colors, not only categories

In this section, I will present data that pertains to semantic domains and/or to different cognitive processes. The aim of this will be to show that, while so much of the current and past research focuses on color categories, the claim that language processes and information can indeed make a cognitive difference in non-linguistic cognition can find empirical support of another kind. One specific area, that of social cognition, will be the main focus of the rest of the book; as a consequence, it will not be treated here. Instead, I will be focusing on presenting some data about other cognitive domains, where it is possible to see that the available empirical research documents the important role of labels, not only for color cognition and recognition, but also for other cognitive mechanisms.

As we saw in the previous section, studies regarding novel created categories can be very useful in assessing the contribution of language. Lupyan et al. (2007) compared the learning abilities of participants that had to learn new categories with and without labels. Sixteen aliens interacting with a space explorer had to be classified into two categories based on body features: “approachable” and “better-to-avoid”. During training, a figure representing a space explorer would approach or escape from an alien, and an auditory stimuli would describe the act as right or wrong. In the label condition, the stimuli was accompanied by one of two different labels, one used for good aliens (to be approached) and one for bad ones (to be avoided). After training, participants were presented with the aliens and had to decide whether or not they were of the approachable or the dangerous kind; participants in the label conditions performed significantly better and faster than those in the non-label condition. In this case, labels seem to act as a powerful memory device, in line with the *language as an enhancer* idea spelled out above.

The role of language in facilitating categorization is further explored in Lupyan and Mirman (2013) in relation to aphasia. Here participants had to cat-

egorize objects according to a given criterion, which could be a high-dimensional criterion, like being an animal farm, or a low-dimensional criterion, like being red. The crucial difference between the two criteria is that, for two things to be of the same color and be in the same category, the category has to be formed on the basis of only one dimension, whereas to be two animal farms, more dimensions have to be taken into consideration. The hypothesis was that such an operation is more difficult without the aid of labels. On the other hand, the category "farm animals", even if more sophisticated semantically speaking, can be formed on the basis of the more general representation of "things related to a farm", which entails that less cognitive control, and indeed less selective ability, is needed. The assumption, then, is that language helps with controlling and selecting the representations that are not automatically grouped together. The hypothesis was confirmed, as aphasic subjects performed more poorly with low-dimensional categories. Moreover, naming performance in a naming task was positively correlated with performance in the categorization trials regardless of lesion site, thus suggesting, according to the authors, that the effect was not due to an overlapping of areas dedicated to cognitive control and areas involved in the task,but that there is a causal link between naming abilities and categorization of the low-dimensional kind. Once again, this supports the idea of language as *enhancing* cognitive processes.

Similar results were achieved with a verbal interference task on normal subjects (Lupyan, 2009). In the first experiment, participants performed an odd-one-out categorization task with and without verbal interference, where they had to exclude the object that did not belong to the group, on the basis of color, size or theme (e. g. "birthday party"). The experiment was conducted with both words and pictures as stimuli, and verbal interference was a number-rehearsal task. Results showed that verbal interference slowed down overall performance, but further analysis highlighted that verbal interference significantly increased reaction times for color and for size, but not for theme, thus supporting the hypothesis that language helps with isolating perceptual dimensions and using them for categorization. The results were further confirmed in the second experiment, where the interference task was a spatial one (remembering how a series of dots was arranged), and no effect was found, suggesting that it is indeed the verbal interference that causes a delay in the categorization.

In Yoshida and Smith (2005), Japanese children were tested in a situation where they were given redundant labels, to see if the role of language can be extended beyond simple naming. Japanese does not mark the difference between objects and substances in the same way as English, which means that the mass-count distinction is not grammaticalized in the same way (Erbach et al., 2017). Children were presented in training with couples of novel objects,

named with the same label, that had shape in common in the case of solid objects and material in the case of non-solid objects. In the linguistic clues condition, different markers were also used for the solid and non-solid pairs. The idea was that learning a correlation between a perceptual cue and category structure in the context of a redundant linguistic cue led to better performance. In the test, children were presented with an entity, told its name, and then asked to pick, from among three other entities, the one that shared the name. (The label was a novel one.) The idea was that, if redundant linguistic cues can help the formation of categories, being aided by the cues in the training phase would have led to better generalization in the choice of the name for a novel solid thing on the basis of shape, and for a non-solid thing on the basis of material. This prediction was confirmed by the results, suggesting that teaching the correlation between words and perceptual properties can bootstrap category learning.

Arithmetic is a more abstract domain where some results were achieved in proving that language has an influence[3]. In a study by Pica et al. (2004), speakers of Mundurukù, which is a language spoken in Amazonia that does not have numbers bigger than 5, fail in some simple basic arithmetic tasks with numbers above 5, despite their ability to perform other comparisons between numbers without problems. The authors discuss the possibility of relating these results to the idea that humans have two counting systems in a fashion similar to that which will be explored later in treating the theory set out in Carey (2009). Fundamentally, the idea is that language allows for counting routines that require a one-to-one pairing of objects and representation, which allows more efficient numerical representations used in arithmetic.

Fascinating data comes from the Pirahã speaking community in Brazil (Everett, 2005b). This community has remained consistently monolingual and, in this sense, isolated linguistically, to the point that claims by Everett have been the object of several controversies. These are mostly due to the fact that the linguist most familiar with the language, Daniel Everett, has claimed that the existence of Pirahã, which he defines as *non-recursive*, is the ultimate proof that many assumptions of the Chomskyan theoretical framework are wrong. This goes well beyond the scope of this book, but the data should be mentioned because of

3 The debate regarding numerical cognition is incredibly vast and complex, and it should not be addressed lightly. Later in this chapter, I will present a theory belonging to Susan Carey (Carey, 2009) that makes numerical cognition a core example of the interaction between language and non-linguistic cognitive resources. For a review of relevant findings in numeral cognition concerning both human and non-human knowledge, a useful read is Hubbard et al. (2008), where many different empirical findings are reported. For a full introduction to the debate, the most useful source is probably Kadosh and Dowker (2015).

its interest[4]. Color terms have been discussed in Pirahã in both Everett (Everett, 2005b) and Everett (2005a), where the main disagreement is over the existence of color words themselves: whether we can speak about basic color terms in this language, or just about phrases used as color terms, is then a matter of discussion. Some relevant data for the issue of language and cognition, however, come from numerical cognition; Pirahã has a very limited counting system, with two words denoting loosely *one* and *two* (Gordon, 2007) and an additional word for "many", accompanied by finger counting. In a series of experiments, Gordon (2007) tested Pirahã counting abilities through simple matching tasks: participants had to replicate the arrangement of sticks laid out by the experimenters by, for example, putting batteries in a one-to-one correspondence with arrays of nuts varying in orientation and arrangement complexity. Results showed a very limited ability to perform in these tasks when the quantity was raised above 3, with difficulties increasing as the quantities increased. Interestingly, it seems that participants improved on some tasks when the number of elements exceeded 8, compared to their very low performance when elements were between 3 and 8: this seems to suggest, according to Gordon, that they were able to employ strategies like *chunking* the units in larger units (3 elements, for example), to meet task demands. This is in line with the fact that non-human animals and language-impaired humans seem to use an *analog* strategy (Nieder and Miller, 2004), not based on counting, which is also employed by typically developed humans with large quantities, with limited accuracy (Agrillo et al., 2012). It is therefore possible that, lacking another strategy (which Gordon calls parallel individuation), Pirahã speakers employ analog individuation. At the same time, their lack of a counting system seems to imply difficulties with processing medium quantities. A fundamental point about these findings is that they seem to suggest that, in this case, language is not only enhancing performance in already-present mechanisms, but actually providing the mind with a new cognitive tool for performing arithmetic tasks.

In the same direction, a study by Dehaene et al. (1999) found that training Russian-English bilinguals in exact number addition in one of the two languages led to faster addition only in a subsequent task using the same language as the training. As Wolff and Holmes (2010) notices, this suggests the influence of lin-

4 However, it should be noticed that these controversies are very important for this kind of research. It could be argued that Everett's research about the Pirahã community is not rigorous enough, and that a database for data collection should be available in order to verify his claims. Work in this domain and on the language has been carried out by Gordon (Gordon, 2007) too, which might be considered a confirmation that something relevant can be said about the Pirahã number system, and relevantly more research was presented in Everett (2017).

guistic coding on the task, and even a *language as an augmenter* effect, but not that the cognitive process was definitely occurring in linguistic form.

Motion verbs and action events are another area that is often explored in relation to the language-thought debate. Relevant work was done by Boroditsky et al. (2002) comparing Indonesian and English. While English, along with many Western languages, has obligatory tense marking, i.e. it is specified by the verb whether the action is taking place now, will take place in the future, or has already taken place, the same is not true for Indonesian. The language has temporal words available to mark tense, as it is easy to imagine, but these are optional and not included in the verb. In the first experiments, participants had to look at pictures of actors who were either about to perform an action (e.g. preparing to kick a ball), performing an action (kicking a ball) or had just performed one (having already touched the rolling ball with a foot). Two images were shown at the same time, with different actors at different action stages, and participants had to rate their similarity. English speakers significantly rated images at the same action stage as more similar than Indonesian speakers. In the subsequent experiment, bilingual Indonesian-English speakers were tested in both languages separately. Interestingly, bilinguals rated same-action-phase (so, same tense) pictures as more similar when tested in English, and different-tense pictures as less similar when tested in English. In the third experiment, a memory task was administered, where English and Indonesian speakers were shown an actor at a particular stage in the action and, subsequently, asked to choose, from among the pictures representing all three stages, the one matching the previous picture. English speakers showed better memory performance, and the result was confirmed in the last experiment, where bilinguals showed better performance when tested in English as opposed to Indonesian. Note that this difference in performance for bilinguals is interesting because it leaves room for interpretation: while it seems to indicate that a *language-as-a-meddler* effect could be taking place in aiding memory, one might also argue that the results are compatible with what is sometimes called the *inner speech* hypothesis, i.e. one might argue that the tasks were performed with some linguistic encoding. However, the authors interpret the results as suggesting that setting up the task in a particular language, despite the task being non-linguistic, would prime a thinking modality. Similarly, in a study by Papafragou et al. (2008), eye-tracking revealed that Greek speakers, watching animations of people moving, were more focused on *path* than *manner*, and that English speakers tended to do reliably the opposite. This is attributed by the authors to the fact that, contrary to English, Greek encodes path (entering and exiting, for example) more than manner, contrary to English (for instance, jogging, rolling, and marching). Importantly, this effect was found both when participants were told they would have to

talk about the video afterwards, and when they were not instructed in this sense, albeit in the second case the differences in eye movement seemed to appear towards the end. This suggested, according to the authors, that language was acting as a *meddler* in the coding even when linguistic code was not prompted. The relation between the inner speech hypothesis and the views on language and thought will be discussed later in this chapter, and put in relation with this kind of data.

The spatial domain has been largely investigated in this respect too, deserving more space than can be devoted to it here. However, some main results can be mentioned. Levinson explores in a series of studies (Levinson, 1997, 2003) the different frames of reference used in several languages. Frames of reference refer to coordinate systems used to specify the location of objects in relation to the speaker or other objects. Languages differ dramatically in the frames used, varying from Absolute Frames of reference, which use fixed points in the universe (cardinal systems, but also landmarks like coasts or areas inland), Intrinsic Frames of reference, which are related to the features of the object considered (and based on volumetric properties, for example, or functional criteria, like the front of a car), and Relative Frames of reference (for example, relative to the speaker: in front of, left of, right of, all keeping the speaker as a referent). In a large variety of experiments, language speakers using different frames of reference were tested in non-linguistic tasks (Levinson, 1997, 2003, 1996); in most of the studies, participants were shown a spatial relation (i. e. two objects in a specific relation to one another) or a trajectory of an object and were then rotated 180 degrees to perform a spatial task. These tasks were either memory tasks, where the configuration or the path had to be remembered, or inferential tasks. For example, participants had to observe two objects, A and B, in a certain relation to one another. Subsequently, after rotating 180 degrees, they observed object B positioned in relation to object C. Finally, after rotating again, participants had to position object A and C in a way that was consistent with what they had been shown. Participants tended to apply in these tasks the same frames of references that are preferred in their own language, despite the fact that no linguistic coding was necessary to perform the purely spatial tasks. The experiments altogether confirmed Levinson's hypothesis, that different frames of reference used in language entail different spatial coding, which is reflected in cognitive tasks involving spatial cognition and therefore indicate the role of language in non-linguistic cognition.[5]

5 Li and Gleitman (2002) have addressed these studies maintaining that they do not support Levinson's hypothesis. However, Levinson et al. (2002) argue very convincingly that Li and Gleit-

More cross-linguistic evidence comes from studies on acquisition of verbs and nouns in Korean and English children. Korean has a very rich verbal morphology, where endings of verbs mark significant semantic components. Moreover, the presence of nouns is often non-obligatory, with the possibility of ellipsis. English presents the opposite pattern; while verbal morphology is extremely simple, nouns are mostly obligatory. In a series of studies, children speaking these languages were studied to confirm that verb morphology develops earlier in Korean learners than English ones, and that English children use more and more varied names (Choi, 1995; Choi and Gopnik, 1995). In a longitudinal study, Korean children and English children were tested in categorization tasks and means-ends tasks that involved the use of an "action insight", e. g. using a rake to move an object. The hypothesis was that English children would be better in categorization tasks, given their familiarity with naming and nouns, and that Korean speakers would be better in the problem solving tasks involved in understanding action performance, which was confirmed by the results. According to the authors, this shows that linguistic coding of actions and categories has a direct impact on conceptual acquisition related to actions and categorization skills.

A final domain that is worth mentioning is that of space and time. In a series of experiments, Casasanto et al. (2004) explored differences in time metaphors. English and Indonesian, among others, use as a preferred metaphor for time linear distance (long, short time) whereas Greek and Spanish tend to use quantity metaphors (much, little time). In a line-growing experiment, participants had to observe lines growing across a screen, and then either estimate their duration (clicking one time, waiting the appropriate amount of time, and re-clicking when they thought they had replicated the duration of the stimuli) or estimating their length (clicking, moving the cursor as far as they thought was necessary, and clicking again). In a container task, participants saw empty containers filled gradually with lines symbolizing water, and had to estimate either the amount of water or the amount of time the container would take to fill (always by clicking appropriately). English and Indonesian speakers were strongly influenced by the length of the line when estimating the time for the growing line task, whereas such an effect was not found for Greek and Spanish speakers. Conversely, Greek and Spanish speakers were influenced by the volume of the container when judging the time for the container task. The results were interpreted as showing that language might be responsible for *conceptual mappings* between

man dramatically misunderstand the use of frames of reference and fail to produce a valid counterargument to Levinson's point.

spatial and temporal information in a way that consistently changes cognitive performance.

2.3.2 Limits of the influence of language

As already pointed out, many of the effects of language on non-linguistic cognitive processes seem to be limited: they are task dependent, or related to specific requirements and domains. This, as argued, does not mean discarding the effect as non-interesting. However, it is worth focusing on some inherent limits of the influence of language on cognitive processes.

The fact that the influence of language does not seem to run as "deep" as it possibly could is not the only reason why one should be careful in evaluating the role of language in cognition. A rather obvious point, which has been mentioned by prominent voices in the debate regarding cognition (Bermúdez, 2005; Fodor, 1983), concerns infants, language-impaired adults, and non-human animals. The case for language-impaired adults is, in one sense, the trickiest. There is good evidence that many important cognitive processes are left intact in subjects suffering from strong cases of aphasia, as is documented at length in Donald (1991) and in different domains by Varley and Siegal (2000); Siegal et al. (2001). Dennett (1991), on this subject, points out that aphasic patients have been *exposed* to language before, and used to master it: as a consequence, their cognitive skills might have been permanently shaped by language in a way that makes their loss of linguistic abilities irrelevant to the evaluation of the role of language in cognition.[6] In the following chapters, the case of aphasia will be dealt with in more detail, but it is worth mentioning that, in an extensive review, Fedorenko and Varley (2017) list a long series of studies that show how aphasic patients can solve logic problems, solve navigation tasks, and process music like healthy individuals, which suggests that natural language abilities and thought can be dissociated and thus are independent from each other.

A different case is posed by infant cognition and animal cognition, and these cases have been presented as evidence that thinking cannot be fundamentally linguistic (and as a case for Fodor (1975) hypothesis, which I will mention in the next few paragraphs). Infants show signs of categorization abilities way before they learn language (Rakison and Yermolayeva, 2010). They also show very

6 Note that Dennett addresses Donald's claim, but not Varley and Siegal's studies, which were published after Dennett's discussion. However, it is, I believe, a safe speculation to assume that he would present the same argument, since the point made by the authors is rather similar.

early development of fundamentally sophisticated joint attentional skills (Eilan et al., 2005; Tomasello, 1995). Pointing behavior, which is of fundamental importance in communication, emerges during the preverbal stage (Carpenter, 2009). More saliently, non-human animals are able to engage in rather sophisticated cognitive activities, as is the case for counting in birds (Hirai and Jitsumori, 2009), social behavior in dogs, primates (Soproni et al., 2002; Call and Tomasello, 2008; Pack and Herman, 2006) and categorization for both dogs and primates (Range et al., 2008; Vauclair, 2002). Fodor underlines that either one decides not to see any continuity between non-human animals, infants and speaking adults, or "[...] some thinking, at least, isn't done in English" (Fodor, 1975). I believe that this point is rather uncontroversial: however, as it will be shown in the next sections, there are theories that do attribute to language a very significant role, which go behind recognizing the possibility of language to "interfere" with other cognitive processes. In what follows, I will analyze theories of the interface between language and thought that, while recognizing the role of language, vary on which kind of role they in fact attribute to it.

2.3 Theories about language and cognition

2.4.1 Communicative views: Fodor for all

The main purpose of the empirical review in the previous part of the chapter was to point out that there exists some interaction between linguistic skills and cognitive processes; what still needs to be addressed is the issue of how to characterize this relationship.

This does not only imply discarding a view that assigns no role to language (more on this below). It also implies giving language the *adequate* role. At the end of the previous section, I summarized some motivations for discarding what is often called the "inner speech" hypothesis. The simplified version of this view is that thinking occurs in natural language. This can be interpreted in two ways: in a weak sense, *some* thinking occurs in natural language. This view is not in contradiction with the data exposed above, and a version of this claim will be addressed below. In a strong sense, the view entails that most of what is called "higher cognition" occurs in natural language. In the rest of the chapters, some theories will be presented that attribute to language a similar role. In general, it is worth pointing out that the view that Natural Language (NL) is constitutive of thinking develops, in a sense, as a reaction to a position that is virtually at the other extreme of the spectrum, which is the Language of Thought view developed by Fodor (Fodor, 1975, 1978, 1985, 1987,

2008). While a complete analysis of the Language Of Thought proposal is beyond the scope of this work, I will now highlight which aspects of the theory are essential to the issue and why, in its original formulation and taken with all the implications, the Language Of Thought Hypothesis might be considered incompatible with the data presented above.

According to the original LOT (Language Of Thought) proposal, the vehicles of propositional attitude contents are physical structures, composed of re-combinable and separable components that are in a one-to-one correspondence with the structure of the sentence expressing the content of that belief. In other words, the vehicle of a belief has the structure of a natural language sentence. Language Of Thought, then, has a language-like structure; however, it would be a big mistake to identify LOT with natural language. To further unpack the claim, let us consider the summary that Egan (1991) provides of the tenets of LOT, recalling the two central claims exposed by Fodor (1987):

1. Propositional attitudes are relations to meaning-bearing tokens;
2. The representational tokens in question are quasi-linguistic.

As Egan (1991) notices, 2 means that the symbols processed are "quasi-sentential", meaning they have semantically evaluable sub-parts as constituents. Egan (1991) also lists the kind of requirements that empirical evidence has to fulfill to actually support LOT:

1. The mental representations involved have to be assumed to have psychological reality;
2. Such representations have to have appropriate content, meaning they "must be interpreted *in the theory* as the contents of independently ascribable propositional attitudes" (Egan, 1991, p. 384);
3. The representations have to be language-like.

Fundamentally, as Bermúdez (2005) points out, the gist of the LOT Hypothesis (LOTH) is that for an agent to have an intentional attitude towards a state of affairs means for the agent to direct attention to the state in question through vehicles of thought in the appropriate way, and the appropriate way is, in this case, a sentence-like structure. This is what is implied by the above criteria individuated by Egan (1991) as well. This view is connected to the idea of an *isomorphism* between the *vehicle* and *content* of propositional attitudes: this is, according to LOTH, what grants a casual dimension to propositional attitudes. Naturally, this is not a neutral assumption: on the contrary, it can be disputed at length, both by calling into question the role that *intentional content* has to play in mental representations (Egan, 1991) or by calling into question the role of propositional attitudes (Matthews, 2007). What is fundamental for this chapter's purpose,

however, is not this aspect of LOTH, but rather the view regarding language that accompanies it in the original formulation, namely the view that natural language has little if any role to play. As specified, thought is supposed to make use of language-like structures. However, this cannot be a natural language structure; Fodor relies on the arguments mentioned above, like the necessity of attributing the possibility of thought to non-human animals and infants, to argue against this idea. Language Of Thought is rather, in the Fodorian view, the *means* through which natural language is acquired; in a sense, LOT is necessary for the acquisition of NL. Thinking, according to LOTH, means tokening representations with syntactic structure and the appropriate semantics: this happens in an internal symbol system that, albeit similar to NL in having a syntax and a semantics, has nothing to do with the language acquired by the child during development. This can also be connected to the strong modularist view that Fodor proposes: there is a *language module* that, like any module understood is a strong sense (see for instance Bermúdez (2005) for a discussion on what features of modules are essential in Fodor's understanding of cognition), is *encapsulated* : this implies that the module does not have access to information used in other modules, and does not interact with the world-knowledge a subject has. This, in other words, means that there is no direct interaction between knowledge understood as categorical or perceptual knowledge and knowledge as linguistic knowledge. NL is only useful, in order words, to *communicate* thoughts that pertain to a completely different domain.

This might seem like an extreme position to hold, but one should notice that it is far from being unique. While his contribution to theories of communication is the most important and discussed legacy, Grice (1957, 1968, 1989) holds a position that is similar in attributing to language a role that is exclusively communicative. The function of language is, evolutionary speaking, that of making possible a correspondence between the psychological states of different people (Grice, 1982): language is used to influence each other's mental states and thoughts, and this is done by communicating the content of the internal mental states. In this sense, language comes *after* thought and it is merely a tool to express its content; it does not have any role in determining it[7]

7 Note that there is a strong sense in which this picture of communication is not only challenged in respect to a more substantial role for language in thinking, but also as a result of a more general concern with the way communication works. Assuming that language is a means for reproducing our thoughts, while appealing in many ways, also implies that communication is a mostly conscious, intentional process that follows quite specific interaction rules. This is arguably not always the case; see Clark (1998); Brennan and Clark (1996); Horton and Gerrig (2005, 2016); Horton and Keysar (1996).

This quick detour around the classic Fodorian take on language and thought should have made it clear why, despite its objections to the idea of NL as LOT, Fodor's assumptions about the role of language as an encapsulated, communication-only device are not compatible with the data presented in the empirical review section of this chapter. Empirical evidence from a relatively large variety of experiments seem to point in a different direction, since, far from being encapsulated, linguistic information has at least some role in many non-linguistic tasks. While this should be clear, it is worth underlining that this is not an argument against everything that LOTH entails: what I have just listed as a motivation for discarding the communicative view of language Fodor proposes does not touch a series of fundamental issues involved in LOTH, such as the representationalism it entails, the modularism that accompanies it, and the causal validity of folk psychology as an explanation. As will be seen, it is possible to maintain that language has a fundamental role in cognition and still hold that LOT is part of how cognition works. Before diving into theories of this kind, I will briefly underline why a view at the opposite end of the spectrum to communicative views is worth rejecting.

As anticipated, the *inner speech hypothesis* is partially best understood in opposition to LOTH, but has a very different scope. While the Language Of Thought Hypothesis is, as the name suggests, a hypothesis regarding cognitive architecture in general, the *inner speech hypothesis* is a thesis about *propositional thinking*, understood as thinking that (1) can be expressed in propositions, and (2) can be evaluated for truth conditions (Bermúdez, 2005). This kind of cognitive process might not cover, for example, calculating whether or not a car fits the garage. The idea of the inner speech hypothesis, then, is that, while a calculation of space based on perceptual space and some kind of simulation might be possible without language, the same is not true for thoughts that are subject to inferences and logical transitions. In this case, we need to have access to a linguistic code to formulate this kind of thinking. This implies at least two things. First, it implies a strong discontinuity between mental processes, since, while propositional thinking occurs in NL, non-propositional thinking has another format of representation, whatever that is. This might imply a very strong distinction between, for example, semantic-processes and perceptual and motor processes. While this might sound like a truism, one should be careful in assuming that this is a neutral position, as many theories in the tradition of embodied cognition would straightforwardly reject this idea[8]. The second implication of such a position is also a discontinuity, as it implies attributing to non-human animals

8 As an example and for an overview, see Zipoli Caiani (2010).

and infants only one kind of thinking, i. e. non-propositional thinking. This is a salient contrast with LOTH; as has been seen, a strong point in Fodor's idea is the possibility of attributing the same format of thinking to linguistic and non-linguistic agents. While this might be acceptable for many, one issue with the inner speech hypothesis is definitely the implication it has for human adults whose language capability is impaired. What does that imply for their thinking processes? While something along the lines of the inner speech hypothesis has been developed by Sellars (1997) and Carruthers (1998)[9], I will focus, in what follows, on the most prominent account that attributes to language such a strong role in cognition, namely the one developed in Carruthers (2002)[10]. His theory will open the section dedicated to theoretical frameworks that try to account for the empirical data presented at the beginning of this chapter.

2.4.2 Carruthers and LF: language as conscious and/or cross-modular thought

It is important to see Carruthers' idea of language as *constitutive* of thought as specifying that only conscious, propositional thought occurs in NL. Carruthers (1998) already contains the core of the proposal, since it is there argued that certain patterns of thinking and reasoning, which are acquired linguistically, are only available in language form. However, the relation is slightly more complicated, since Carruthers makes a distinction between what is called Logical Form (LF), on the one hand, and NL on the other.[11]

A point that is worth stressing from the very beginning is that Carruthers attributes a relevant role to syntax, which is fundamental for various theories of the interaction between language and social cognition that will be presented later. LF, as suggested by the name, has a language-like syntax. Consequently, it could be described as language without a phonological representation. The proposal by Carruthers is then that this representational format of thinking is

9 And might be present in Wittgenstein (1953), according to Bermúdez (2005).

10 As will be seen, however, this might not be classified fully as an instance of the *inner speech hypothesis*, since it is not always natural language in *all* of its features that plays a role in cognition.

11 Note that Carruthers slightly varies his view frequently and has fairly recently declared that *conscious thought* might be an illusion (Carruthers, 2017). It is not completely clear how the various views Carruthers held through the years are connected, but I will here focus on the most detailed accounts of the relation between language and thought that he has put forward, namely (Carruthers, 2002, 2013a).

a core process in some forms of cognition: *some* thoughts have linguistic structure, and in particular, while thoughts in LF can have a linguistic structure but can lack a phonological realization, they are not identifiable with NL – i.e. they do not occur in inner speech, despite their syntax-like form. On the other hand, LF thoughts can also have a phonological realization, and hence be conscious, and therefore be identifiable with inner speech. In those cases, thinking actively occurs in natural language.

To better understand the claim at stake, it is essential to specify that the thesis presented in Carruthers (2002) rests on the assumption of modularism, inherited from Fodor (Fodor, 1983). The idea of modularism is that the mind can be divided into subcognitive systems that have at least a few of the following characteristics:

- Domain specificity
- Mandatory operation
- Limited central accessibility
- Fast processing
- Informational encapsulation
- "Shallow" outputs
- Fixed neural architecture
- Characteristic and specific breakdown patterns
- Characteristic ontogenetic pace and sequencing

While not all of these features have to be present, fundamental features of modules as used by Carruthers are *domain specificity* and *fixed neural architecture*. Without going into too many details about the many "modularisms" that are possible in the literature, it is essential to remark here that Carruthers assumes central-process modularism; in addition to standard input-output modules, there are also *conceptual modules*, taking conceptual input and delivering conceptual output.

Carruthers also assumes a distinction between modular thinking and intramodular thinking, i.e. thinking that is not domain-specific but integrates different kinds of information from different modules and allows for sophisticated forms of reasoning. In addition to such a system, Carruthers assumes the existence of a *pre-language practical module:* this would have, as input, "raw" desire-like and belief-like states of the form DESIRE[Y] or BELIEF[IF X THEN Y], with *x* being an action for which a motor program is available. The *output* of the pre-language practical reasoning systems will be indexical in form, and is going to be able to produce *intentions* of actions, i.e. to provide for the basis of goal-directed thinking.

However, humans are also able to integrate information coming from different modules, i. e. to integrate information that is not domain-specific. Carruthers' examples are, in this case, geometry. Compared to rats, humans seem to be able to integrate geometrical and non-geometrical (related to colors, for example, or smell) information when performing a search task (Cheng, 1986). To Carruthers, this is a sign of inter-modular thinking: the *geometry* module deals with information that is different from the other modules involved in the task. This cognitive ability must have evolved at some point in human history, and Carruthers argues that it was absent in *Homo erectus* and archaic forms of *Homo sapiens* (citing Mithen (1990)); hence, the argument goes, it must have evolved either after language or at the same time as language. The first option is deemed implausible by Carruthers because of time-constraints; the second option, then, is what is considered most plausible, assuming that language and the ability to integrate information from different cognitive modules must have evolved at the same time.

According to the argument, then:

> [...] it is hard to discern what the separate selection pressures might have been, which would have led to the development of two distinct faculties at about the same time (language and domain-general thought), when just one would serve. (Carruthers, 2002, p.17)

Consequently, the use of language *presupposes* input from various modules, which further leads to Carruthers' idea that language is indeed the medium of the non-domain-specific cognitive ability he talks about.

The "loop" depicted implies that language plays a double role; it can compute the input received by conceptual modules and form sentences in LF; and it can serve as input for other conceptual modules when a phonological representation of the sentence is given, forming "inner speech". This in particular is processed when neural pathways dedicated to the hearing of actual speech are recruited to generate a "quasi-auditory input" (Carruthers, 2002, p.27). Quoting:

> So the suggestion is that language, by virtue of its role in unifying the outputs of conceptual modules, and by virtue of our capacity for auditory imagination, can be used to generate cycles of central-modular activity, hence recruiting the resources of a range of specialized central-modular systems in seeking solutions to problems. This may be one of the main sources of the cognitive flexibility and adaptability which is so distinctive of our species. (Carruthers, 2002, P.27)

Central process modules, on the other hand, will take input from perception, in terms of *conceptualized perceptions* and propositional descriptions, which derive from linguistic input. The hypothesis, then, is that the *output* of central-process modules are sent to a central cognitive mechanism that is dominated by lan-

guage; according to the hypothesis, all non-domain-specific reasoning of a non-practical sort is conducted in language. The format of the resulting representations is LF, the *logical form* advocated by Chomsky (Chomsky, 1995), i.e. a level of linguistic representation interfacing between the language system and other cognitive systems. In Carruthers' model, cross-modular thinking, i.e. the cognitive domain dealing with the output of the central process modules, happens in LF. Language, then, is an input-output module that receives input from conceptual modules and transform them in speech and receives speech input and presents it to the other modules in a format that can be accepted by the conceptual modules.

Most importantly, Carruthers underlines how his account is far from a view like LOTH and in general from communication-only views of language, since it avoids the classic formulation-of-thought + linguistic resources that convert the thought in NL. This is because the thought itself cannot be formulated without NL, in this picture.

A rather different spin, but one that goes in the same direction, is given in Carruthers (2006, 2012, 2013a, 2013b), where Carruthers tackles the issue of what are called "dual system" theories. Dual system theories emerge in the literature as solutions to some of the problems emerging with LOTH and other monolithic accounts of cognition: as will be shown in later chapters, this kind of solution is adoped in, for example, the theoretical literature on social cognition. What is relevant about Carruthers' new proposal is that, in the version of the theory, language is no longer called into question for conscious thought, but is instead considered more like a "tool" used for thinking and, as I will explain, *rehearsing* mental processes.

The general idea of a dual system theory is a division between processes that are fast, automatic, and shared with non-human animals (System 1) and processes that are slow, flexible, and potentially more complex, which are the prerogative of humans (System 2). Depending on the version (see Evans (2008) for an overview), System 1 can be thought of as a collection of modules, fast and mandatory, and it is sometimes considered innate. On the other hand, System 2 usually is not modular, it is acquired with experience and social interaction, and dependent on language or other higher cognitive skills. System 2 is often considered domain-general, able to deal with abstract representations, sometimes conscious, and linked to general intelligence. In the case of the account presented by Evans and Stanovich (2013), System 2 (or, as they name it, Type 2 processing) is linked to working memory and the ability to simulate counterfactual scenarios. Carruthers criticizes the general idea of a dual system theory on the basis of a worry regarding the kind of cognitive architecture that the division entails. If System 1 is mandatory and automatic, it seems like it would constantly

be the default choice, for it is the most cognitively economical one, in requiring less processing power and less cognitive resources. However, it is unclear if in this picture System 1 only gets involved in reparative strategies, and why such a system would evolve in the first place. As I will explain in chapter 5, this is very similar to the objection raised by De Bruin and Newen (2012) to dual system theories of mindreading: there seems to be something problematic in assuming that two systems which have different functions and do not communicate with one another should be involved in the same processes. Carruther's solution, then, is to classify System 2 as operating within the same resources of System 1: the more sophisticated and flexible system, then, operates on the result of a variety of cycles of the more automatic system. System 2 *rehearses* System 1 cycles, connecting the information resulting from different System 1 modules. A fundamental role is given to language, in this picture, since language is the format of representation that links together the results of different encapsulated modules, making it possible for System 2 to operate. In this sense, System 2 can be put in continuity with System 1, because it uses many of the same resources. The potential theoretical advantage of this is, then, that it explains rather concisely the difference between non-human and human cognition. In Carruthers' view, language provides a new representational format. Animals can generate "images of actions": they can use their cognitive resources not only to engage in goal-directed activity, but also to consider the consequences of an action they would perform. Having generated these "simulations" of action, they can then engage in practical reasoning. Humans, however, can also generate *phonological* images of linguistic actions, i. e. they can engage in inner speech, which bears a cognitive advantage.

The story goes like this: the language production module receives a command for producing a string, but in this case it is *quasi* executed: there is a phonological image of speech (but no utterance happens). This phonological image is given as input to the language comprehension module, which *decodes* it and gives it as input to other conceptual modules available.

There are several remarks to be made about Carruthers' view. Firstly, it is the heavy reliance on modularity that makes the account not necessarily the most appealing one: as Carruthers himself points out, the system stands or falls with empirical proof of modularity, which is not uncontroversial. (See, for an example of a full-blown critique of modularity, Tomasello (2014).) However, this is not the only concern.

One kind of worry is underlined by Tillas (2015a), who points out that aphasia seems to be problematic for an account that relies on this kind of division of labor between language and non-linguistic thought. The studies by Varley and Siegal (2000) seem to point very clearly to the fact that aphasic patients,

whose mastery of language is significantly limited in many ways, have intact abilities even in sophisticated mental tasks like mindreading and false belief reasoning. (More on this in chapter 4.) The constitutive role of language in System 2 seems to be in contrast with this piece of evidence or, at least, it seems to deserve some reconsideration after studies regarding language deficit.[12] In Carruthers (2002), the idea that an LF account can be maintained by adopting a diachronic rather than a synchronic view of the role of language is briefly defended, but it is not completely clear how this applies to the view of language as *lingua franca* of cognition. While this may be disputable, there are other worries.

The second kind of worry is related to the nature of language itself, and it is partially envisaged by Beaulac (2014). He points out that language is conceived as rather monolithic in this picture: although there is a division between a comprehension and a production submodule, this might be a rather simplifying picture implying a rigid distinction between processing and production processes: it is not completely clear how this would be compatible with the (empirically funded) models of language that see an interaction between language production and language comprehension (for example Pickering and Garrod (2013)). Moreover, it seems to entail a vision of language as a system which never operates automatically, fundamentally different from any other module, and whose computational strength relies only on the possibility of producing strings with propositional structure. However, there are good reasons to think that the linguistic faculty emerges from many different systems and sub-systems that rely on different processes, including memory, executive control, and so on. Dividing between a comprehension and a production system seems to be limiting in the sense that, even if only for the production of a phonological-like string, what is probably needed is efference copies resulting from motor commands. This is not necessarily in contradiction with what Carruthers suggests but, as Beaulac (2014) underlines, there is a need for further specification of how different cognitive resources employed in language participate within the mechanism as a whole.

Thirdly, a rather relevant point is raised by Machery (2008); it is not clear how language can be the "lingua franca" of cognition if, in this model, none of the other modules has the possibility of "parsing" the content of the representations that are produced by the language module itself. For, if the conceptual modules have domain restrictions, as they would in such a strict modular architecture, and if the format provided by language is different in representational format, it is not clear how this kind of information can then be used by other

12 In chapter 4, several studies that shed light on these issues will be presented.

modules that are not language-based. As Machery points out, a plausible reply might be that the modules in question have these representations already in their conceptual repertoire. However, this comes with granting that language modified the conceptual repertoire of the modules, thus making the cycles of inner speech unnecessary, since the modules already possess the conceptual resources for interacting.

Finally, whether or not the theory effectively explains the empirical data presented above is not clear. Some of the tasks used in the studies explained above were purely perceptual or visual, which implies that an explanation would be needed of why the processes involved are to be ascribed to different encapsulated modules that cannot interact without a linguistic code as Carruthers predicts. Some of the tasks described above seem to entail abilities that are available without linguistic information, but where labels or other bits of language modify the performance. Moreover, this interference seems to be contingent on the task and, in some cases, to have to do with lateralization too, which is not straightforwardly explainable in Carruthers' view. Some of these points will become more relevant after more data, i. e. data regarding the interaction between linguistic skills and social cognitive ones, is presented. For now, we will focus on accounts according to which the role of language is "more modest".

2.4.3 Associating language and concepts

The previous section explored Carruthers' account, which can be seen as an account on which at least some cognitive processes are carried out in natural language (or in some linguistic form related to it). In what follows, I will focus instead on theories that go in another direction, exploring the interaction between linguistic processes and non-linguistic ones rather than assuming an identity between natural language and (some) forms of thinking. As will be shown, going for this option implies, on the one hand, scaling down the influence of language. On the other hand, it means accounting for a variety of different effects on cognition in a way that is flexible enough to fit with some of the presented data.

In what follows, I will firstly briefly present an account of the interaction between language and thought that has a large scope, LASSO, and subsequently present a hypothesis with a narrower scope, i. e. LHF. The reason for presenting them together, as will be seen, is that they make similar predictions and present similar advantages.

2.4.3.1 LASSO: Associations between labels and concepts

Tillas presents his account of the interaction between language and thought as partially overlapping with Lupyan's one (exposed in 2.4.3.2) and partially as a counter-proposal to Carruthers idea, which he names among the most prominent hypothesis about a *constitutive* role of language in cognition.

Now, LASSO (Labels and ASSOciations) relies on associations of three different kinds:

- Word-word associations
- Word-concept associations
- Concept-concept associations

Firstly, there is a level of interlexical association.[13] Words are associated with each other; these links are present thanks to frequency of co-occurrence. This claim is neutral in many respects, since it is common to both associationist accounts and Fodorian-views (for instance, Fodor (1987)); however, while modularism is assumed in a classic Fodorian perspective, these linguistic associations are connected to the conceptual level in LASSO. As a matter of fact, following Hume (1748), concepts are associated with each other as well; in LASSO, concepts form a network where the connections are weighted, and where the strength of the weightings determines activation patterns in the network. This is possible following the Hebbian learning principle (Hebb, 1949), according to which *what fires together, wires together*: in other words, frequent co-activation of two different concepts establishes a link between them, which is what grants an association. Note that repetition of stimuli is not only essential for the connection between different concepts, but also for the formation of the concepts themselves, at least in LASSO; the idea is that, far from being innate, concepts form through abstraction from several occurrences of the same stimuli or situation (Tillas, 2014, 2015b). The storage of the representations is driven by mechanisms of selective attention that focus on determined features of the stimuli; selective attention also influences in a top-down manner how new representations are stored, guiding perception, which implies that a series of representations will be stored in the same memory locus. (Tillas appeals here to Perry (2001)'s mental folders.) A category is then formed by abstraction on the basis of the features across the stored representations in the same focus.

Aside from the details of the empiricist account, available in Tillas (2014, 2015b), what makes this account non-modularist, and what is central for the issue of language providing a representational format, is that the association

13 Note that this kind of association is also normally assumed in studies about lexical priming.

network at the conceptual level and at the linguistic level are not independent, but in communication thanks to the lexical-conceptual association. Words act as labels for concepts, following the same Hebbian learning mechanism based on activation and frequency; this means that the activation of the phonological representation of a word can trigger the activation of a concept, which can sub-activate a connected concept, and so on.

Fundamentally, syntactic information is not lost because words do not solely act as labels for concepts, but also contain information about the structures they most frequently appear in, and they form an association net that results from exposure and use of natural language. In this sense, information about language use and information about word-concept association are actually part of the same network; the relation between syntactic and semantic information is "flat" and mostly built probabilistically. It is through continuous activation of the same clusters of units in the net that we end up with reliable links between words and non-linguistic constructions, i. e. conceptual structures.

Crucially, this kind of relation is what plays a fundamental role in solving the issue that every account of thought that does not identify it with NL has to face, i. e. propositional thought. While an account like LOT naturally provides an answer to how thought can have propositional form (as thought is conceived as fundamentally *having* propositional form), the same cannot be said for an account like LASSO. One solution is, naturally, to assume a modified version of LOT, assuming that the combinatorial rules can be applied to representations that are not innate. However, this is not the path chosen by Tillas. The answer, in LASSO, lies in the possibility that thought has to "*piggyback* " on language in the structured content.

The following quote is useful for unpacking the claim:

> Thus, given that sentences are unified syntactically structured entities, they unify and structure concepts associated with their components into a propositional thought in a way that mirrors their unity and structure. A thought gets to have propositional content in virtue of concepts (for objects or features) being associated with individual words or phrases; the sentence provides a kind of unity. (Tillas, 2015a, p. 227)

Hence, the association between the concept and the word ends up having more than one relevant role: in a way, it is what connects the semantic interpretation of that word to the phonological form. In another sense, it also provides the "piggybacking" mechanism that is, in the model, responsible for the compositional nature of thought, without making thought itself propositional. Notice that, in this sense, Tillas' view is highly in line with approaches like that in Camp (2009), where is argued that language provides a possible insight into the combinatorial capacities of thought. In LASSO, the idea is rather that language pro-

vides a mechanism for acquiring *endogenous* control over thought. In other words, while conscious thought without language is possible, language is necessary for directing and controlling conscious thought. This entails, naturally, discarding the possibility of a *constitutional* relation between language and thinking. What is gained, on the other hand, is an account of how "raw" associations can still be part of an account that predicts propositional thought.

The key detail of an association-based analysis like the one above is the notion that linguistic labels can carry information about the structures they appear in – in other words, syntactic information has a place in the net. This is what forms the foundation of the claim, that language provides enough scaffolding for thought to be propositionally structured.

> A thought gets to have propositional content in virtue of concepts (for objects or features) being associated with individual words or phrases; the sentence provides a kind of unity. (Tillas, 2015a, p.27)

An essential feature of the analysis provided by Tillas is that language is not thought to be the *only* possible way in which thought can be endogenously controlled, but just one way among others. In non-human animals and pre-verbal children, for example, associations between representations and goal-directed actions might guarantee endogenous control, since actions are indeed something over which an agent can exercise control without linguistic symbols. On the other hand, to acquire endogenous control tokening on a given concept through language means being able to *activate* a determined representation in the absence of its referent through the use of linguistic information, i.e. a label.

It has been shown above that Carruthers' idea commits us to a version of modularism: on the contrary, Tillas' association account seems to commit us to two different assumptions. The first is that of the absence of modularism. Language and thought (and more generally information of a perceptual, motor and conceptual kind) interact, in the presented account, in a way that is more dynamic. Encapsulated modular processing is not compatible with the idea, of connectionist inspiration, that labels and conceptual representations interact on an associationist base. The other view that is tightly related with LASSO is that thought in general has an associationist nature. In this sense, LASSO's claim is twofold: associations dominate the relation between labels and concepts, on the one hand. But associations also dominate thinking, on the other hand, which is the reason why the view is fundamentally Humean. While the associationist picture presented here has a certain unity, it is not necessary to adopt both claims to see an advantage in conceptualizing mental representations as associated with their labels in a productive way. To better make this point, I will

present a more modest view in the literature, the LFH, that also relies on associations, and makes (arguably) fewer assumptions on cognitive architecture in general.

2.4.3.2 Label-Feedback Hypothesis

LFH is the proposal made by Lupyan (Lupyan, 2012), which bears several similarities with LASSO. In a nutshell, LFH claims that verbal labels play an active role in perception and categorization processes. Lupyan calls this function of labels a *top-down augmentation* of perceptual representation: language guides the activation to the diagnostic features of categories that are labeled. This is not a full-blown theory of the interaction between Language and Thought, but rather a specific hypothesis about how categorization processes can be affected by linguistic labels. As such, it will be presented rather briefly but put in relation with the other theoretical accounts. However, notice that this mechanism, while being conceptually very close to LASSO, is also ultimately compatible (and in line) with the assumptions made by other accounts, as will be outlined below.

As underlined by Tillas (2015a) himself, there are no consistent empirical prediction differences between the two accounts, which differ more in terms of generic scope. This is because LFH is a specific hypothesis about the effects of language labels on categorization, and LASSO focuses on the more general relation between language and cognition. Moreover, in Tillas (2015a) it is argued that LFH does not recognize non-verbal representations. In Lupyan (2012) it is briefly mentioned that "the distinction between verbal and non-verbal representations becomes moot"; however, this is meant mostly to underline how the fact that linguistic effects that are online and contingent to task are not necessarily marginal, and hence that there is an intimate relationship between categorization and labels. Lupyan clearly considers perception to be a mechanism that does deal with representations, and categorization as a more general cognitive process that is not necessarily always linguistic. In this sense, I think Tillas (2015a) remark is intended to draw attention to the fact that Lupyan does not specify what the difference between the two is, nor what the relation between concepts (that he considers at least partially verbally determined) and perceptual representations is. Another potentially unspecified issue regards how a more general interaction between linguistic and non-linguistic representation would work in cases that are not straightforwardly part of the categorization problem he treats.

Importantly, Lupyan stresses frequently how the role of labels recognized by LFH is *transient*, i. e. neither permanent nor to be classified in what has been previously defined as "deep whorfianism". The position, as a matter of fact, is prob-

ably best classified as holding a "language as meddler" stance, where linguistic information can play a role in determined tasks (and even enhance the performance in some cases), thus being a case of "language as an enhancer", but the effect is still temporary and flexible. This kind of effect is what Lalumera (2014), as seen, considers theoretically interesting.

In Lupyan's view, naming (i.e. attributing labels to categories) is one of the many possible categorization processes. Drawing on evidence presented by Goldstone et al. (2001), Lupyan relies on one assumption: categorization can affect perception. In a series of experiments (Goldstone et al., 2001; Goldstone, 1994), Goldstone and colleagues found that training subjects in categorization tasks (e.g. distinguishing between individuals that belong to a club and those who don't, or sorting stimuli in different categories according to some dimension, like brightness, size or a combination of the two) lead to changes in performance in a visual discrimination tasks where subjects had to decide whether or not two stimuli were the same, with increased focus and ability (a "warping effect", in Goldstone words) for the relevant perceptual dimensions, making subjects more sensitive to cases that were closer to category boundaries. Language, Lupyan argues, is one of the ways to implement categorization; consequently, it has the potential to affect perception in a similar way.

More specifically, labels in Lupyan's view facilitate "*transient* modulation of ongoing perceptual (and higher-level) processing" (Lupyan, 2012, p.4). In particular:

> In the case of color, this means that after learning that certain colors are called "green", the perceptual representations activated by a green-colored object become warped by top-down feedback as the verbal label "green" is co-activated. This results in a temporary warping of the perceptual space with greens pushed closer together and/or greens being dragged further from non-greens. Viewing a green object becomes a hybrid visual-linguistic experience. Knowing that some colors are called green means that our everyday experiences of seeing become affected by the verbal term, which in turn makes the visual representation more categorical. (Lupyan, 2012, p.4)

Labels have the function of *re-aligning* representations in categorical perception. This is possible through a mechanism that sees connections between labels and perceptual categories which go in two directions, with feedforward and feedback connections. Between the perceptual layer and the label layer, a hidden layer is present.

A label, according to the model, activates corresponding features, thus pulling apart representations and making the relevant perceptual features more easily identifiable. During training, a model like the one described learns to associate a given label ("chair", in Lupyan's example) with an instance of a relevant

member of the category (e. g. a chair). It also learns to activate perceptual properties related to a chair upon hearing the word "chair". Some of the properties of a chair will be more strongly correlated with being a member of that category than others, i. e. there will be more chairs that have backs than chairs that are brown, so the label will be strongly associated with having a back, as a perceptual feature, than with being brown. In a feedback loop, hearing a label will help the perceptual process to focus on those features that are more strongly connected with it. The perceptual representation, then, will be modulated by verbal labels. The model supposedly also explains cases of verbal interference affecting the influence of labels (as seen for example in Winawer et al. (2006)); verbal activity disables the connection between label and the "hidden layer", not allowing for the label to guide the perception process top-down. Note that this is very similar to the process explained in section 2.4.3.1, where a similar mechanism of association between linguistic labels and perceptual experience was described. On this account, labels guide perception in a top-down manner according to a feedback mechanism.

2.4.3.3 Associations: advantages and limits

The two accounts above rely on a direct relation between concepts and labels, which is dramatically different from what has been proposed by Carruthers. A reason to be inclined to accept such a picture is that it seems to fit rather well with some of the empirical evidence reported above. In both LASSO and LHF, labels are not the only way to access conceptual information, but they do have a facilitating effect, being linked to perceptual information and conceptual representation in a fashion that allows for activation nets.

The fact that these accounts argue that an intervention of language is not mandatory, but possible for determined tasks, is their strength. However, the picture they depict does not come for free, as one has to give up modularity for both LASSO and LHF, and strong neo-empiricist assumptions are necessary in LASSO's case.

Moreover, while it is clear that association links might provide the right mechanism for the effects described in categorization studies (see section 2.2 above), it is less obvious how they can account for effects on the integration between spatial and temporal information seen in studies like Casasanto et al. (2004). Is the association between labels and the conceptual representation of space what makes this "metaphorical mapping" between space estimation and time estimation mediated by language possible? Perhaps. There is a sense in which Carruthers idea, that language provides a medium for cross-modular thinking, seems to be more appealing in this respect. A mechanism that could

also do the job is the piggybacking relation between labels and conceptual structures envisaged by Tillas (2015a), since language might be the format that is responsible for the associations between temporal concepts and space concepts, for example. This, of course, would need to be further elaborated.

2.4.4 "Re-wiring hypothesis"

This section is concerned with another "family" of theories regarding language and thought, which has been at times labeled the *rewiring hypothesis* (Bermúdez, 2005). The label is meant to indicate that the assumption informing these theories is that language is a game-changer in cognition, on both a phylogenic and ontogenic level. The view has been developed in Karmiloff-Smith (1994), and has the most prominent precursor in Vygotsky (1962), but it is best represented contemporarily by Clark (Clark, 1996; Lupyan and Clark, 2015; Clark, 2006), which is the theory I will focus on. Even if Clark does not describe his idea as a form of rewiring hypothesis, I will use the label here to indicate the family of theories that see language as providing essential new cognitive tools for cognition. In any case, it is worth mentioning some of the other theories' strengths, in order to better contextualize Clark's account.

According to Karmiloff-Smith (1994), a function of language is to provide a format that keeps together information of different kinds, which is a point similar to that made by Carruthers (2002). General takeaways from a view such as Karmiloff-Smith's are that (1) language provides a storing representational format that is different from other cognitive coding, and (2) language can make these bodies of information, re-encoded, subject to new thoughts, in a *meta-representation format*. According to this theory, this allows for *self-monitoring*. Consider the two interpretations of this claim given by Bermúdez (2005):

- At the personal level, this would mean that we cannot engage in conscious and reflective self-monitoring without language;
- At the sub-personal level, this would imply that cognitive systems not participating in public language cannot engage in self-monitoring.

A similar view is also endorsed by Clark, who suggests that second-order cognitive dynamics and processes can only manipulate natural-language-like representations, which can be abstract and fixed in meaning, not context dependent and amodal. Importantly, none of these features are possible in Clark's view without language, because conceptual representations that are not linguistic, in his view, are fundamentally embodied. Similarly to the view proposed by Tillas (2015a), however, the rewiring hypothesis gives us a way to satisfy the nec-

essary conditions of LOT for productive systematic thinking without committing to the LOT hypothesis – this would mean, according to Bermúdez's interpretation (Bermúdez, 2005), that the real proving ground is what the differences between language-endowed and non-linguistic animals are. LOT's prediction is that animals and humans potentially both have a language of thought, and in this sense have a similar cognitive architecture in many respects. This is not the case for the rewiring hypothesis, since the idea is that language has shaped the evolution of the human mind as much as it shapes the development and growth of the human newborn brain. This, as already pointed out, also constitutes the main argument that Fodor presents against the rewiring hypothesis: the fact that certain kinds of thinking are clearly available for non-linguistic animals (children, and non-human animals), but that this still (allegedly) requires lots of what LOT has to offer in terms of representational means. Note that, under this description of the rewiring hypothesis, while the first view offered by Carruthers is substantially different in expecting a *constitutive* role of language for *conscious* thinking, the proposal that language fundamentally enables System 2 operations over System 1 results is, on the other hand, very similar.

This kind of approach has a precursor in Vygotsky (1962), according to which language acts as *cognitive scaffolding*: higher mental functions are made possible by language. On the one hand, language and other cognitive abilities have separated developmental paths. On the other hand, these paths converge in the learning child when language starts being a means through which inner speech can aid, for example, problem solving. In this case, it is overt interaction with caregivers and engagement in language that provides a cognitive boost to the child: instructions and help from the caregivers provide the learning child with examples of cognitive achievements and with an aid, language, that is actively used to solve various tasks. Hence, Vygotsky considered overt self-directed speech as a fundamental aid in cognitive development, and its direct evolution, inner speech, as playing a fundamental role in adult cognition as well.

Let us go now to a more modern interpretation of the rewiring hypothesis. Clark's view, available in Clark (1996, 2006, 2013) is often described as a standard example of an *anti-expressivist* view (Sutton, 2002; Tillas, 2015a), since it openly challenges the idea that the most prominent function of language is to merely express someone's thoughts, which is, as underlined, something that is openly supported by accounts like Fodor's. In Clark's view, language is a computation-transforming instrument that expands the domain of what is subject to human thought. Recalling the classification in 2.1, this view would be classified as seeing language as *enhancing* cognition, but perhaps most relevantly as a view that sees language as attributing new cognitive power. Note that more recently, views concerned with the scaffolding function of thought have been de-

veloped, for example, in Dove (2017); I will focus on Clark as a prominent representative of this approach.

Words have, for Clark, the value of *tools:* this is a strong point of similarity with other theories that have been developed against the backdrop of embodied cognition, like WAT (Borghi and Binkofski, 2014). In both frameworks, linguistic labels exercise the role of *glue* for thought: in Clark's case, this is reflected in the ability of labels to make thoughts possible objects for further consideration and operations in computation. While representations are not linguistic in form (which strongly differentiates this view from Carruthers' one), linguistic input given by labels allows one to form representations that are more easily recallable and reusable: this allows for memory processes to be more efficient. (This role for language as "gaining control" over thoughts in terms of memory processes is also essential in LASSO, as seen in section 2.4.3.1.) Additionally, thanks to language, more complex and structured representations are formed. Language is then conceived as complementing thought.

In Clark (2006), an important point is made that should not be overlooked: in this version of the rewiring hypothesis, words are not only relevant for cognition because of their content, i. e. because of the conceptual units they are related to. They are also relevant in their vehicle/physicality, i. e. the fact that they are perceivable in speech and readable on paper (or on screen). In this sense, language is for Clark a *cognitive niche:* it enables the building of physical structures that make thinking and reasoning more efficient. Consequently, there is a double role for language: on the one hand, to activate other representational structures (in a fashion similar to that of LHF or LASSO), and on the other hand to provide a representational structure itself.

There are at least three ways, Clark argues, in which language can significantly enhance our cognition processes:

1. *Labelling as short-cut*. Labeling something makes available a new perceptual representation (the sound) to be associated with other perceptual input. This can facilitate the retrieval of information connected to the labels in a way that makes the retrieval faster and more efficient;
2. *Hybrid thoughts*. Language can *combine* with previously existent and language-independent resources for the formation of new skills. An example is number cognition, where the non-linguistic capacities of approximating magnitudes and recognizing small quantities are combined with linguistic number systems to produce new skills, like the ability to think of an unlimited set of quantities[14];

14 The quantification example is also essential for Carey (2009); I go into further detail below.

3. *Words as anchors*. The ability to *think about thinking* is, in Clark's picture, a good candidate for a language-only ability. Formulating our thoughts on paper or in words makes it possible to have a physical format of representation for them that is otherwise unavailable, and creates an *object* that can be considered for further thinking. This idea will become especially relevant in chapters 4 and 5.

Clark's view can also be better comprehended by addressing his more general proposal about predictive processing. In Lupyan and Clark (2015), it is claimed that words not only act as tools, but also as *artificial contexts*. Predictive processing is a view of cognitive architecture that dramatically differs from the Fodorian take presented above, and is much closer to the connectivist and distributed-parallel processing idea. According to the model, proposed in Clark (2013) among others, the best way to think about cognitive processes is in terms of forward and backward flows of information. *Percepts*, i. e. perceptual representations, are the result of *top-down* predictions the framework makes on the basis of previously acquired knowledge and expectations; these predictions are confronted with incoming physical stimuli, i. e. sensory data, and updated to generate better predictions. These predictions are not always going to perfectly match reality, which is why the *bottom-up* flow of information carries prediction error signals, which in turn select better top-down predictions, and so on. The system is conceived in terms of hierarchical signaling processes, where cycles of these predictions and failed expectations constantly update the information used. In this framework, linguistic input acts as additional information that influences the predictions. This is because of various features: firstly, words are associated with concepts that are thought of as being (in the kind of framework pushed by Clark) connected to specific experiences. Hence, my representation of dog is related to specific experiences I have related to dogs. Hearing or pronouncing the word "dog", then, activates not only expectations related to the category of dog, but plausibly also information coming from different experiences I associate with the word. Secondly, words are found in syntactic environments, and can be modified: when hearing "small dog", I also activate expectations related to size.

These kinds of context have an essential role in shaping the kind of expectations that enter the predictive processing top-down flow of information. In this sense:

> Language directed at others and at oneself (e. g., in verbal rehearsal and other forms of self-directed speech) provides a powerful tool for manipulating thought and reasoning. The main mechanism by which language accomplishes this manipulation is through flexible

> modification of both what top-down information is brought to bear, and (by selectively influencing the precision-weighting of prediction error) how much influence this top-down information has on specific lower-level level processes. Language can thus help constrain what representations are recruited, and what impact they have on reasoning and inference. On this view, language becomes a powerful tool for cognitive self-manipulation, providing a huge boost to intelligence. (Lupyan and Clark, 2015, p.7)

In a predictive coding system, information is recruited in a context-dependent, flexible way, depending on what the contextual clues generate as predictions, and on the information that is already stored in the system in virtue of previous experience. Hence, language can act as an *anchor*. Since new information acquired in the system is automatically clustered with previously existent information, there is a sense, according to Clark, in which advanced thinking requires inference and reliable "trajectories in representational space" (Clark, 2006, p.372). Words and sentences in natural language, then, do exactly that, constituting the way we reliably operate on clusters of representations.

The account provided by Clark seems to cover most of the empirical data presented above, in having at least the *potential* to explain the results in categorization and perceptual tasks, along with the influence of language in spatial mapping tasks. This should not come as a surprise, since Clark's account is the most "generous" as far as the role of language in cognition is concerned. In the picture of language as tool that Clark delineates, linguistic information has the double function of assisting perceptual information in lower cognition (via directing attention, and aiding memory) and of providing a new, more abstract format of representation for other tasks that are presented as essentially impossible without the aid of language. In this case, while there are no strong assumptions about modularism, there is indeed a heavy cognitive load on language. In the concluding section, I will highlight some of the advantages and shortcomings of this kind of stance in terms of empirical and theoretical investigation.

2.4.5 Carey (2009) and Gopnik (2001): theories, concepts and language

An account that is worth citing, even if it's concerned with the restricted case of conceptual development, is that presented by Carey in her book "The Origins of Concepts" (2009). The book presents an influential proposal regarding how conceptual resources are acquired. However, Carey also makes a very specific proposal regarding the role of language in what she calls *Quinean bootstrapping*. In parallel to this account, I will also briefly present Gopnik's proposal for the theory-theory. The reason for listing these theories in this chapter is that they ad-

dress the problem of the interaction of language with thought in a way that is different from the other perspectives explained in the previous sections. The focus is, in this case, not on linguistic coding interacting with perception processes, but rather on conceptual scaffolding. However, this is not irrelevant to the theories presented so far, given that Carey and Gopnik's accounts both address questions related to how language interacts with pre-existing resources. Although the theories differ in a significant sense, they share some core assumptions and approaches that I will underline.

Firstly, some specifications should be made about Carey's project. The idea has been described by Gopnik (2011) as a proposal to bring two different traditions together; on the one hand, the empiricist tradition, which sees concepts as acquired representations, and the innativist conception exemplified by Fodor. Carey's proposal can be summarized in three claims:

1. Core cognition is identifiable with innate modular perceptual-input devices;
2. There are two types of conceptual representations, those embedded in core cognition and those that are embedded in explicit knowledge systems;
3. New representational resources emerge in development, over time, and they emerge as a result of *Quinean bootstrapping*.

A central assumption in Carey's work is that of *discontinuity* : it is assumed that there are fundamental differences in the representational repertoire of children and adults, and these discontinuities can be explained in terms of a change in the representational abilities. The puzzle that emerges, then, concerns how these new representational capacities come about: assuming that there are innate conceptual primitives but also that there are discontinuities in representational abilities in development means having to explain how these "gaps" get filled.

The answer to the question "How do children acquire a new representational format?" lies for Carey in the human ability to formulate explicit symbolization. In this sense, language is naturally involved. The ability that humans have to create symbols that are not related to anything in the world is a key feature of Quinean bootstrapping: the idea is that these symbols, i. e. words and strings of language, allow the child to establish mental symbols that correspond to these explicit, socially determined symbols. These linguistic symbols will not be mapped for meaning to the mental symbols automatically; rather, they will acquire their first interpretation in relation to one another, and so they are called by Carey "placeholders", and they form "placeholder structures". These structures are then mapped by the child to the limited conceptual repertoire that she already possesses. Subsequently, a modeling process allows the child to more ef-

ficiently map the conceptual representations to the symbolic representations, to the point that, via induction, a new representation is formed.

The best way to understand Carey's proposal is to use an example of the bootstrapping mechanism she advocates.[15] The idea is that children come equipped with more than one innate system for dealing with counting and mathematical representations. There are three systems of representation with numerical content in core cognition, which are parallel to the individuation of small sets of entities in working-memory content, analog magnitude representations of number, and set-based quantification. These abilities are thought to be innate and to belong to the conceptual repertoire of the child, which is in line with what has been argued briefly above about mathematical representations in non-human animals, for example (see 2.3.2). However, Carey argues, the child's system is not yet able to represent integers: relying on a large body of empirical work, she maintains that, while children learn counting routines very early, they do not learn that counting represents cardinal values of sets.

This is achieved by Quinean bootstrapping. Initially, the string of number words is learned and only understood in terms of relations among the units: one comes before two, three comes after two, and so on. A partial meaning then gets assigned by the child to this list, mapping "one" to a single individual, and the word "one" is understood as applicable to sets that can be put in 1–1 correspondence with it, and the same applies to "two" and so on. Importantly, the counting list is acquired linguistically and this placeholder list is understood and memorized separately. Then the bootstrapping mechanism proceeds as follows:

> The child notices the identity between the singular, dual, trial, and quadral markers and the first four words in the count list. The child must try to align these two independent structures. The critical analogy is between order on the list and order in a series of sets related by additional individuals. This analogy supports the induction that any two successive numerals will refer to sets such that the numeral farther in the list picks out a set that is 1 greater than that earlier in the list. (Carey, 2011b, 8.1)

This counts for Carey as the acquisition of new representational means, as the child can now *divide* numbers, represent cardinal values of sets, imagine quantities indefinitely, and so on. Thanks to the bootstrapping mechanism, the child acquires a new ability, which is only possible in virtue of the symbolic structure built with language.

---

15 This is also the way that Carey develops her argument in her book, which focuses on case studies.

The power of the symbolic representations available in language is even more evident and more relevant when complex concepts like *weight* are concerned. Children's minds, Carey argues, do not possess the same concept of weight which adults possess: it is *degree of heaviness* that can be attributed to entities, usually big ones, but is not conceived of as a property of any object. It is with complex mapping between numerical relations and formulas that children can arrive at the concept of weight. Given this is, arguably, a mathematical concept of a quite abstract nature, it might not come as a surprise that it involves language. However, relying on the educational resources used in actual educational curricula, Carey provides a plausible story for how the concept of weight is learned through representational language and mapping to the physical properties of objects. In the curricula she cites, children are first familiarized with empirical explanations of phenomena, e.g. they are told that objects of the same size can have different weights, and they are brought to formulate thought experiments that compare sizes and estimation of weight. The child for example is forced to acknowledge that, if a feather weighs 0 grams, then there is a contradiction in thinking that 50 feathers do indeed have some weight. Subsequently, the child is familiarized with different systems of representations: boxes of a constant size are accompanied by dots. The child is brought to familiarize herself with the fact that she can derive the number of boxes from the number of dots and the number of dots per box, and vice-versa, experimenting in this way with the way the three measures are related to each other. Density, at this point, is represented as the relation between weight and volume, passing first through the visual and perceptual representation of the relation between boxes and dots, and then through the linguistic formula D= W/V. This is initially understood, thanks to the combinatorial properties of language, as expressing relations among the symbols it employs, and only subsequently mapped to the independent representation of dots and boxes through an analogy process. In this way, the more complex concepts of weight and density are acquired.

This kind of process in development is mirrored in historical processes regarding more complex concepts, like the formulation of relevant concepts in mathematical and astronomical sciences, and Carey explains at length how the two kind of processes rely on roughly the same analogies. While children acquire the placeholder structures from language and interaction with adults, scientists pose them as theoretical entities; this is the case, for example, with the idea that the sun causes the motion of the earth. This idea was a placeholder, Carey argues, for Kepler, until an analogy with magnetism allowed him to formulate the more elaborated concept of *vis motrix*, which was not possible before. While this analogy between theory-making in science and learning might

sound peculiar,it is as a matter of fact not unusual among theories of cognitive development, e. g. Gopnik (2001).

Gopnik is one of the many proponents of the theory-theory take on concepts, whose main tenet is that conceptual acquisition and use are to be understood in light of what could be named the "scientist assumption": children operate like "little scientists", testing hypothesis regarding their understanding of the world (and their understanding of the meaning of words) and adjusting their theory, or model, according to the results of their experiment. This process of hypothesis formulation and subsequent testing is thought to be *analogous* to theory formation as carried out by scientists.

As do many others, Gopnik recognizes the role that labels can have in grouping together perceptually different stimuli in categories, i. e. using the same label "animal" for so many different instances of different looking beings. The fact that a certain word is used in a different context is then a "clue" that something similar underlies the nature of the named entities. This is confirmed by studies that prove that linguistic similarities act as clues in children's induction of similarity between objects (Gelman and Coley, 1991).

However, the role of language in more complex conceptual categories is deemed more general. According to Gopnik, language acts as *evidence* that aids theory formation and theory change. While in early infancy children are acquiring evidence from observation of the environment and of other agents, language provides a new source of evidence. This allows the child to get evidence that has been collected from others: in a sense, this means less cognitive effort and less need for perceptual and observational evidence. On the other hand, Gopnik argues, this will also introduce more variation in the information possessed by individual children, given that the linguistic stimuli is more variable.

Gopnik characterizes her position as a weaker version of Whorfianism; language does shape cognitive resources and behavior, she argues, but does so in a way that it is not as "anti-realist" as she sees Whorfian positions as being. This is because children are already equipped with theories and conceptual means to navigate the perceptual and conceptual world: however, language constitutes a new kind of evidence. Additionally, Gopnik classifies cases in which linguistic information seems to be influencing non-linguistic domains, like spatial language, as a case in which, while language does contribute to theory formation, it can do so because none of the theories is itself "better" than the others. In other words, she argues that these effects of language on cognition are possible because the linguistic evidence is not in conflict with other kinds of evidence and perceptual information. While Greek and English might induce slightly different "theories on movement", then, the results are not such that any of the resulting systems will be more effective than the other in general: depending on

the situation, one might be more suitable for a task, but they will stay roughly equivalent in value. One "deeper" effect of language in theory making is hypothesized by Gopnik as far as moral or social concepts are concerned; these, being related as they are to socio-cultural norms, and acquired linguistically, will be more prone to change.

The two views presented in this section are in many ways different, but do share more than one assumption. Firstly, as Carey herself underlies, theory-making as a cognitive tool is central to both proposals, to the extent that it is part of the cognitive processes that underlie conceptual acquisition and use. Secondly, both Carey and Gopnik assume that a lot of the conceptual resources are innate, and that a computational view of the mind is indeed apt to describe the interaction between non-linguistic and linguistic abilities. Finally, the role of language in these views is somehow limited. Although Gopnik recognises a role for language in acquisition and thought, the role attributed to it in lower-level cognitive processing is limited. Moreover, while Carey stresses the fact that language plays a fundamental role in her conceptual development proposal, there is also a sense in which her emphasis, more than on language *per se*, is on the ability to create symbol systems. While this is undeniably part of what is remarkable about language and linguistic abilities in human, Carey explicitly recognizes, in her reply to Gopnik and other commentators (Carey, 2011a), that there is a sense in which the symbolic ability does not need to be linguistic in nature, and that her account grants at least the logical possibility that this might not happen through *public* symbolic representation, but through mental symbols. Carey ends up rejecting this possibility, but there is a sense in which this is relevant to the present discussion, as Carey's consideration of language relies on one dimension, that of symbolic representation. While this implies a view to some extent similar to that proposed by Clark, where language boosts and enhances cognitive abilities, it does leave us with nothing particularly relevant to say about other possible kinds of interaction between language and thought, like those that seem to be at work as far as categorization processes or perceptual tasks are concerned. This, as seen, is not the case with other accounts, and thus makes Carey's proposal rather unsatisfactory for the purpose of this analysis. However, Carey's hypothesis does shed some light on the fact that assuming LOT does not necessarily imply discarding a significant role for language. More on this will be said in the concluding section.

2.5 Conclusions

This chapter has given an overview of what the theoretical context is for the debate about language and thought. After presenting the main coordinates of the debate and highlighted how a pluralistic approach to the problem of the interface between language and thought can be useful in disentangling the claims and empirical findings, I have given a description of some of the relevant studies that have been produced. The overview included a focus on color cognition and concluded with data from other cognitive domains. Subsequently, I have presented some prominent theories about language and thought.

The choice of the theories has been determined by several factors. Firstly, I selected theories that are representative of the possible spectrum of positions. While Carruthers assumes a constitutional relation between language and some kind of thought, Clark considers language to be a fundamental tool for the advancement of cognition. These two positions are rather clear in pointing in the direction of a deep role for language in cognition, but they do so in rather different ways: for Clark, this presence is pervasive, since language rewires the brain in a significant way, providing it with tools, short cuts, and enhancements that are transformative of the way the human brain processes information in a large variety of tasks. For Carruthers, language has a rather specific role, operating on the output from otherwise encapsulated, self-sufficient modules. This shift in perspective is relevant, because it brings with it different predictions, as natural language is present in Clark's view as a non-modular cognitive tool, whereas LF operates under the restriction of a modular system.

Tillas and Lupyan present associations between labels and non-verbal representations as a fundamental feature of the role of language. While Tillas adds to this a crucial feature, i. e. the role of language as providing thought with a propositional structure, the same cannot be said about the Label Feedback Hypothesis, which focuses instead on a specific mechanism. While very specific, the mechanism actually effectively explains many of the findings above. It is worth mentioning that the view Lupyan and Clark push together, in Lupyan and Clark (2015), integrates this role of labels in a picture of the mind, based on predictive processing, where language does have the reshaping role that Clark predicts. In this sense, LFH and LASSO, while sharing many assumptions, depart on the general cognitive picture.

Finally, Carey and Gopnik's theories are worth mentioning because, while sticking to a Language of Thought idea of cognition, they recognize an active role for language in the realm of conceptual acquisition and development. While the theory-theory stance that they both take is rather in line with a computational, traditional view of cognition, it attributes to language a role that is

much more significant than that of a simple communication tool. At the same time, compared to the other views analyzed above, the role they attribute to language acquisition is minimal and clearly confined to specific domains.

2.6 Theories, predictions and assumptions

There is one thing this analysis reveals, which is that putting these views on a spectrum depending on how "deep" the role of language is can be useful for explanatory purposes, but also somehow simplifying. In the table below, I highlight some of the assumptions and predictions made by the theories:

Table 2.1: Theories' claims and predictions.

	Language directs attention	Language aids memory	Language "directs" thought	Language as rep. format	Language as propositional thought
LASSO	√	√	√	x	√
LFH	√	√	√	x	x
"Inner speech"	?	√	?	√	√
Carey	x	x	x	√	√
Gopnik	√	x	√	x	√
Clark	√	√	√	√	√

A fundamental message resulting from the above analysis of the theories is that none of them comes for free: it is, trivially, impossible to choose a theory of the interaction between language and thought without subscribing to the (dramatically non-neutral) theoretical assumptions that these theories make about cognition as a whole. While this book is not meant to solve the rather challenging issue of how human brains work in general, it is meant to cast light on how the available theories, including their theoretical assumptions, do when it comes to explaining the data regarding the interaction of language and thought.

It has emerged as a result of the discussion in this chapter that, while weaker claims about the role of language, like those by Gopnik and Carey, might be sufficient to explain some effects described by the empirical literature (for example, the fact that arithmetic skills seem to depend at least partially on linguistic development), they do not cover all the cases outlined above: it is hard to see how

these theories would account for a role for language in lower-cognition tasks like visual searching. At the same time, Carruthers' claim, that language has a predominant role in certain kinds of thought, also does not extend to cases in which language seems to be fundamental in directing attention and enhancing memory. This job seems to be easier for Tillas, Clark, and Lupyan, who all assume a prominent role for language in many lower-level tasks, and also recognize its role as a tool for acquiring different representational abilities and new skills.

While the main focus of this book is how the acquisition of mental state terms aids social cognition, one of the results of this work will be an assessment of how this topic can be positioned in the current literature on language and thought. As a consequence, once the literature regarding mental state term acquisition and mentalizing skills is analyzed, I will briefly return to assess some of these theories.

2.6.1 Open questions

Most of the empirical results that have been mentioned in the review section have focused on semantic domains that are rather concrete. This is the case for color cognition and for spatial tasks especially. In these cases, it seems that labels and linguistic information exert an influence on a variety of tasks, many of which are not supposed to entail linguistic coding. This is telling, because it points in the direction of an empirical research program that continues to investigate what the possible non-linguistic tasks can be with respect to which linguistically-acquired information can determine a change in performance. Clearly, a lot remains to be investigated. On the one hand, the fact that many studies reveal lateralization effects might be telling: it might be a sign that, while the effect of language is indeed relevant, it is also strictly dependent on whether or not the left hemisphere is recruited.

Cross-linguistic studies have enormous potential to reveal crucial insights into how linguistic information influences memory tasks as well as visual search tasks, and this is definitely a direction that needs to be explored further. Luckily, most of the research goes exactly in this direction. Now that the ghost of a radicalized Whorfian view, where speakers from different languages think in completely different ways, has been defeated, research is ready to recognize that some features of our communicative systems might relevantly influence how we remember, store, categorize, and form expectations about the external world. While most of the research focused on arithmetic, color and spatial infor-

mation, there are good reasons to think that these studies will effectively keep exploring what the role of language is in various cognitive domains.

Lesion studies, as mentioned, are of fundamental importance: however, they also should be considered with care, and they are intrinsically more difficult to come across. While it is always dangerous to generalize a correspondence that exists between a single lesion in a single patient and performance in a certain task, it is undeniably useful to come into possession of data that so strongly points in the direction of a *functional* relation and does not rely on mere correlational data. In this sense, the ground is open for new research that will undoubtedly shape and inform the debate.

As mentioned, however, most of these semantic domains have been concrete. There is a fundamental insight that emerges when analyzing Carey's and Gopnik's position: not only it is important to distinguish which processes can be potentially influenced by language. It is also essential to keep in mind that, while perception tasks and visual searching *might* be influenced by language, there are other cognitive skills that seem to be independent from linguistic abilities. For Carey and Gopnik, this entails at least the acquisition of very abstract concepts and skills, like full-blown arithmetic and complex concepts regarding social relations. While it is easier to account for some lower-level cognition skills, which we clearly share with many other animals, in terms of non-linguistic cognitive resources, it is harder to straightforwardly exclude a role for language in the development of other cognitive skills.

This chapter has revolved around a specific issue: how complex the relation between language and thought seems to be, once one accepts that different cognitive mechanisms and different skills will be differently influenced by linguistic information. If this is true for rather concrete domains like the ones analyzed above, it is clear that matters will be even more complex in the case of a specific domain. The very complex domain that will be the focus of the rest of this book is that of the relation between the acquisition of specific linguistic information, i. e. that related to mental states (verbs in particular), and the development of social skills, specifically those involved in the attribution of mental states. In what follows, I will firstly present an up-to-date review about the acquisition of these lexical items, and the development of their use in children. Subsequently, I will present the empirical studies that have been conducted to investigate the relation between the acquisition of mental state verbs and nouns and the development of social cognition, followed by a review of the philosophical and psychological theories that address these issues. Finally, I will present my own account of how language can influence the development of these cognitive skills. At the end of the book, the theories presented in the current chapter will be situated with respect to the results of these analyses.

Part II: Talking about minds, thinking about minds: mental terms acquisition and the role of language in mentalizing

Chapter 3:
Mental state verbs and constructions: acquisition timeline, semantics and syntactic occurrences

This chapter is dedicated to the most relevant characteristics of mental terms, including their semantic features, the syntactic structures they appear in, and their acquisition timeline. Given that the purpose of this book is to investigate the relation between the acquisition of language and cognitive development, and that I focus especially on a particular class of verbs and linguistic units, I think it is essential to have an overview of what these units are and how they behave in language structures. As will be explained in later chapters, in which the theories regarding mental term acquisition and cognitive development will be analyzed, the nature of mental state verbs and their syntactic behavior is a central concern for some frameworks.

Most of the research regarding mental state verbs focuses on English, and the same applies to research regarding conceptual acquisition; this is the reason why I will use the English language as a baseline for the analysis. However, I will integrate cross-linguistic data in the analysis whenever possible, for different reasons. The first reason is a methodological and deontological one, as it is important for me to avoid (or to mitigate) an anglocentric bias, when possible, and to declare its presence, when unavoidable. This kind of problem becomes of central importance when empirical research into psychological phenomena is the subject matter. A study titled "The Weirdest People in the World?" (Heinrich et al., 2010) has presented a meta-analysis that shows that a big part of behavioral psychology research has been focused on what is referred to as WEIRD people – Western, Educated, Industrialized, Rich, and Democratic parts of the world population. This bias occurs naturally in psychological studies, when the subjects recruited for testing are college students, i. e. mostly Western, or Western-raised, come from industrialized countries, are rich enough to afford an education and (mostly, at least superficially) democratic. This would not be a huge concern if it was not for the fact that most psychological studies are presented as testing hypotheses and theories whose validity is allegedly universal. In general, an advantage of typological linguistic research is that it brings a great theoretical and methodological contribution to the table, allowing us to enrich the analysis with other languages, populations, and environments.

The second concern is a theoretical one: cognitive theories that highlight the role of language acquisition in cognitive development have to be concerned with cross-linguistic data in order for their claims to be universal and solid, and to

https://doi.org/10.1515/9783110748475-004

make sure their generalizations can be applied at a level that concerns the faculty of language in general, and not only the most studied and culturally hegemonic language.

In a similar way, and as a third reason, language relativity, as explained in the previous chapters, is a claim that rests on the idea that different languages might entail differences in cognitive development; even though it is not the central issue in this research, this is connected to the kind of analysis that it pursues. Similarly, claiming that language has an influence on cognitive development means necessarily taking into consideration the fact that there is a substantial difference between claiming that a specific language shapes cognition and the idea that the development of language as a cognitive skill influences cognition. This also entails being careful when formulating claims about the cognitive mechanisms involved in language in universal form. For all of these reasons, attention will be given to cross-linguistic data, when available. Given that data in this sense is not abundant, the contribution of the cross-linguistic data to this book is still somewhat relative; however, I will report on what is available in order to allow the placement of English data within a larger context.

One might wonder why, in a philosophical inquiry, the preferred term is *mental state verbs* rather than *propositional attitude reports*, which is the proper philosophical term for sentences attributing/reporting a propositional attitude for a given subject. Many of the issues concerning propositional attitude reports are very relevant for the current work, since (1) they shed light on linguistic components that according to some theories are essential for conceptual development and (2) they give an idea of how complicated the cognitive operations underneath the acquisition of the syntactic structures involved are. The focus of this book, however, is on the interrelation between *natural language* and cognitive processing; this means that attention should be given to the linguistic data we are concerned with, to better understand its nature and its potential role in cognitive processing. Moreover, propositional attitude reports are only one of the many ways mental state verbs can be used.

Similarly, it is necessary to specify something about the following two sections. There are semantic properties of mental state verbs constructions that are especially complex from a formal semantics point of view. The next two paragraphs are by no means meant to be exhaustive of the complex logical problems arising from referential indeterminacy and opacity, but just a series of brief specifications about the issues related to mental state verbs in the literature.

3.1 Referential indeterminacy, opacity and other semantic issues

3.1.1 Referential indeterminacy

The problem of referential indeterminacy is not unique to mental state verbs and constitutes a puzzle for language acquisition in general. The idea underlining the puzzle arises when considering the ostensive learning paradigm, i.e. the idea that an effective way to teach a child a word is by pointing to a determined object and naming it. However, an important objection can arise, which is partially due to Quine's elaborations on referential indeterminacy (Quine, 1960). Let us imagine that a caregiver interacts with the child and, pointing to a box full of pencils, says "pencil". To the child's eye, the word might be referring to a series of different things: the pencils, the box, the box and the pencils together, one pencil and so on. How do children distinguish among the very different hypotheses they can formulate about what the words are actually referring to? This combines with another problem that lies at the core of research regarding language and concept acquisition alike, which is that of abstract terms. Referential indeterminacy is not the only issue that arises when considering word learning in general; another issue is that the correct referent of the word is not a concrete object, but something arguably more abstract – an action, for instance, or a non-perceivable entity. In the case of mental terms, this becomes even more marked since their referent is, at least intuitively, an "internal" state, or an attitude, that can even be attributed to another person and not necessarily to the subject; in this sense, the problem of a lack of reference becomes even more evident. Language acquisition theories like constraint theories (Markman, 1992) focus on the formulation of determined constraints that can help children to narrow down the possibilities, while the Social Pragmatic Approach (Tomasello, 2000) points in the direction of a resolution of the conflicts based on social-pragmatic cues and the functional distributional analysis of elements. Whichever theory of language acquisition is adopted, referential indeterminacy presents a problem, and how children acquire the meaning of mental state verbs, as will be seen, is far from a trivial question.

3.1.2 Referential opacity

Another characteristic of mental state verbs that is worth mentioning in terms of reference problems is so-called *referential opacity*, which can also be traced back to Quine (Quine, 1960), as well as to Frege (Frege, 1892), and which constitutes a

classic topic in formal language models. Using an often reported example (in this formulation in de Villiers (2005)) example:

(1) Oedipus married Jocasta.

(2) Oedipus married his mother.

(3) Oedipus thought he married Jocasta.

(4) Oedipus thought he married his mother.

In the four sentences, "Jocasta" and "his mother" all refer to the same individual, the character of the Greek tragedy. However, substituting the co-referential terms clearly does not retain the truth value. In other words, "Jocasta" is not substitutable *salva veritate* in the sentence (3). This characteristic is used in some theories of mental state term acquisition (de Villiers, 2005) to argue for a specific role for mental state term constructions in providing the child with a "Point of View", as will be spelled out in the following chapters. For now, it will be sufficient to consider the logical implications of these problems, sometimes referred to as Frege's puzzle.

This kind of problem is central to the work in Frege (1892), where the difference between *sense* and *denotation* is introduced. The debate couched in these terms regarding propositional ascriptions is long and complex, and entails discussion about the nature of propositions, the nature of the relation between propositions, citation theories, possible words semantics and of course theories of reference (Soames, 2002; Schiffer, 1992; Richard, 1989; Kripke, 1979; Perry, 2001; Crimmins and Perry, 1989; Davidson, 1968; Church, 1950; Bach, 2002; Stalnaker, 1984; Lewis, 1986). As such, it goes beyond the scope of this book. It will suffice here to delineate briefly the problem, and to point out that another way to think about it is in terms, not of reference, but "cognitive value"; in the sentences above, "Jocasta" and "Oedipus' mother" have the same reference, but not the same epistemic value for Oedipus, who would accept (5)

(5) I married Jocasta.

but not (6)

(6) I married my mother.

This kind of puzzle is telling in terms of how complex the acquisition of structures like the ones above can be. As it will be seen in later chapters, the kind of linguistic stimuli used in experimental tasks to test children's skills in attributing mental states has a relevant impact on the results.

3.1.3 Classes of mental state verbs

Any work concerning mental state verbs would be fundamentally incomplete without some specification of the classes which these verbs cover. Mental state verbs can, for a start, be divided into those which are factive and those which are non-factive (Kiparsky and Kiparsky, 1970). Good examples of the first class are verbs such as *forget* and *know*. Good examples of the latter are verbs such as *think* and *assume*. Factive verbs require their tensed complements (the complements they embed, see below) to be true, whereas the same does not hold for non-factive complements. For example:

(7) He forgot that Elisa was out with friends.

(8) He thought Elisa was at home waiting for him.

In the case of (7) the statement can only be true if Elisa is out with friends; in case of (8), the statement describes a belief, so the embedded clause "she was at home waiting for him" can be either true or false. This makes mental non-factive verbs naturally more interesting, as they can be the vehicle for false belief.[16]

Note that, obviously, mental state verbs are not the only class to include non-factive verbs, as should be evident when considering communication verbs. Observe (9):

(9) He said she was at home waiting for him.

The sentence has the same structure as (8). Some hypotheses have been advanced concerning the acquisition of mental state verbs that are based precisely on this similarity (de Villiers, 2005), but they have been further elaborated and modified and they are fundamentally bound to a specific theory of false belief reasoning, so they will be addressed in the chapter dedicated to the topic (5). It is also important to consider the data regarding languages in which the boundary between mental state verbs and communication verbs is not as strict and defined as in English: these cases will be treated in the section dedicated to cross-linguistic data on mental terms (3.4).

16 Note that it might even be argued that a verb like *think* is presuppositional, i.e. it might suggest an interpretation, in certain contexts, according to which the following sentence is more likely to be false. I will come back to this topic in chapter 4, when dealing specifically with some empirical studies like the one in Cohen et al. (2015).

3.2 Syntactic structures

In this section, I will briefly present some syntactic facts regarding mental state verbs. As will be explained and analyzed later on in this book, syntactic factors play a prominent role in many theories analyzing the contribution of language-learning to conceptual acquisition. Therefore, it is useful to keep syntactic factors in mind, along with considerations of a more philosophical and semantic nature.

3.2.1 Argument structure

Mental state verbs have a peculiar structure in terms of argument. Let us compare three different argument structures in English:

1. Camilla ran;
2. Andrea watered the plants;
3. Gabriele thought that Camilla was gone.

The intransitive verb in 1 has a single argument; the transitive verb in 2 has a two-argument structure; the mental verb in 3 has a complement structure, i. e. it takes a sentence as the argument. The possibility of a sentential complement structure is what makes it possible for the verb to embed another sentence and to allow for the subordination relation. In semantic terms, this is what makes it possible to express the propositional attitude of the subject towards another sentence. Once again, this applies to communication verbs as well. It is essential to notice that sentential complement structures with mental state verbs are learned in English *after* sentential complement structures with desire verbs (Custer, 1996; Perner et al., 2003). As will be seen in later chapters, this fact has some importance for various theoretical and empirical issues.

3.2.2 Syntax and semantic learning

Given some of the peculiarities explained above, it might not be surprising that mental state verbs are frequently labeled "hard words" (Gleitman et al., 2009), i. e. words which are particularly challenging in L1 acquisition and that are therefore learned later than others. While this group of words includes several kinds of verbs, mental state verbs are a particularly interesting and studied case (Papafragou/Cassidy/Gleitman, 2007; Shatz et al., 1983). According to a popular

view, the Conceptual Change proposal, hard words are acquired later because prior conceptual knowledge is required for their understanding (Gopnik and Meltzoff, 1997). This view is inherently connected to a specific position regarding Theory of Mind and mentalizing that will be explained in later chapters. Similarly, this view is connected to what has been treated in chapter 2 as the developmental model that sees conceptual acquisition as resulting from the refinement of theory-like knowledge, where the rules and representations used by children to learn about the world are similar to those used in scientific theory (Gopnik, 2001; Carey, 2009). As previously underlined, according to these theories, concepts are to be seen as abstract entities related to one another in complex and coherent ways (Gopnik, 2001). As has been pointed out in chapter 2, this theory does not straightforwardly argue that language development always comes as a result of conceptual development; on the contrary, it envisages, at the general level, co-development and mutual influence. However, in respect to mental verbs precisely, Gopnik has been supporting a different position, given her commitment to a certain conception of mentalizing (often characterized as "Theory-theory"):

> [...] changes in the childrens spontaneous extensions of these terms parallel changes in their predictions and explanations. The developing Theory of Mind is apparent both in semantic change and in conceptual change. (Gopnik and Meltzoff, 1997, 121)

The position is in line with an account championed by Smiley and Huttenlocher (1995), which predicts that conceptual development necessarily comes before language acquisition; a word-to-concept mapping is necessary, for the semantic content to be fixed.

Regardless of the idea of mentalizing that this view entails, the fundamental focus for now is on the fact that these theories see the acquisition of mental state language as, ultimately, a result of underlying conceptual abilities. The opposing account, named Informational Change Hypothesis, presented by Gleitman et al. (2009); Gleitman and Landau (1985), relies on an increase of informational resources instead, claiming that the acquisition of hard words is only possible when relevant syntactic structures are acquired. Learning to master specific grammatical structures allows for a *syntactic bootstrapping* mechanism that provides the structure-to-world mapping necessary for semantic acquisition. As will be further explained in later chapters, the *syntactic bootstrapping* account of hard words acquisition can be connected to, but fundamentally differs from, the syntactic bootstrapping hypothesis put forward in the context of Theory of Mind acquisition presented in chapter 5. Regarding the dynamic between conceptual change and informational change, empirical evidence seems to support

both accounts. For instance, while training in sentential complement constructions use promotes mental term acquisition (Papafragou/Cassidy/Gleitman, 2007), learning of both complement structures and mental state verbs, independently, correlates with cognitive tasks that imply mastery of mental-state-related concepts (Lohmann and Tomasello, 2003; de Villiers and Pyers, 2002). This suggests that, at the critical age of 4 years old, both important semantic and cognitive changes, which are reflected in their conceptual repertoire, underlie children's comprehension and use of mental state verbs.

Gleitman et al. (2009) present the following outline of the Informational Change argument:

1. Several sources of evidence contribute to solving the mapping problem for the lexicon;
2. These evidential sources vary in their informativeness over the lexicon as a whole;
2. Only one such evidential source is in place when word learning begins; namely, observation of the words situational contingencies;
3. Other systematic sources of evidence have to be built up by the learner through accumulating linguistic experience;
4. As the learner advances in knowledge of the language, these multiple sources of evidence converge on the meanings of new words. These procedures mitigate and sometimes reverse the distinction between "easy" and "hard" words. (Gleitman et al., 2009, p.28)

There is an obvious mismatch between the number of basic structure types and and the meaning of many other sorts of verbs. However, it is argued in Gleitman et al. (2009) that the syntactic structure can provide enough *focus* and establish saliency:

> When paired with a scene, the structural properties of an utterance focus the listener on only certain aspects of the many interpretations that are always available to describe a scene in view. (Gleitman et al., 2009, p.39)

The idea is that there is a "zooming" effect of the syntactic structure on relevant semantic features. In this sense, what syntactic structures do is mostly a matter of *narrowing down* the possible meanings that can be attributed to the verb. The account rests on the idea that there are some systematic correspondences between semantic and syntactic features that enable children to pick up on important cues and to disentangle the puzzles raised by the mapping problem. In particular, attention is given to:

– A tendency to align noun phrase and argument (this includes signing languages);
– A universal bias of agent and source semantic roles to be indicative of subject position.

The second is particularly interesting in terms of perspective taking; perspective verbs such as buy/sell or chase/flee are of particular interest because they highlight different perspectives on the same event. The perspective's choice, of course, is in some sense "the speaker's job"; the learner has to differentiate between different instances of the same event told by different perspectives. This seems to be facilitated by word order, or by general syntactic positioning, in a reliable way in English (Fisher et al., 1994). Note that this seems to imply that children are highly sensitive to perspective taking in language comprehension; this is used to argue in favor of the role of syntax as a "zoom lens"for picking up on certain semantic features.

A very important topic addressed in this context is that of attentional states, where the question is whether tracking of attentional state can be a relevant variable in the case of verb semantics, as it seems to be for perspective verbs like *chase*. In the study in Nappa et al. (2004), giving clues about the speaker's attention's direction influenced the verb chosen to describe a determined situation. For example, a scene was presented where a dog was chasing a human and the participants had to describe what they saw. In the condition with clues, a boy observing the scene was also represented, clearly looking at one of the actors (the dog, or the human), which proved to be influential with respect to the choice of the verb used to described the situation (for example, *chase* versus *flee*).

Two other principles are fundamental on this perspective (Gleitman et al., 2009, p.52):

1. The lexical and phrasal composition of arguments is related to the meanings of their predicates.
2. Sentence complementation implies a thematic relation between an animate entity and a proposition (semantically, an event or state of affairs).

The first principle is fundamental when one assumes that syntactic cues are essential in mapping new semantic information. Principle 2 is exactly the same principle that is at the center of theories that will be exposed in chapter 5. As support for this principle, a study (Papafragou/Cassidy/Gleitman, 2007) is often cited: cues related to the type of argument (sentential complement versus noun phrase) and the salience of a mental state (true belief versus false belief) had an impact on the frequency of mental state verbs (or credal verbs, as they are called in Gleitman et al. (2009)) produced in a verb-guessing paradigm.

Going back to the syntactic bootstrapping approach, the evidence for the argument presented by Gleitman comes from a series of studies performed using the Human Simulation Paradigm (HSP) (Gillette et al., 1999; Snedeker and Gleitman, 2004); this could be described as an attempt to make adults learn words in

a similar way to the way in which toddlers do, by presenting them with recorded dialogical situations between caretakers and infants, where they have to figure out the meaning of a word marked with a sound. As reported in Gleitman et al. (2009), it is much easier to guess the meaning of a noun compared to verbs (45% compared to 15% accuracy). In the mentioned studies, it was also found that adding other cues was essential for increasing the accuracy score for verbs missing. In particular, the informational sources used in Gillette et al. (1999) and reported by Gleitman et al. (2009) were the following:

- Scenes: video clips of mother-child interactions (no audio, beep played at time of unknown word).
- Nouns occurring in the six maternal utterances: (*gramma, you*); (*Daddy, Daddy*); (*Daddy, you*); (*I, Markie*); (*Markie, phone, you*); (*Mark*).
- Frames for the six maternal utterrances: *Why don't ver GORP telfa?*; *GORP wastorn, GORP wastorn*; *Ver gonna GORP wastorn?*; *Mek gonna GORP wastorn?*; *Mek gonna GORP litch*; *Can ver GORP litch on the fulgar?*; *GORP litch*.

The account proposed by Gleitman goes along these lines; initially, word learning occurs, for roughly the first hundred words, via ostension-learning for concrete words. This corresponds to the period in which children are mainly one-word-at-a-time speakers; subsequently, however, the vocabulary expands incredibly. This occurs when syntactic knowledge is also evident in speech, as children start uttering full sentences. Most importantly, in Snedeker and Gleitman (2004), concrete verbs and abstract verbs were compared, showing that the influence of syntactic clues was more informative, i.e. led to better learning results, for abstract verbs compared to concrete ones. All of this is interpreted in terms of structural clues providing a major contribution to semantic learning. This is possible, according to the syntactic bootstrapping hypothesis, because structural information relies on different cues. These are, firstly, distributional; processing of a word narrows down the possibilities for the words that follow it, based on frequency data; e.g, a verb like "drink" will likely be surrounded by a noun referring to animate subjects and direct subject nouns (Harris, 1957; Pinker, 1989). Moreover, there is a correlation between syntactic features like argument, type, positioning of verbs and verb meanings (see, among others Chomsky (1981); Fisher (1996); Jackendoff (1990)); a verb that describes an action will frequently have two noun phrases as arguments, a verb which denotes movement will have three places (source, object, and end point), and so on. The same happens, crucially, for the argument type: some verbs will often take nouns as complements, others will accept sentences, as is the case for mental verbs and communication verbs. It is essential to notice that the syntactic bootstrapping account does not predict that cues like argument type or correlation alone are sufficient for the

learning of the meaning of verbs; crucially, probabilistic cues are multiple and, if possibly jointly necessary, individually not sufficient.[17]

Let us return to the specific case of mental state verbs: according to the syntactic bootstrapping hypothesis, their meaning relies at least partially on the fact that they accept sentence complementation, which implies a thematic relation between an agent and an event, or state of affairs, according to Gleitman et al. (2009). Adults, as mentioned, are sensitive to these kind of regularities; moreover, sentential complements are more restrictive, it is argued, compared to noun phrases, as fewer verbs allow them. To this extent, then, they can be considered a stronger cue for the verb with which they appear. In a study (Papafragou/Cassidy/Gleitman 2007), the role of syntactic clues in suggesting the use of mental state verbs was investigated. Children (4 years old) and adults were presented with stories in two conditions: *false belief* and *true belief*. The descriptions of the stories were given verbally; however, pseudo-words were used as substitute for verbs in the stimuli, which consisted in either transitive or complement cause structures. Children had to guess the meaning of the "mystery word". Results showed that both false belief situations and sentential complement environments would more easily trigger responses in terms of mental verbs. The result was reinforced in an additional condition for adults, where every content word was substituted by pseudo-words. This allowed for some syntactic clues to be present (e.g. "that"): even in this case syntactic clues were more reliable than extra-linguistic context. These studies, according to the syntactic bootstrapping proponents, point to the fact that syntactic clues can be fundamental in learning the meaning of verbs, especially if they belong to the "hard words" group, both in adults and children.

Several things have to be noticed. While the experiments and arguments presented are convincing in showing that structural information has some role in aiding identification in the meaning of a given verb, there are some clear limitations. First of all, the assumption is that learning in adults and children is more or less comparable. While this is indeed plausible, it is not a neutral assumption. This is linked to the second point that should be taken into consideration, i.e. the fact that syntactic information plays a role does not completely exclude a conceptual change interpretation. 4-year-olds have already mastered mental state verbs, as will be repeated below; this means that, while it is clear that syn-

17 Note that even very recent work on the syntactic-semantics interface confirms the possibility of finding a very significant correlation between syntactic distribution of mental state verbs and semantic classes. A good example is White et al. (2018). While this might not be considered direct evidence for the syntactic bootstrapping hypothesis, it definitely is encouraging for its proponents.

tactic information helps to narrow down the possible space of interpretation for a given unknown verb, it is less clear that this kind of mechanism excludes the idea that some kind of understanding of what the mental state verbs denote is necessary in order to actually comprehend the verb and the stimuli in which it appears. Note that this is not in contrast with what is argued in Gleitman et al. (2009); Gleitman (2009), since they do recognize that syntactic bootstrapping might be aiding other learning mechanisms. Importantly, evidence that syntactic-based clues are used to identify the meaning of an unknown verb has also been found in other studies, like Harrigan et al. (2016), where the experimenters tested the impact of different syntactic constructions of children's interpretations of *hope*, a verb that is relatively unknown to children aged 4 and 5. Despite these limitations, the findings on which the Informational Change hypothesis relies seem to be convincing in pointing out that the acquisition of the meaning of verbs relies on a variety of clues, among which syntactic ones seem to be very relevant. As will be seen in chapter 6, how the syntactic abilities emerge can also be a matter of debate.

In chapter 6, I argue that, while syntax might be enabling semantic learning as proposed by Gleitman (2009), it also aids the development of cognitive resources, providing the learner with clues and patterns that make an abstract schema possible. Crucially, I will argue that other factors also have a relevant role in the formation of this schema, along with syntactic input. I will also present, in the following chapters, views and empirical evidence supporting the idea that it is not necessary to assume that a *concept* of mental states is available for young children. As will be argued, this model is compatible with several theories of the interaction between linguistic abilities and non-linguistic cognitive skills (for instance Tillas (2015a)).

3.2.3 Evidential constructions

A related piece of linguistic evidence that has to be taken into consideration when talking about the lexicon regarding the mental domain is evidentiality. Evidentiality is grammatical marking of information source (perception, inference, hearsay) and quality (strong or weak certainty): in other words, it is the linguistic tool used to mark where a piece of information comes from and how certain the information is considered to be. The connection between evidentials and lexicon regarding mental states is quite complex; firstly, assessing the quality of a source of information implies tracking the sources of one's beliefs and mental states, which usually implies an extensive use of both communication, perception and mental state verbs ("I know Alexandra took a day off" or "I was told Alex-

andra took a day off"; "I saw Kurt leaving the building" or "I think Kurt left the building"); moreover, this implies, as noticed in Papafragou and Li (2001), the complex cognitive abilities to recognize the sources of one's own beliefs and knowledge. In Papafragou and Li (2001), three different evidential relations conveyed in English are underlined:

- Mental states and perception verbs like know, think, see, guess, look, and hear;
- Epistemic modals, like must and may;
- Adverbials, like allegedly or reportedly.

Clearly, the first group is of most importance for the current analysis. However, as will be discussed later in the chapter, the expression of evidential relations varies across languages, which creates some interesting differences in terms of how the vocabulary associated with this kind of information is concerned. In the current section, I will highlight another interesting feature of the debate about evidential constructions.[18] Source monitoring is interesting developmentally speaking, as it seems to follow a rather incremental arch. While the connection between *seeing* and *knowing* seems to be understood around 3 years old (Pratt and Bryant, 1990; Pillow, 1989), and the fact that having directly witnessed an event is a good source of knowledge seems to be clear to 3-year-olds (Robinson and Whitcombe, 2003; Mitchell et al., 1996), they do not always manage to attribute knowledge gained through basis of verbal report (Wimmer et al., 1988), or to attribute greater knowledge to somebody that has been told something in contrast to somebody who has not (Mossler et al., 1976). Some relevant data on this will be presented in later chapters, as it is important for the development of social skills that might be connected with the acquisition of mental state verbs.

Note, furthermore, that interesting comparisons are possible when looking at the differences between acquisition of evidential use in English and acquisition of evidential morphology in other languages, like Korean and Turkish. In Turkish, there is obligatory marking in past tense to indicate the source of knowl-

18 Note that the debate around evidentiality acquisition, its connection to source tracking and its cognitive consequence extends well beyond the scope of this book. For further reading, the following works are to be recommended among others: Aikhenwald (2004); Fitneva and Matsui (2009); Chafe et al. (1986); Anderson (1986); Satoh et al. (2008); Willet (1988); for some overviews of the issues: Aksu-Ko and Slobin (1986); Aksu-Ko (1988, 2000); for the specific issue of evidentiality in Turkish: Brosseau-Liard and Birch (2011); Huffman et al. (1994); for English: Choi (1995); Lee (2010); for Korean: Lee and Law (2000); for Chinese: de Villiers et al. (2009); for Tibetan: Faller (2007); and for Cuzco-Quechua: McCready and Ogata (2007) for evidentiality and probability.

edge for a past event: *-dl* is used to indicate that the event was witnessed first hand, whereas *-mls* indicates that the speaker has indirect knowledge (Papafragou and Li, 2001). Studies like Aksu-Ko and Slobin (1986) show that, while a trend in differentiating the use of the two starts to show around 3.8 years old, the differentiation becomes more stable only when children are 6 years old. Choi (1995), on the other hand, shows that a good percentage of Turkish evidentials are acquired by children at the age of 2 and a half. In this sense, acquisition timeline varies a lot across languages.

Once again, the contrast is between the position that sees language learning as subsequent to conceptual development, as maintained in Smiley and Huttenlocher (1995) and Gopnik and Meltzoff (1997) and a more multi-faceted idea. According to the Conceptual Change hypothesis, the late acquisition of linguistic means to express evidentiality is due to the difficulty entailed in mastering the underlying concepts; according to the rival hypothesis, presented in Ozturk and Papafragou (2016), it is due to difficulties in mapping the concepts to the language stimuli. The experiments carried on in Ozturk and Papafragou (2016) are focused on Turkish evidentials, given the complex evidential marking of this language and, are aimed to verify whether the mentioned source monitoring skills develop after, before, or in parallel with the linguistic acquisition of evidential marking. The experiments test children (aged 5, 6, and 7) in tasks where they had to report how their own knowledge was acquired, how others acquired information, and how reliable other agents were on the basis of their information sources, while comparing these results with their linguistic mastery of evidential marking. Results show that the complex morphology is acquired well after basic abilities in source monitoring become part of the cognitive repertoire of the child, leaving however some room for conceptual development to play a role in the drive for linguistic acquisition as far as indirect evidence is concerned; the direct-indirect source marking in language seems to be acquired hand-in-hand, chronologically speaking, with the abilities in distinguishing between direct and indirect source of knowledge. Recall what has been argued in the previous chapter 2; while there is room to think that language acquisition plays a role in the development and change of cognitive skills, caution should be exercised when assuming that sophisticated cognitive mechanisms only arise as a consequence of language learning.

3.3 Desire verbs?

As will be clear, the book will focus mostly on verbs like *to think*, *to know*, *to believe* and so on. The verb *to want* might look like another potential candidate for

the current analysis. In what follows I will give some explanations of why I have decided to focus on mental state verbs strictly speaking, and on why the verb *to want* is nevertheless a valid example to keep in mind when analyzing their peculiarities.

At first glance, there are elements that make beliefs and desires fundamentally different in terms of the relation between the world and the mind of the child, or direction of fit (Searle and Vanderveken, 1985). Classically, beliefs are thought to have a world-to-mind direction, whereas desires have a mind-to-world direction; in other words, we are supposed to adjust our beliefs to the world, whereas to want something means wanting the world to "adjust" to our mind. Moreover, this is related to a difference in terms of *realis/irrealis* dimension. *Want* takes *irrealis* objects: the verb characterizes a wanted change in the state of the world (the possession of an object, or an event) as projections in the future and not current states of affairs (de Villiers, 2005). In this sense, the objects it takes are "irrealis": this is not necessarily true for "think" and "believe", whose object can be a current state of affairs. This has particular relevance for the debate regarding specific theories about mentalizing and language (de Villiers and de Villiers, 2009; de Villiers, 2001, 2005) and cross-linguistic data (Perner et al., 2005) so it will be addressed in more details in chapter 4 and chapter 5.

Secondly, there is a subtle yet important difference in the abstractness of the two mental state verbs: *to want* has a quite straightforward relation with behavioral evidence that verbs expressing beliefs lack. Expressing a desire or a need is easily enough one of the first activities children engage in, and is likely to be one of the first behaviors they learn to control. Evidence shows that children learn quite early to point at something they want in order to focus adults' attention on their behavior (Carpenter, 2009; Leung and Rheingold, 1981; Masataka, 2003; Morissette et al., 1995). Moreover, two things should be kept in mind: firstly, non-human animals might engage in communicative acts expressing desire too. More importantly, there is a sense in which *to want* is less prone to referential indeterminacy in comparison to *to think* : an act of wanting something can easily be exemplified by pointing, reaching, or struggling to get something similar, while this is more difficult in the case of a verb denoting the having of a thought. In this sense, "wanting" is associated more strongly with a definite set of behaviors and external expressions, which is not the case for believing. Consequently, one might see this kind of verb as occupying a middle ground between abstract verbs like *think* and more concrete action verbs like "run" for instance. In chapter 6, I will argue that a fundamental difference is also the pragmatic and communicative context of use for desire and belief verbs, which is a

hypothesis backed up by other theories and approaches (for example, Van Cleave and Gauker (2013)).

The two mentioned points, namely the difference in abstractness and the different relation that desires and beliefs exemplify in terms of agent-world interaction, make the two verbs in question fundamentally different in terms of what they denote, which is the first relevant reason to set them aside. However, other differences, more strictly related to their linguistic nature, can be noted. At least as far as English goes (more will be said below), the following is important:

1. *Want* is acquired before *think*, with want being mastered by children around the age of 3 years old and used around this age in all of its forms.
2. There are differences in how the complements behave in syntactic terms, at least in some languages.

As a matter of fact, the already mentioned feature of sentential complementation is relevant here. Let us observe the three examples:

1. He wants a car.
2. He wants to buy a car.
3. He wants Jane to buy a car.

As noted in de Villiers (2005) we have either a noun complement or an infinitive complement, and, while the first two forms are mastered by the third year of age, the third one becomes frequent a bit later. (De Villiers reports here evidence from Bloom et al. (1975); Bartsch and Wellman (1995).) The acquisition of the verb, either way, seems to be faster than in the case of *think*. The infinitive structure is also relevant for the second point; in English, there is a difference between the two structures:

1. He wants her to leave.
2. *He wants (that) she is leaving.
3. He thinks (that) she is leaving.
4. *He thinks her to leave.

This does not hold for every language, as noticed by Perner et al. (2005), and it is the subject of a heated debate explained in chapter 4.

This brings us directly to the another fundamental issue, which is the differences across languages in the way and order that mental terms are acquired; the above difference between syntactic structure is far from being a universal fact

about language. In the next section, this and other relevant variations across languages reported in the literature will be analyzed. Finally, a relevant piece of evidence should be mentioned: the CHILDES database (MacWhinney, 2000), a collection of child-directed speech, reports that *want* occurs 22.012 times per million utterances, whereas *think* occurs 10,187 times per million utterances. While the value of frequency should not be overestimated, it is to be kept in mind that children are significantly less exposed to sentences containing *think*, which might be highly influential in its late acquisition. Not surprisingly, Booth et al. (1997) found evidence that the use of the verb *know* in children is correlated with its parental use. Note, on the other hand, that lower-frequency action verbs are learned earlier than more frequent mental state verbs (Shatz et al., 1983), which is why frequency of stimuli seems not to be sufficient to explain order of acquisition.

3.4 Mental terms cross-linguistically

A fundamental concern with treating mental state verbs is that most of the examples and the theories focus on English. However, mental state verb polysemies and variation in syntactic structures across languages have to be considered. This is a particularly urgent concern, as will be clear in the following chapters, for the theories that attribute to the peculiar syntax of mental terms and expressions a fundamental role in conceptual development and acquisition of mentalizing skills. As a consequence, I will review the available data regarding mental vocabulary and assess the impact on the claims that were previously made.

3.4.1 Syntax

Let us go back to the example given above, of differences between the structure for *want* and *think* in English:

(10) Andrea wants her to leave.

(11) Andrea thinks (that) she is leaving.

Let us compare these sentences to German (examples from de Villiers (2005), slightly modified):

(12) Die Mutter will dass Gabri ins Bett geht. (*The Mother wants that Gabri goes to bed.)

(13) Die Mutter glaubt dass Gabri ins Bett geht. (The Mother thinks that Gabri goes to bed.)

At least on the surface, there is no substantial difference between the two syntactic structures and, significantly, a similar phenomenon can be observed in Cantonese and Mandarin (de Villiers, 2005). However, this seems to have little influence on the acquisition timeline of the verbs; children seem to acquire the want-structures before the know-ones (Tardiff et al., 1997). As explained in section 3.3, moreover, fundamental differences in meaning and in the possible connection with external stimuli have to be taken into consideration: in the following chapters, the relevance of this argument for theories of mentalizing will be further discussed.

3.4.2 Semantics

Not every language has mental state terms that are as easily identifiable as English ones. For example, Vinden (1996) presents interesting data about Junín Quechua, a dialect of Quechua spoken in the Andean highlands, in Perù. In this language, the verb *say* is used often in combination with *yes* to express the concept of belief, which is otherwise not explicitly linguistically encoded. The same can be said for *thought* and *denial*, which are also related to the use of *say*. Interestingly, Vinden reports that these verbs and combinations are remarkably absent in story-telling as well. As will be shown in chapter 4, this seems to be connected to the performance of children in this community in some social cognition tasks. In another relevant paper, Vinden (1999) presents evidence regarding the social cognitive performance of children whose primary language is Tainae, which does not lack mental state language, like Junín Quechua, but where it is out of the ordinary to ask about somebody's mental states without specifying the content. In other words, it is possible to say "The person thinks that he will clear the garden", whereas it is odd to ask "What does the person think?"

Remarkably, some cultures have been documented as talking about mental states and folk psychology considerably less than Westerners are used to. An example is given in Lutz (1982); the Ifaluk people, living in Micronesia, use a generic term roughly equivalent to *insides* to refer to thoughts, emotions, desires, illness, and physical sensations; something, in other words, denoting internal states of some kind. Along with this term, words like *nunuwan* denote both thoughts or emotions indiscriminately. Similarly, Gerber (1985) reports that Samoan culture is sometimes reluctant to give folk-psychological explanations of behavior; more likely, actions and intentions of agents will be talked about in

terms of external driving factors. Something similar can be said for Baining in Papa New Guinea (Fajans, 1997).

The number of studies exploring this kind of cross-linguistic differences is far from abundant, but it helps to point out a relevant factor, in that there is more than language at stake: a lot of anthropological work has to be done to analyze the cultural and social factors that coat both linguistic and social development. However, for the reasons listed at the beginning of the chapter, it is useful to keep in mind how data regarding languages and cultures do not always align with the most studied Western ones. In chapter 4, I will discuss some data on social cognition that will cast some light on theories that are extremely relevant for the present book, and in the rest of this work I will underline how cross-cultural data provides valuable support for theories that consider cultural factors as central.

3.4.3 Syntax and semantics: Mandarin and Cantonese

Mandarin- and Cantonese-speaking children produce more verbs than English-speaking children at the same age (Tardif, 1996; Tardif et al., 1999; Tardif and Wellman, 2000). Moreover, Mandarin employs less complex structures for complementation, where "that" and finite tense markings are both absent in Mandarin constructions. If in English there is a syntactic difference between (15) and (14):

(14) Who did Big Bird forget to invite?

(15) Who did Big Bird forget that he invited?

in Mandarin it is possible to have:

(16) Big Bird m4-gei3-dak1 zo2 ceng2 bin1go3 heoi3 party?

The meaning in (16) could be equivalent to (14) or (15) (Tardif et al., 2007). Moreover, Mandarin has mental state verbs whose meaning is not clear-cut as they are in English, with the same lexical items that can express *to think* as well as *to want to do something*, *to believe*, and others (Tardif and Wellman, 2000). As a consequence, Mandarin is an interesting testing ground to investigate differences with English, where verbs expressing desires are acquired earlier than those expressing belief. The study in Tardif and Wellman (2000) addressed detailed use and reference of the verbs in question in order to disambiguate their contextual meaning in child production in both Mandarin and Cantonese. The results revealed that, despite the different semantic features of the languages involved,

Mandarin- and Cantonese-speaking children still linguistically code desires earlier than cognitive states like thinking and knowing (Tardif and Wellman, 2000).

3.5 Acquisition timeline

3.5.1 First uses

Children start acquiring mental state verbs relatively early, with the first uses emerging around 2 and a half years (Limber, 1973), and by the end of the third year of life they seem to be successfully using them in some structures. However, several specifications are in order.

First of all, some uses of mental words by children can be described, as Diesel and Tomasello (Diessel and Tomasello, 2001) phrase it, as qualifying "the degree of certainty-uncertainty of the complement proposition"; hence, they do not use them assertively. This only happens at the age of four. Let us look at the examples:

1. "Where is my unicorn?" "I think it is in the garden."
2. "Where is my unicorn?" "Maybe in the garden."

where the utterance 1 can actually be read as 2.

Another relevant fact seems to be that these verbs are used initially as conversational hedges (Shatz et al., 1983): this is the case with forms like "you know" and "I think I want an ice-cream". These cases, while related to the function of expressing mental states, have a fixed conversational function that does not necessarily reflect an understanding of the semantics of the verb itself. In other words, they are also classifiable as non-assertive, and their main functions seem to be to direct the interaction in terms of, for instance, getting attention ("You know, this is my hat!"), to introduce or get information ("Know what?"), and to introduce an activity ("I thought we could go to the bakery"). In other words, although these expressions seem to have a variety of conversational and communicative functions, they do not seem to have the same functional role that structures with sentential complements have. These uses have been called by (Diessel and Tomasello, 2001, p.98) epistemic markers, *attention getters*, or *markers of illocutionary force*.

These data are in line with a finding reported in several studies regarding sentential complements like the ones described above; despite the fact that children seem to be mirroring these structures frequently, they do not seem to master sentential complementation in a more-than-formulaic way for the longest time

(Diessel and Tomasello, 2001; Kidd, 2003; Bloom, 1993; Limber, 1973). According to the analysis run by Diessel and Tomasello (2001), as a matter of fact, they constitute merely a "constructional island" around specific verbs; in other words, while children are able to pick up on the frequent use of these constructions and to replicate them, they are not able to apply the structures productively until after their fifth year. Evidence that children still have difficulties with complement constructions at age 4 is presented in de Villiers (2005, 2007) as well. Evidence about referential substitution in opaque constructions, on the other hand, like the ones in the Oedipus-Jocasta example above, pushes the full mastery of sentential skills further, with children still not handling the distinction when 6 years old (Russel, 1987; Kamawar and Olson, 1999; de Villiers, 2001).

In de Villiers (2005), children had to reply to a question of the following kind:

(17) Mom said to Dad, "I'm so happy that Bella is washing the dog."
(a picture of Bella mopping the floor)
Did Mom think Bella was washing the dog?

3-year-olds consistently replied incorrectly to question (17). Also, three years old fail the "memory for complements" task (de Villiers and Pyers, 2002) as in (18):

(18) The boy said he found a ring. But look, it was a bottle top. What did the boy say he found?

As will be discussed in chapter 4 and chapter 5, this kind of evidence has major roles in the discussion regarding the role of mental state verb acquisition in social cognition.

Regarding the different uses of mental state verbs and their appearance in development, it is perhaps beneficial to adopt the suggestion made by Shatz et al. (1983) on the matter, and characterize "pure" use of mental state vocabulary as those sentences in production where the mental state verb is used to talk about other people's thoughts, memories, or knowledge (Shatz et al., 1983, p.307). This seems to capture at least some of the semantic characteristics of the domain, since it implies considering those expressions that refer to internal (thus not directly observable) states, and definitely goes hand-in-hand with some of the most important syntactic features we discussed above (most importantly, embedded clauses). Another relevant note in this respect is made by the same authors in terms of contrastives: in their study, Shatz and colleagues pay special attention to contrastive sentences, which are sentences that highlight the difference or discrepancy between two different states, such as:

(19) "Before I thought this was a crocodile, now I know this is an alligator."

(20) "I am just pretending" in response to the question "Are they really dead?"

(21) "I did not know you went to the store." (Shatz et al., 1983, p. 309)

In all these cases, what emerges is a contrast between two different mental states. The ability to tell apart two different perspectives on a matter, and to consider one of the perspectives as fundamentally false is obviously essential in social cognition, as will be especially clear in the next chapters. As a result, checking whether or not children are able to encode this linguistically is of central importance.

3.5.2 Distinguishing between know and think

The issue of when different mental state verbs are distinguished by children is more complex than initial appearances suggest. While production data is easier to investigate, being mostly based on observation, it is trickier to find measures of comprehension that delineate when and how children understand adults' utterances.

Classical studies cited in the literature as evidence for the comprehension of verbs like *think* and *know* are Moore and Davidge (1989) and Johnson and Maratsos (1977), but previous experiments have used different measures. In Macnamara et al. (1976), for example, children were asked rather complicated questions regarding whether or not speakers using mental state verbs in a given story were lying. The results confirmed the fact that, despite the rather complex meta-cognition language requirements, 4-year-olds were able to pass the task, giving good evidence of their being able to master at least to some extent the meaning of *know* in terms of being a factive verb, i.e. a verb suggesting the truth of the following statement. Notice that this kind of evidence might suggest that a story-telling context aids the comprehension of the presuppositions and implications of the use of mental state verbs, an observation made by Abbeduto and Rosenberg (1985) as well.[19] Evidence along the same lines was found in Abbeduto and Rosenberg (1985), where three different taks were submitted to 4-

19 As will be seen in the following chapters, there are reasons to think that a connection between story-telling and cognitive skills plays a role in mental state understanding and mentalizing.

year-olds. In one of the tasks, the stimuli was composed of a given context, a test sentence containing a complement structure, and a question to be answered:

(22) context: I have a friend named Mary. Mary has a cat. test sentence: Mary forgets that the cat is slow. presupposition question: Is the cat slow?

In the second task, children were given a story, in which it was either underlined how the main character had good access to the information that was relevant (e. g. who was riding a bike) or how the access to the information was, on the contrary, limited (e. g. there was not enough light to recognize the person properly). Afterwards, children had to answer questions about the speaker's certainty ("Does the speaker know that it was Tommy riding the bike or does the speaker think it was Tommy riding the bike?"). In the last task, children had to give definitions for mental verbs. It emerged that solid performance could be achieved by children in these tasks by the age of 4, especially for *know*, *forget*, *remember* and *think*, with some difficulties with the verb *believe*, specifically, which caused problems up to the age of 7.

The very cited study in Moore and Davidge (1989) tests children's ability to deal with the implications and presuppositions of mental state verbs in a practical task, where children listened to puppets giving them clues about where to look for a toy, using either "know", "I am sure" or *think*. Children at the age of 4 but not those at age 3 were able to successfully rely more on *know* and *be sure*, which was interpreted by the experimenters as showing a better comprehension of, at least, the pragmatic roles played by the verbs. Notice that this also seems to indicate a mastery of sentential complement structures, at least at some level.

It should also be noticed that children show a late understanding (5 years old) of the differences in the meaning of different mental states like *know*, *guess*, and *remember*, a result reported in a series of studies by Johnson and Wellman (1980), in which the children had to express how they got to know about the location of an object in different conditions. In this case, children's performance improved in the 4-year-old age group, and kept improving with increasing age. (The oldest children in the study were 8 years old.)

The study conducted by Dudley et al. (2015) explicitly criticizes the classic paradigms for testing mental state verb comprehension used in the tasks described for being extremely hard and presupposing significant mastery of verbal and non-verbal information. In their task, children (aged 3) were told that the experimenter would hide a toy in one of the two boxes in front of them, and that they were going to be given a clue. The clues were divided according to verb type and negation type. So for both *think* and *know*, test sentences were given in neutral format (e. g. "The puppet thinks/knows that's in the red

box"), in matrix format (e.g. "The puppet does not think/know it's in the red box") and embedded (e.g. "The puppet thinks/knows that it's not in the red box").

The results were multi-faceted: children's performance suggested that some difference between *think* and *know* was known to the children, as they reliably chose the box mentioned by the suggester when the sentence was "LC thinks that it is in the red box" but reliably chose the alternative box when the sentence was "LC doesn't think the toy is in the red box" or "LC thinks that it's not in the box". However, some difficulties are met with *know* sentences, where children failed to perform as adults, and did not reliably chose the mentioned box when the sentences where of the kind "LC doesn't know that the toy is in the red box". The results are discussed in Dudley et al. (2015) as evidence that children have an "adult-like" understanding of *think*, while still failing to fully understand *know*. While it seems reasonable to conclude (with caution, since the study has not yet been replicated) that there is a different treatment of the two verbs, it is a stretch to assume that the results imply a full, adult-like comprehension of *think*, since children might be applying a rule based on negation more than any semantic comprehension of the verb. When a negation is present, they choose the non-cited box, since they attribute some knowledge to the suggesting puppet. While this is not grounds for definitely dismissing the results achieved in Dudley et al. (2015), it is worth underlying that it is a possibility worth exploring. On the other hand, their conclusions about 3-years-old having some difficulties with the use and comprehension of *know* and *think* is in line with the other cited results.

A final remark has to be made. It is a rather uncontroversial result that mastery of action verbs comes quite early; in particular, children seem to master motion verbs around their second birthday (Huttenlocher et al., 1983; Gentner, 1978; Bloom et al., 1975). This means that, while there is a significant sense in which most words acquired at the beginning of language development are nouns and not verbs (Gleitman et al., 2009), there is reason to maintain that acquisition of specific categories of verbs does indeed occur later than others, thus further supporting the idea that mental state verbs might be, at least to some extent, "hard words".

3.5.3 A timeline

The following table summarizes the data concerning the acquisition of mental state verbs given so far. This table will be referred to later on, to compare it with the performance in other important cognitive tasks:

Table 3.1: Summary of the acquisition timeline.

Age	Production	Comprehension
2.6	*know*, *say*, *think* (directing conversation function)	*know*, *say*, *think*
3	Higher frequency of use of mental state verbs	
3.5		First signs of distinguishing between different mental verb meanings
4	First references to own false belief; first assertive use of mental verbs	Comprehension of false belief sentence; distinction between *believe* and *know*; improvement in sentential complement construction
6	Mastery of complement structures	Better distinction between *know* and *guess*

3.6 Summary

This chapter has presented data and issues related to the acquisition of mental state verbs, underlying their peculiarities from a semantic point of view and their characteristic features with respect to language acquisition. The problems of referential indeterminacy and referential opacity in relation to this class of verbs have been highlighted, and their classification in terms of factivity has been underlined. The syntactic features of these verbs have been presented in relation to argument structure, and to the specific issue discussed in the psychological literature regarding their acquisition. It has been discussed how the acquisition of this class of verbs, which comes particularly late in development, is the subject of a debate regarding conceptual development and the role of syntax in promoting semantic learning. The debate on conceptual change and informational change will be especially relevant in the rest of the book, and is especially related to chapter 2, but it will also be referred to in chapter 5 and chapter 6. The debate about evidential constructions has been cited as a related field of study: an ongoing field of research explores the connection between source monitoring abilities and the acquisition of evidential morphology. A small space was dedicated to the relevant differences between expression of desire and belief, and cross-linguistic data was presented. Finally, I have given an overview of what the relevant acquisition timeline is, highlighted the differences between different

uses of mental state verbs, and underlined some relevant differences between verbs and when they are actively learned.

This chapter had two goals: to present an overview of the problems and issues related to mental state verb acquisition in general, and to make clear how their peculiarities are of interest not only to linguistics, but also to philosophical and psychological research. In chapter 4, a rather different literature review will be given, where I will explore the data that more systematically investigates the relation between mental state verb acquisition and social cognition.

Chapter 4:
Language and social cognition: data about the role of language in mentalizing

4.1 Introduction

In the previous chapter, I presented the main characteristics of the semantic domain I am investigating, i. e. that of mental state verbs. The more general purpose of this book is to investigate the role that acquisition of the described structures has in developing non-linguistic skills that might be related to them. As argued in previous chapters, it is reasonable to assume that language interacts with non-linguistic skills in a variety of ways, and that how these effects develop, and the impact they have, will depend on a variety of factors. It has been showed in chapter 2 that there is reason to think that labels play a role in categorization and perception mechanisms. The challenge is to investigate which, if any, are the effects of learning the language related to the mental domain, i. e. mental state verbs and nouns, and the constructions they appear in, when it comes to processes that are related to *understanding* the domain of the mental, i. e. the domain of social cognition.

The chapter is divided as follows: firstly, I will delineate some main issues related to the specific topic that is mentalizing. There are two reasons for doing this, one practical and one theoretical. Practically, the debate about the role of language in mentalizing skills has its own specificity and scope, and has been evolving for a long time, producing a large body of literature that needs to be considered on its own. Understanding this literature and assessing its contribution means situating it in the context of mentalizing skills in general. The second motivation is related to the first, but has a theoretical nature: as I argued above, taking seriously the role of language in cognition entails investigating the effect that it has on *specific* mechanisms, and in the case of mentalizing it appears to be the domain of social cognition that constitutes a good starting point for investigating the role of linguistic information. I will start by defining what is mindreading, why it is at the center of the debate on social cognition, and what are the cognitive skills related to understanding others that are studied in the literature. Defining mentalizing, I will make an effort to specify how different traditions categorize some features in different ways. I will give special relevance to the false belief task, as it is considered an essential issue in the literature regarding mindreading, and because it is related to different important philosophical questions: the issue of mental state attributions, on the

https://doi.org/10.1515/9783110748475-005

one hand, and the nature and specificity of the representations involved, on the other.

Subsequently, I will present the findings concerning language acquisition and the development of mindreading. I have reviewed some of this data in Berio (2020a) as well, but I here present a much more detailed analysis. The empirical review will consider first data in typically developing children. I will disentangle the different components (pragmatics, semantics, syntax, etc.) that have been studied. I will also briefly report on the literature about bilingualism and cross-linguistic differences in mentalizing, and give an overview of the data about the relation between a linguistic practice like storytelling and mentalizing skills. In the last part of the chapter, I will dedicate space to clinical data. As it will be seen, specific conditions like ASD, SLI and Schizophrenia offer informative windows on the relation between language and social cognition skills.

The purpose of this chapter is twofold; on the one hand, it is meant to provide the background against which theories in the following chapters will be analysed. On the other hand, I will argue through the chapter that a role of language acquisition in the development of mentalizing skills is demonstrated by the data in more than one important way. This will prepare the ground for the coming chapters, where current theories which interrelate the two will be analyzed.

4.2 A picture of the debate

4.2.1 What is mentalizing?

The terms *mentalizing*, *mindreading* and *Theory of Mind* are often used in the literature in relation to each other and, as will be highlighted later, are not theoretically neutral. For now, it will be useful to understand what these words refer to, in order to be able to consequently disentangle their uses and implications. The description given below of the relevant theoretical stances regarding mentalizing simplifies, in many ways, a very rich and long debate, but it serves (for the moment) the purpose of defining the object of inquiry in order to understand the empirical literature in social cognition.

Mentalizing refers to the ability to think about beliefs, desires, knowledge, and intentions (Apperly, 2011). This ability is thought to be at the basis of what is frequently called *folk psychology*, i. e. the practice of explaining people's behavior in terms of their mental states. Whether or not folk-psychology explanations have causal value and/or explanatory power is, obviously, a matter of debate (Goldman, 2006), and a positive answer to this question has been chal-

lenged from many directions[20] While this remains an issue that has to be solved, it is not the subject of the current chapter: I will assume here that humans engage in predictions and explanations of other agents' behavior, and that they do that by doing *something* like attributing them mental states. What this "something" is, however, is up for debate. The received view, which is, however, being put increasingly under stress by empirical findings and theoretical challenges, is that this *something* is literally attributing mental states to other people, and doing so while applying a so-called *Theory of Mind*, or ToM. This is supposed to be a theory-like process: the way adult and children understand other agents is by having (1) a set of concepts, like *belief*, *thought*, *desire* and so on, and (2) a theory about how these concepts can be used to attribute mental states and how they relate to each other. This view has been championed by many, among the most prominent being Gopnik, Meltzoff, Perner, Wellman (Gopnik and Goldman, 1993; Gopnik et al., 1999; Wellman, 1990; Perner, 1991). The view assumes that discrepancies in the performance in social cognition during development, like those that will be explained below, are due to the fact the children have not yet developed a Theory of Mind and/or they lack the conceptual tools that are useful to understand other people's behavior. When understood in this light, mentalizing is having a *Theory of Mind* and being able to master *concepts* related to mental states.

Goldman describes Theory-Theory this way (Goldman, 2006, 70):

1. Young children's performance in mentalising tasks changes over time as a function of changes in their grasp, or understanding, of mental concepts;
2. These changes in concepts, or conceptual understanding, reflect successive stages in childrens theories of the mental;
2. Therefore, mental concepts must be theoretical concepts;
3. Hence, all determinations of the instantiation of mental concepts, in both self and others, must be inferential in character.

This approach is sometimes coupled with assumptions that come from the modularist tradition, and was mostly pushed by Leslie (1994a; Leslie and Thaiss 1992; 1987) and by empirical work like that of Baron-Cohen (1991). In this case, the idea is that there is a *module* dedicated to mindreading. This implies, normally, that it is an innate module (however, as will be shown in the next chapter, this is not the only option, as in Garfield et al. (2001)), and that pre-programmed stages of this module are predicted in development. There is a sense in which a traditional modularity approach to mindreading, interpreted strongly, is

20 See, among others, Churchland (1981, 1988).

straightforwardly incompatible with even the purpose of this book: if ToM is an encapsulated and innate module, and the information needed to carry on the tasks in social cognition as they are explained below is a *domain-specific* matter, then, unless one is prepared to make an uneasy assumption that the language module and the ToM module are the same, language acquisition should have hardly any influence on mindreading.

ToM is not the only possible approach to mentalizing, which is one of the reasons why the choice of words has to be more neutral. Simulation Theory (ST) has been proposed, in more than one variant, by many researchers from different fields, and it is tied to discoveries regarding mirror neurons. The fundamental upshot of ST is that mental state attribution occurs through a mechanism that entails using our own internal states in a simulation process to attribute mental states to others. While this was conceived initially in phenomenological terms in Gordon (1986), Collingwood (1946), Heal (1986), and partially Goldman (1986), the approach took a different turn when combined with the discovery of mirror neurons, due to the research carried on by Rizzolatti and Gentilucci (1988); Rizzolatti et al. (1987, 2001). Mirror neurons are neurons with visual-motor properties, i.e. they are active both during perception and action, and they are especially relevant in their visual-motor properties in that they are active both when executing a determined action and when perceiving the same action, in a very specific and selective way (i.e. they are sensitive to the goal of the action, they are sensitive to specific features like the kind of grip used to perform the hand-action, and so on (Rizzolatti and Sinigaglia, 2006)). They are thought to be sensitive to *actions* and not movements (Rizzolatti and Sinigaglia, 2006), which is, among other features, what makes them particularly interesting for theories of affordances (Gibson, 1986) and action-perception mechanisms (Zipoli Caiani, 2010). A fundamental tenet of ST, then, is that mentalizing occurs as an upshot of these mirroring mechanisms (Goldman, 2006); in other words, attributing mental states to others involves using our own mental states to simulate their behavior, in continuity with our ability to understand people's actions by recruiting our own motor and perceptual resources to do it. In this sense, *mentalizing*, while it may involve the use of conceptual knowledge, does not involve *theory-like* knowledge.[21]

21 Note that one strong point of departure between ST and TT is on the issue of whether *self-attribution* is to be understood as different from attributing mental states to somebody else. While the version of ST in Gordon (1992, 2005) does not include introspection, Goldman (2006) famously attributes to introspection a central role in mindreading, arguing that one has privileged access to one's own mental states. In the case of theory-theory, the assumption is often that theoretical explanations about oneself are made in a way which is similar to the

This book is not meant to solve the long-standing dispute between ST and TT, especially because, as will be pointed out in the following chapters, the accounts that seem more likely to be correct in assessing mentalizing actually combine features of both. However, the aim of this book is to assess what role, if any, language acquisition has in mentalizing. As will be seen in the next chapter, theories from both sides of the field have tried to incorporate the following data. Before analyzing the data about the interaction between language and mentalizing that is central to this book, however, some specifications have to be made about other aspects of the classical literature on mentalizing, like the most common tasks and tests used in the literature.

4.2.2 The false belief task: the standard mentalizing test

A fundamental measure of mentalizing skills comes from the so-called *false belief task* (henceforth FBT). The extremely influential paradigm was developed in Wimmer and Perner (1983). It should be noticed that the use of the false belief task as a standard mentalizing test is not uncontroversial: Bloom and German (2000), for example, argues that the false belief task should be abandoned in this sense, because of the many complex abilities that it seems to entail, on the one hand, and because of how partial it is as a representation of social cognition skills, on the other. While both reasons are certainly valid, I do believe that abandoning the false belief task as a paradigm in the present analysis would be impossible without compromising the debate. Moreover, as will be seen, the variety of skills that are involved in FBTs are actually very relevant for the debate regarding the role of language in mindreading, and hence constitute an appealing aspect more than a discouraging one.

In the original version of the task in Wimmer and Perner (1983) children are familiarized with the puppet Maxi, who has a piece of chocolate and hides it in a cupboard, then leaves the scene. Maxi's mother enters the room when he is not present and changes its location, moving the piece of chocolate to another box. When she leaves the scene, Maxi returns, and declares that he will look for his piece of chocolate. Children are then asked where he will look for his chocolate, whether in the first box (the cupboard) or the second box (the refrigerator). Famously, children fail this task until the age of 3 and a half years old, and pass it

way in which attributions to others are made, i.e. by applying ToM. However, the issue of first-person attribution is to be explored separately, being connected as it is with introspection and consciousness. Like many other theoretical contributions, this book focuses on the attribution of mental states to others.

reliably when they are 4. A very famous variation of this task is Anne and Sally's (Baron-Cohen et al., 1985), where Sally places a marble in a box and leaves the room. In her absence, Anne moves the marble to a second box. Children are then asked 'When she comes back, where will Sally look for her marble?"[22].

The task has been largely replicated and modified in the literature, with the purpose of assessing which features make it so hard for children, on the one hand, and which features make it interesting from a mentalizing point of view, on the other. The Maxi and Anne and Sally variations are verbally elicited, explicit false belief tasks that entail an *unexpected location*. Variations on verbally elicited, explicit false belief tasks that do not include location variations have been used as well. In the *unexpected content task* (Perner et al., 1987; Gopnik and Astington, 1988; Flavell, 1993) children were shown a box of "smarties", candies that actually contained pencils, and were then asked to report back on their previously held belief about the contents of the box ("What did you think the box contained before we opened it?"). The control task consisted in showing the child a doll house that contained an apple at first; later, a doll was put in place of the apple, and the child was asked to recall what he saw inside the doll house before the change. Note that, while in this task mental state verbs like *think* were used, the same cannot be said of the Anne and Sally task above. In the *unexpected identity* task (Gopnik and Astington, 1988), children were also asked a specific question containing a mental state verb. In a similar set of tasks, participants were shown, for example, a picture of a cat that appeared to be black until a plastic cover was lifted, revealing that the cat was in fact green. Children were then asked what they thought the color was (in the change of representation task) and what did they think a person that saw the cat as covered up would think about the cat's color. In all of these cases, results showed a sharp difference between 3-year-old children and 4-year-old children, with the latter group performing successfully when the former group could not. The location task results, in the explicit verbal version task above, are considered very reliable and solid, and they have long been considered the most telling piece of evidence when it comes to revealing a fundamental change in the development of mentalizing skills in children. Fundamentally, results from the verbal FBT have been used to argued that, at age 4, something changes in children's cognition: they become able to decouple their own beliefs and expectations about reality from those of others, and consequently learn to attribute false beliefs to somebody else. Crucially, other forms of the explicit verbal false belief task were also developed, along with the change of location and unexpected

22 Among the many examples, see Byom and Mutlu (2013).

identity tasks described above. Among these, the deception task and the emotion attribution tasks are less used but very interesting. In the deception task, children often have to predict what a character would say if they wanted to deceive another character, or they have to hide something from other agents (Guajardo and Watson, 2017; Hughes and Dunn, 1998). In emotion attribution tasks, on the other hand, children have to predict the emotional state of a character (Harris et al., 1989) in particular situations, often when they encounter something not expected.

Crucially, not all the variations of the FBT delivered the same results. In what have been named *implicit* FBTs, children seem to perform better. In Clements and Perner (1994), for example, children look at Sam, a little mouse, who stores an object in front of one of two mouse holes and then falls asleep. Similarly to the Anne and Sally task, the location of the object is changed when Sam cannot see the change. When Sam comes back, children were not always asked to predict his behavior: instead, their gaze was monitored after the experimenter said the prompting sentence "I wonder where Sam is going to look". Children aged 3 (but not 2) looked, first, at the old location of the object: this was interpreted as a sign that the children, albeit non-consciously, knew at some level that Sam was likely to look for his object were he left it. On the contrary, children were not able to verbally formulate a prediction about Sam's behavior, failing the part of the task that was similar to the Maxi one. In a similar study, Southgate et al. (2007) performed the same experiment, but without the prompting sentence: in this case, 2-year-olds also seemed to look, first, at the previous location of the ball.

Onishi and Baillargeon (2005) famously discovered that the age of a "minimal" FBT can be lowered even more, tracking the gaze of 15-month-olds. In their paradigm, a longer looking time is taken to mean that the child is surprised by the outcome. In familiarization trials, a protagonist hid a toy in location A to allow the child to develop an association between the location and the agent. Then the usual change of location was made in absence of the protagonist. Looking times were measured for children who were subsequently observing the protagonist looking for the object, and results showed that longer looking times, reflecting surprise and hence expectation's failure, were observed when the protagonist was looking in the new location of the object.

Notice that, both in Southgate et al. (2007) and Onishi and Baillargeon (2005), an association was established between the agent and the location in the familiarization task; this, as Low and Perner (2012) notices, might be what guides their behavior. It is possible, in other words, to interpret the behavior of the child, not as an expectation about the protagonist's belief about the location of the object, but rather as the reflection of an association between the pro-

tagonist and the box, or location, where the protagonist has been looking. Perner puts it in terms of behavioral rules: the child might be operating according to an association rule "actor-yellow box" that has nothing to do with understanding the subject's mental state.

A similar objection could be made regarding the Song and Baillargeon (2008) study, where 14-and-a-half-month-olds were familiarized with an agent constantly choosing a doll with a blue hair over a stuffed skunk. When the agent was absent, the blue-haired doll was placed in a plain box, and the stuffed skunk was placed in a box that had a blue tuft of hair above it. Children were reliably looking longer when the agent was reaching for the plain box, suggesting that children were expecting the agent to look for the toy in the location that suggested the presence of the doll, i.e. the blue-tuft box. In the study, the location (right-left) of the boxes was changed in a control trial to make sure that children were not associating the agent with one of the two locations instead. This, according to the authors, along with the fact that children were not familiarized with the boxes before the task, suggested that no association rule was formed. However, this remains to be established: since the box had a perceptual characteristic in common with the doll, and a prominent one like blue hair, I do not think it is completely clear how they could exclude the fact that an association was indeed formed, between the agent and a preference for blue-haired objects.

In a study by Surian et al. (2007), results in the same direction were obtained with a clever variation of the paradigm. In the first experiment, children were familiarized with a caterpillar reliably choosing cheese over apple (and hence always taking the same path to get its favorite snack, reaching out for the location where cheese was and not for the one where the apple was). In the last familiarization trial, the position of the cheese and the apple were swapped, and the caterpillar did not enter the room; this was supposed to exclude a prominent perceptual relation between the object and the location. In the seeing condition test trial, the caterpillar could see the objects that were placed as in the last familiarization trial; in the non-seeing condition, the caterpillar could not see them. Children reliably looked longer (hence, looked more surprised) when the caterpillar went along the "wrong" path, the one leading to the apple, when the caterpillar had seen the location of the objects. This was interpreted as demonstrating that the infant could take into consideration the caterpillar's perspective. In the second experiment, the caterpillar was not able to see the objects in either condition: this was supposed to make the caterpillar *perceptually* disconnected from the objects, in case such a connection was what was guiding the child's expectations. In the *knowing* condition, the caterpillar could see when the objects were placed behind the screens; in the *non-knowing* conditions, it would be completely ignorant about the location of the objects. In the knowing

conditions, participants were "surprised", i.e. they looked longer, when the caterpillar was taking "the wrong" path to the cheese; in the non-knowing condition, no effect was found. According to the experimenters, the joint results of the tasks suggest that children do attribute both false belief and true belief to the caterpillar. However, the conclusion is far-fetched. While it is certainly the case that children were making some association between the caterpillar and its goal, i.e. the cheese, and that some sensitivity to a goal-directed action is supported by the study, this falls short of providing hard evidence for any false belief reasoning. Contrary to the expectations of the experimenters, children did not look reliably longer in experiment 2 when the caterpillar was choosing a new path in the *non-knowing* condition. This, as they underlined, would have been rather strong evidence for the fact that children do attribute some false belief to the caterpillar. However, the fact that they showed no expectations is compatible with assuming the child was expecting the caterpillar to act on perceptual information (received before, but still perceptual) in the *knowing condition*, and not in the non-knowing condition: no appeal to beliefs is really necessary. Moreover, while in the first experiment there was a longer looking time for trials where the caterpillar was going to the wrong path in the *seeing-condition* (a difference of over 20 seconds), the pattern was only partially the opposite for the *non-seeing* condition (fewer than 9 seconds).

The interpretation of the implicit false belief task is not as straightforward as it is often presented in the literature. One of the problems with this kind of task is the violation of expectation paradigm, based on the assumption that a longer looking time implies a violated expectation, or surprise at an unforeseen situation. However, the validity of the paradigm itself has been discussed at length[23], since it is still relatively unclear how to interpret looking times in a straightforward way. Moreover, a more general worry regarding this kind of task is that looking time patterns might simply reveal that something in the scene is novel (Poulin-Dubois and Yott, 2017). Additionally, not all of the results of these paradigms have been replicated, suggesting that they might not be incredibly solid, as demonstrated by the groundbreaking report by (Kulke and Rakoczy, 2018)[24] which lists which of the most discussed papers on implicit false belief understanding have not been replicated.[25] Importantly, studies like Surian and Geraci

23 See for example Mukanata (2000).

24 But also Dorrenberg et al. (2018).

25 The report is one of the reasons why I am not dedicating a lot of space to the so-called helping paradigm, in which false belief attribution is measured by observing the helping behavior of children in response to the experimenter's actions. The results in Buttelmann et al. (2009) and Southgate et al. (2010) were mostly not replicated or only partially replicated. While this does not

(2012), for example, were not replicated directly, and only replicated conceptually in older children and adults (Meristo et al., 2012, 2016). This shows that lots of caution has to be dedicated to the evaluation of these paradigms and that, while the "developmental paradox" identified by many researchers might indeed be existent, and children might be able to attribute mental states in some form at a stage earlier than the classical view expects, the consequences might not be as groundbreaking as they first appeared.

A few words should be spent on the helping paradigm used for the implicit false belief task as well. The paradigm is used in studies at the center of the replication crisis (Buttelmann et al., 2009), and hence gave rise to a lot of controversies (Baillargeon et al., 2018; Poulin-Dubois et al., 2018). The general idea of the helping paradigm is that the child is presented with a situation in which she can help the actor or the experimenter solve a particular problem (e. g. reach a particular object) and that the helping behavior can vary depending on whether the child (successfully)attributes knowledge to the actor and the experimenter. In the study by Buttelmann et al. (2014) for example, 18-month-old children are familiarized with an adult playing with blocks to be found in a "block box". In the true belief condition, the actor is present when a spoon is found in a block box, whereas they are not present in the false belief condition; in both cases the box goes back to the shelf. Upon returning to the scene, the experimenter holds a bowl, and makes an ambiguous request to be handed the target box (now containing a spoon). Since the experimenter does not help, it is the child's turn to help the actor, and there are two objects at her disposal: a spoon and a block. The idea is that the child will handle the spoon in the condition in which the actor knows that the box contains a spoon, and a block in the false belief condition, hence demonstrating an ability to keep

mean their results are not interesting or worth further investigation, the absence of solid replicated results has to be taken into consideration, and I believe it is best to carry on the research, which focuses mostly on *explicit* false belief tasks anyway, without referring to these results when possible. However, notice that there is a heated debate about the importance of the replication crisis. It has been argued by Baillargeon et al. (2018) that this has been inflated, and that results are solid enough to be considered reliable, since the failed replications are actually due to methodological flaws. However, Poulin-Dubois et al. (2018) have convincingly argued that this is not the case, and that it is indeed impossible to exclude the possibility that the results in the studies are fundamentally false positives. The debate really deserves, in this sense, attention on its own. Being interested in false belief reasoning abilities the onset of which is during verbal age, and on how language interacts with explicitly measured mentalizing abilities, I will limit the treatment here to what is necessary.

track of the actor's knowledge[26]. The results in Buttelmann et al. (2014) are encouraging because they show that infants reliably hand over a spoon in the true belief condition and a block in the false belief condition. However, the helping paradigm is highly discussed. Poulin-Dubois et al. (2018) for example, point out that the results in precedent studies by Buttelmann et al. (2009) have been only partially replicated or showed no replication in other studies, as for example Priewasser et al. (2018). Most of the critics are based on the Buttelmann et al. (2009) study and not on Buttelmann et al. (2014), leaving the discussion open. In general, however, at least one observation used for Buttelmann et al. (2009) still applies: in Poulin-Dubois et al. (2018) and Priewasser et al. (2018), for instance, the authors argue that a teleological interpretation is possible, i. e. that the child does not need to entertain explanations about false belief, but only to understand that objective reasons drive people's behavior. In the location task, the target toy changes location, with or without the actor present, and then the actor comes back and tries to open the box where the toy was: the child either helps to open the box where the toy is, or continues helping to open that box. The explanation given by Priewasser et al. (2018) is the following:

> In the FB condition E2 is continuously playing with the toy (which makes her happy and thus for a pleasant, desirable situation). That E2 has to interrupt her play in order to fetch a key suggests that she will continue to play with the toy on her return. Hence, enabling her to do so will make for a better situation (shell be happy; a goal to be achieved) than preventing her from doing so (shell be nervous and grumpy). Consequently, when she is looking for the toy in the wrong box children have good reason to help her find the toy in box B to achieve a better situation (a goal). In the TB condition the agent interrupts her play with the toy and watches the child and E1 play with it and then briefly moves away to close a door. This gives less clear indication that the agent is likely to resume her play with the toy. So when she tries to open the empty box the teleologist child perceives no compelling reason to direct her to the box with the toy. (Priewasser et al., 2018, p.71)

The explanation above does not include reference to beliefs. Now, in the case of the block boxes, it might be that something similar is at play; since the actor leaves the room apparently when the game is still going, the child might infer that the actor wants to keep playing when he/she comes back. However, if the actor leaves the room after a spoon is found, this might be read by the child as a sign that the game is over, thus making the spoon the most salient thing for the actor. Regardless of the the teleological reading, another point is relevant: of the 57 children tested in Buttelmann et al. (2014), 21 children did not count

26 Note that in Buttelmann et al. (2009) the helping paradigm is applied to a change of location and not to an unexpected content task.

towards the final analysis because they did not understand the task, i.e. they gave objects to the parents, kept them for themselves, touched them, etc. This might be a sign that the task is not that well understood, and weakens the significance of the fact that 66.7% of children acted according to the experimenters' predictions in the two conditions. (18 children were left for each condition.) The small number of participants and the exclusion rate might indeed present some problems[27]; however, since the results seem to be in line with other studies, the most reasonable thing to do is probably to wait for replication and further discussion of the study to verify the validity of the paradigm. Finally, it is not completely clear that the teleological and mentalistic interpretations are the only options; another possibility is that the children mostly acted by giving the actor the last object they thought was salient for them. In the true belief condition, the actor comments with the experimenter on the presence of the spoon; in the false belief condition, the last object the actor focuses attention on is the block. It is not to be excluded that children dealt with a more general association of this kind.

Despite all of these problems, at least some points are clear. Firstly, depending on the task, different results can be achieved: explicit FBTs, where an explicit prediction of behavior has to be made, seem to be more complex for children, who reliably fail before their fourth birthday. This seems to be true independently of how the task is carried out; however, implicit FBTs, where no explicit formulation of a prediction is required, sees to bring different results. Crucially, the difference cannot be explained solely in terms of the verbal component. In Scott et al. (2011), children's preferential looking time was measured while they were listening to stories about false belief while looking at pictures. In this case, while the stimuli was in verbal form, 2-and-a-half-year-olds still performed well, suggesting that it is not solely language that makes a difference in their performance. The second part of the chapter will be explicitly dedicated to the data about the relationship between language and false belief, so this will be left aside for the moment. However, the most important takeaway of the section is that data seem to suggest that solving tasks involving false belief and false perception is not an all-or-nothing problem, which has been a challenge for many theories, as will be explained. Before diving into data about language specifically, I will first outline other problems that lie at the core of the literature about mentalizing.

---

27 Note that the authors say, "In addition, for what it is worth, including the 11 excluded infants who chose an object for themselves or their parents gives a similar pattern of results." (Buttelmann et al., 2014, p.125). This does not clear the air on how solid the finding is, but it is a bit reassuring.

4.2.3 Representations

4.2.3.1 A specification about concepts of mental states: the folk-psychology way

Another matter of debate, considering the task above, is what kind of representations are necessary to solve the FBT. The questions "What exactly is a belief?" and "What exactly is knowledge?" are obviously loaded with epistemological issues. Difficulties arise when recalling the standard definition of *knowledge* as justified true belief, as it brings with it a complex debate on the validity of the definition (Steup and Neta, 2020). However, as relevant as it is, what is at stake here is slightly different, since the focus is on how beliefs and knowledge are understood in everyday life and used in folk-psychology explanations, and when this happens during childhood. Now, it is of little doubt that epistemologically complex concepts used in philosophy (and in science) have a relevant social role and that they are heavily influenced not only by linguistic practices, but by a complex theoretical net that belongs to the theories that involve them. In other words, the question of whether the concept of truth as justified belief is dependent on language is, although worth asking, probably trivial, since it depends on a definition. In what follows, however, our focus will be on the concepts of *belief*, *desire* and *knowledge* that are involved in folk-psychology explanation.

4.2.3.2 Domain specificity

One of the debates inherently connected to all the issues treated so far is that of the representations involved in mindreading. According to a "domain-specific approach", dealing with mental representations involves a specific mechanism, different from other representational and metarepresentational abilities, and which deals with different sets of problems compared to other mechanisms. The opposite approach, which is "domain-general", sees problems arising from mental representation processing as ones that are to be read in light of a general-domain capacity for dealing with (meta)representations. In other words, according to a "domain-general" approach, what is really at stake in the false belief task is the ability to deal with (meta-)representations in general.

The literature regarding this issue is rich in presenting a very diverse array of tasks that are supposed to be compared or comparable to the false belief task; the idea is that comparing performance across these tasks and FBTs can help to disentangle what kind of representations are involved in mentalizing. In what follows, I will examine the various tasks and the related evidence.

4.2.4 False photography, false sign, false note

The false photography task was the first to be compared to the false belief task, and was the starting point for a lot of research regarding which kind of cognitive ability underlies passing the false belief task. The false photography task was proposed in a study by Zaitchik (1990), in which preschool children were tested with a "false representation" problem. In the study, children were shown a scenario, in which an object is placed in a determined location. The experimenter then took a Polaroid picture, and the object was moved to another location. The children were then asked where the object would be in the picture, which was still developing. It is easy to see the similarity between this task and the false belief task (at least in the change of location version): there is a representation that does not match reality anymore, and the child is asked to recognize the difference between the representation and reality. What is interesting about the results is that preschoolers, in Zaitchik (1990), found the false photography task as hard as the false belief one. The result was used to argue against the "domain-specific" approach; children under the age of four, it was argued, have a problem with manipulating representations in general, rather than a more specific problem with accessing information about other people's mental states. These results were replicated by Perner et al. (1998).

Despite the intuitive similarity between the two tasks, caution should be exercised. In the case of the false belief task, there is a mismatch between a representation's content and the thing that is supposed to be represented (in the false belief location task, the location of the object for example). In other words, the representation is inaccurate, or wrong. In the case of the false photography task, something slightly different seems to happen: the photograph, as pointed out by Perner Perner and Leekam (2008), is actually not false in the same sense, since it simply represents the landscape as it was before, and is therefore not properly *misrepresenting* the situation.

Several other things should be noticed: ASD children perform poorly with beliefs but fairly well with photos (Leekam and Perner, 1991), so the impairment might be *domain-specific*, as argued in Leslie and Thaiss (1992). Moreover, fMRI evidence that a specialized circuit reacts more to beliefs than to photos was found (Saxe et al., 2004), which seems to confirm this conclusion. At the same time, *false sign* tasks also imply difficulties for young children (Leekam et al., 2008), and their similarity is supported by neuro-imagining studies (Aichhorn et al., 2009). In a false sign task, children are confronted with a sign that points to an object in location 1; the object is then moved, but the sign is not, so it still points in the same direction. Perner and Leahy (2016); Perner (1991) argue that there is a general mechanism concerned with metarepresentations at play, not a

domain-specific one, on the basis of the fact that the false sign shares with the false belief the fact that the sign *misrepresents* reality, in a way that the false photography does not. Comparing false belief, false sign and false photography, Leekam et al. (2008) argue that this is reflected in the performance of pre-school children: while their performance in the false belief and the false sign task correlates in a statistically significant way when age and performance in the false photography task are controlled as variables, the same is not true for their false belief task performance and their false photography task performance once the false sign task performance was partialled out. Moreover, they argue, the fact that the skill is domain-general is confirmed by the fact that most of the children performing very well in the false belief task were performing at the same level in the false sign task.

This result has been challenged by Cohen et al. (2015), in which evidence is presented that the FBT is still different from other false belief representation tasks like the *false note*. Testing adults, Cohen and colleagues compared the reaction times for performance in false belief and false note tasks. Participants had to watch a video in which an actress put a purse belonging to a friend in a drawer and wrote a note that communicated the new location of the purse. When the actress was gone, a second actor would enter the room and either move the purse or leave it in the current location. The video was then interrupted and a test probe was given, where participants had to press "yes" or "no" evaluating statements like "The girl thinks that the purse is in the right drawer" and "The note shows that the purse is in the right drawer". In the second experiment, the probes were instead "In her mind, the purse is in the right drawer" and "In the note, the purse is in the right drawer".

Consistently, there was an advantage for processing the content of a belief over the content of a note, the authors argue, because response times were much faster for answering belief-centered questions. However, there are several worries. Among the "possible objections", the authors list contrastive use of the verbs "believe" and "know": participants might have had a facilitation for false beliefs because "think" suggests inaccurate knowledge, in contrast to "know". The authors argue that the fact that "think" was only used in two of the four experiments excludes this possibility, since experiment 2 and 4 used "In her mind" instead. However, I think a delay in performance for the false note task might actually be due to the choice of language. Consider the target probes, which contrasted in this way:

1. *She thinks* versus *The note shows* for experiment 1 and 3.
2. *In her mind* versus *In the note* for experiment 2 and 4.

There is a strong evidentiality contrast between "thinks" and "shows". While "think" suggests the possibility of false representation, "shows" seems to indicate a completely different degree of evidentiality, such that one expects that something that is shown is true. As discussed in chapter 3, mental state verbs do carry enough information on a pragmatic and semantic level to raise similar worries. A similar point can also be made for the contrast used in the rest of the experiments described in the studies. While "In her mind" is a common expression, accepted by native speakers even without a context, informal investigation led me to think that "In the note, the purse is in the right drawer" appears, on the contrary, to be slightly less acceptable to native speakers, whose judgments were not as positive. Obviously, this kind of worry needs further empirical and theoretical investigation. However, it is useful to remind ourselves that stimuli for these tasks always needs to be carefully chosen and handled.

In summary, the debate about the equivalence of the above tasks is in many ways still open. While tasks like the false photography, false sign, and false note tasks provide good comparison for FBTs, in that they help identify different features that might be of relevance in its resolution, there is a sense in which they do not seem to be equivalent. At least so long as we consider FBTs to test skills that are at play in *social* interaction, skills which are related to human behavior as socially driven and to the ability to explain one another's behavior, it is obvious that they are going to be in one sense specific: while mechanisms of metarepresentations might be at play, they are likely to be accompanied by other processes that are more strictly related to social interaction. This is especially true if there is anything realistic about simulation theories, which see recruitment of one's own resources in order to understand other people's behavior. However, as will be seen below, the fact that mentalizing skills tend to be influenced by other factors is also a sign that considering the mechanisms responsible for it as a modular unit is far-fetched, and of course making these representations related to specific mechanisms does not make them necessarily different in *quality*, or format. As will be seen in the remainder of the chapter, data on the influence of language on social cognition tasks shed further light on the issue.

4.3 Language and social cognition data

The first part of the chapter has given a picture of the debate about mentalizing and what the various issues underpinning it are, giving an idea about the false belief task as one of the main tests for mentalizing skills. In the rest of the chapter, attention will be given to the main theme of the book, i.e. the role of language in mentalizing tasks.

4.3.1 General linguistic abilities and semantics

Firstly, it is worth considering data suggesting a relation between general linguistic ability. In this sense, extremely valuable work comes from the meta-analysis conducted by Milligan et al. (2007). What makes the meta-analysis valuable is not only the fact that it puts together more than one hundred studies on the relationship between language and false belief understanding but also the theoretical approach, which is extremely in line with what has been claimed so far about the possible role of language in cognition. Language, the authors argue, is a multi-faceted, complex set of skills, and even measures that are meant to capture the most general language abilities will have to take this diversity into account. While vocabulary measures can vary in tracking production and comprehension of single words, other tests provide a measure of syntactic abilities, and general language measures combine semantic and syntactic abilities. The meta-analysis then takes into consideration three fundamental factors: 1. the type of language ability involved; 2. the kind of false belief task; and 3. the direction of the effect.

The meta-analysis found a significant correlation between language ability and false belief task understanding in general, with no influence exerted by demographic factors like male/female ratio and mean age. Five types of language ability were tested: general language, semantics, receptive vocabulary, syntax and memory for complements. Receptive vocabulary indicates the appropriate response to vocabulary, often tested with yes/no questions, even in the absence of production vocabulary; semantics is tested with different measures that vary from word order tasks, synonym judgement tasks, and more general semantic item tests like the Bankson Language Test, which covers expressive vocabulary related to body parts, nouns, verbs, categories, functions, and opposites (Bradley-Johnson, 1991). Syntax is assessed by a variety of tests, from reception of grammar to specific tests dedicated to embedded clauses. Memory for complements is specifically measured with complement comprehension tasks and memory comprehension tasks, more on which will be said below. The analysis found an effect for each ability in this order: receptive vocabulary, semantics, general language, syntax, and memory for complements. Four kinds of false belief task were considered: change of location, unexpected identity, deception task, and belief emotion. Note that all of the tasks were verbal tasks, albeit with very different linguistic demands, since some involved a narrative, some sentential complements, and some none of the above, given that the only verbal component was a question. The analysis did find a significant relation between language ability and false belief task performance for all the task types, but no significant difference between the types of false belief task.

The results are definitely suggestive of a complex relation between language and false belief task performance. The overall relation showed that language abilities account for 18% of the variability in false belief task performance across all types (10% once age was controlled for); however, there is also a strong variability across the studies, with a proportion of variance explained by language that goes from 0% up to 77%. Additionally, the higher the number of different false belief tasks for studies, the greater the effect size, which led the authors to speculate that using different types of false belief tasks maximizes the possibility for children to show their understanding of false belief reasoning. Moreover, receptive vocabulary was found to account for 12% of the variance, semantics for 23%, general language for 27%, syntax for 29%, and memory for 44%, but post hoc pair-wise analysis revealed only one significant difference between language ability assessment, which was between general ability and receptive vocabulary. The authors point out that this might be due to the fact that receptive vocabulary tests are the only ones that actually provide a single-ability test measure, since for example syntactic tests require some semantic knowledge, and semantic measures often test vocabulary at the same time. The results of the analysis also showed that the effect was bi-directional but stronger on the language false belief task side: that is, while performance on FBTs was predictive of language abilities developed later on to some extent, the reverse pattern was much stronger, with language abilities having higher predictive power with regards to performance on the FBT.

All in all, the results of the meta-analysis do suggest that language influences performance in the FBT; obviously, what exactly it is about language that allows this and what the underlying mechanism is seems to be harder to determine. The results reported in Milligan et al. (2007) confirmed those of other studies, like the previous meta-analysis conducted by Wellman et al. (2001) and the results in Astington and Jenkins (1999), where a longitudinal approach was taken to verify the correlation between language development, assessed thought tests for production and comprehension of both syntax and semantics, and FBTs and reality-appearance task; in this case, language acquisition predicted performance on FBTs, but the opposite pattern was not found. In a study by Jenkins and Astington (1996), children aged 3 to 5 were tested for general language ability, and their performance in FBTs. Interestingly, the analysis brought up several relevant effects: similarly to Milligan et al. (2007), age was responsible for only part of the variance, and general language ability was a predictor of the performance when age was controlled for. The same was true for verbal memory skills, but not for non-verbal memory. Even more interestingly, the effects of language were checked in relation to family size as another control factor, with the relevant result that, while family size was predictive of better false belief task

performance in general, it was more predictive for children with lower language ability scores: according to the authors, this suggests that the presence of siblings might "compensate" slower language development as far as the developing of false belief reasoning is concerned. This is particularly interesting because it brings up a factor that, as will be seen, is particularly relevant for the current topic, which is social interaction.

The relation with siblings has been investigated in other studies as well. In a famous study, Perner et al. (1994), results with 3- and 4-year-olds on false belief tasks supported a correlation between family size and false belief reasoning abilities: the authors suggested that having a sibling might be a testing ground for social interaction and mentalizing abilities. Interestingly, a study that replicated the results, showing that children with siblings perform better than single children, also showed that children with twin siblings are also outperformed by children with non-twin siblings (Cassidy et al., 2005). The authors speculate that the reasons might be varied, from the fact that twins develop idiosyncratic interaction patterns to the fact that the household's dynamic might change. In any case, non-twin sibling interactions seem to facilitate the development of mentalizing abilities. Moreover, results in Peterson (2000) support the idea that the presence of siblings of a relatable age (i. e. between 12 months and 12 years) has a positive impact on mentalizing abilities in children aged 3 to 5, in contrast to having toddler siblings or teenager/young adult siblings, suggesting that the possibility of interacting in play has a significant role.

Other studies suggest that general language ability has an impact on FBTs. In a series of studies (Ruffman et al., 2003, Ruffman, Slade and Crowe, 2013), four age groups were tested for semantic abilities, syntactic abilities, two kinds of FBT, an emotion-situation task, a desire-situation task, a desire-action task, a wicked desire action task, a second-order FBT, and a display task. The two false belief tasks were a change of location and an unexpected content FBTs. In the desire-situation task, they had to guess the emotional state of a character finding their favorite animal or an animal they do not like. In the emotion-situation task, they had to attribute emotional states to a character in stereotypical situations. In the wicked-desire task, they had to guess the emotional state of a character that missed or hit another character he did not like with a ball. In the display task, they had to guess the right display of emotions for a character hit by a ball who wanted the classmates to think he was brave. Finally, in the second-order-belief-task, they had to attribute a second-order belief, e. g. "The boy thinks that the girl would do x, because.. ".

The results were multi-faceted. Language ability was related to belief understanding over a 2.5 year period, with semantic abilities being a good predictor of variance in belief task but not in desire tasks, and syntax alone did not explain

unique variance for any of the tasks. Note that all of the syntactic measures used in the experiment were word-order-based in the first experiment, as this was considered by the authors as the best way to tap syntactic abilities. In the second experiment, embedded sentences were also tested. Children had to show comprehension of embedded predicates by linking sentence like "The blue ball is under the square" to picture (in the embedded-sentence task); they had to show understanding of word order by linking sentences like "The horse chases the zebra" to the right picture (in the word-order-task). They also had to guess the right word to describe the emotional state of a character that found their favorite animal or an animal they did not like (in the semantic-task) and finally they had to perform in an emotion-recognition task where they needed to connect specific emotions to some pictures. In addition, they performed in two location-change FBTs and in two content-change FBTs. Semantic and syntactic test performances actually highly correlated with one another, making it hard to disentangle the measures. While syntactic performance did not appear to be uniquely related to false belief reasoning in the first experiment, it appeared to be more relevant in the statistical analysis in the second experiment; moreover, it did also correlate with emotion task performance, which remains unexplained, and constitutes for the authors the main reason for discarding a role for syntax alone. Location change performance was explained by embedded syntax by 3% variance, word order syntax by 6% and general language by 30%, whereas content task performance was explained by the embedded syntax task (5%) and general language (7%) and emotion task performance was explained by embedded syntax (4%), by word order syntax control (6%), and by general language by a much higher 30%. All of this data suggests, according to the authors, that general language is a better predictor for the performance.

An essential point brought out by the study is that syntax and semantics are hard to dissociate in language, since language as a dynamic system relies on a consistent interrelation between information about structure, about meaning, and interface between the two. Results along the same lines were presented by a later study by Slade and Ruffman (2010) in which, once again, different tests for semantic and syntactic comprehension and working memory were used and the relation with false belief reasoning was tested. The results confirmed the influence of language abilities on mentalizing performance excluding, moreover, the possibility that working memory explains the relation between language acquisition and false belief reasoning development; the study also found no support for syntactic-specific predictive power for mentalizing abilities, which as it will be seen is a central point for specific theories like the syntactic bootstrapping hypothesis and is not in line with some of the data presented below.

As will be seen, particularly telling in this case of inquiry are those studies that investigate lesions and specific impairments. What analysis like Milligan et al. (2007), Ruffman et al. (2003), Rufflman, Slade and Crowe (2003), and Wellman et al. (2001) do suggest with a good degree of certainty, however, is that a certain degree of influence of language on the development of the explicit false belief task is present.

4.3.2 Syntax and complement structures

The already cited analysis in Astington and Jenkins (1999), in addition to showing the impact of general language abilities, also focused especially on the role of syntax. The test for linguistic skills was TELD (Test of Early Language Development), which was divided into syntax and semantic components, finding that syntax accounted for variance in FBT performance independently from semantics, but that the opposite was not the case. The results were challenged by Ruffman et al. (2003), arguing that TELD does not differentiate semantic and syntax factors enough, since many syntactic tasks do require semantic knowledge, as in the case where verbs have to be inserted in the blanks of a sentence: clearly, semantic knowledge is required to complete the task as well. Note however the fact that the "more syntactic" TELD components were still found to be influencing FBT performance in a more independent way than the semantics components; while it is certainly true that semantic and syntactic components are intertwined, the fact that purely semantic measures did not score the same might be telling. At the same time, tasks used in the TELD assessment that required the child to listen to sentences like "The car hit the truck" and choose between two different pictures, one representing a car hitting a truck and one representing a truck hitting a car, seem to tap mostly syntactic information and might therefore be considered a good measure of syntactic abilities.

Referential opacity was at the center of studies like the one in de Villiers and Fitneva (de Villiers and Fitneva).[28] An example story from the task was the one of Sarah, whose mother prepared a gift, a silver box filled with candy. In this context, since Sarah can see the box but is ignorant of the content, a verb like *move* allows referential substitution, but a verb like *think* behaves differently:

(1) Mum moved the silver box.

28 Unfortunately, the study was never published, so I will focus on other results in this paragraph, but I will describe the task reported because it is used in other studies, that will be cited below.

(2) Mum moved the candy.

both have the same the truth value, but:

(3) Sarah knew the box was on the table.

(4) Sarah know the candy was on the table.

do not. Children, after hearing the story, had to reply to questions containing communication verbs, mental state verbs, and ordinary verbs, about either the mum or Sarah. Preschoolers found this task difficult, suggesting according to the authors that difficulties with the opacity of the verbs reflect an immature understanding of mental states.

The relationship between complement structures and false belief reasoning has been highlighted in a series of studies conducted mostly by de Villiers and her group, where training was introduced. In de Villiers and Pyers (2002), a longitudinal study is reported which explores this relation. Children had to perform a *memory for complements task*, a location change task, and an unexpected content task. In the memory for complements task, children had to report on what characters had said, after listening to stories that contained embedded contents. In this case, communication verbs were used. Crucially, a correlation between performance in the memory for complements task and performance in standard FBTs was found. Moreover, controlling for FBT performance three months later gave similar results, showing that memory for complements predicted FBT performance. The FBTs used in the study were of different kinds, and included the unexpected content one developed in Perner et al. (1987), a location change FBT similar to Wimmer and Perner (1983), and an *explanation of action* task: in this case, the child had to explain the actions of a puppet which, for example, wants to cook eggs and therefore looks in an egg box. Since the eggs were removed from the box when the puppet was away, the child needs to recognize the intention of the puppet and his action as motivated by the belief the eggs are in the egg box. Note however that the child did not need to give a mentalistic explanation, and can provide an explanation like "Because the eggs were in the egg box" and still pass the task. In the memory for complements task, children had to hear stories and report back on the embedded complement as below:

(5) He thought he found his ring, but (second picture) it was really a bottle cap. What did he think? (Pointing back at first picture.)

(6) She said she found a monster under her chair, but (second picture) it was really the neighbors dog. What did she say? (Pointing back at first picture.)

Note that, while these stories and questions sometimes involved mental state verbs, they also made use of communication verbs. Spontaneous speech by the children was also recorded to register the use of complex sentences. The memory for complements measure not only predicted false belief performance in the same round, but was also predictive in the second round of testing three months later, whereas the opposite was not true (the false belief performance was not predictive of the memory for complement performance); also, performance with memory for sentential complements with communication verbs was the most predictive variable for FBT performance, suggesting that the results were not due to a semantic component related to the verbs *know* and *think*.

This line of research has generated a lot of material and has been used to argue for a specific theory that will be discussed in later chapters (de Villiers, 2007, 2001; de Villiers and de Villiers, 2011, 2012; de Villiers and Pyers, 2002). Also, as will be seen, important replications are to be found for deaf children in de Villiers and de Villiers (2011), so they will be dealt with later in this chapter.

While results in correlational studies are good evidence that an interesting relationship between language development and FBTs should be investigated, they provide little information of what this relation might be, and they only hint at a possible casual connection. Longitudinal studies, on the one hand, provide additional insight since they allow us to follow the chronological development of different skills. Training studies, finally, provide us with stronger evidence. In a famous study, Hale and Tager-Flusberg (2003) trained children in the use of sentential complements and compared their performance in FBTs before and after the training, finding a significant improvement in the performance. The study was designed to test sentential complements specifically, and for this reason communication verbs were used.

Three groups of children aged 3 to 4 were trained with *Sesame Street* puppets. In the false belief group, children had to predict where the puppet would look for an object that was moved in his absence; in the sentential complement group, a *Sesame Street* character said he made for another puppet something that he actually made for a third puppet, and children had to report on *what* the puppet said he did to *whom*. In the relative clause training group, a scene with identical twins and a *Sesame Street* character was enacted, and children had to report which twin was the recipient of the action. Before and after the training, children had to perform in a standard location change FBT and a sentential complement test in which they had to report back on what a character said they did, and after the training they performed in a location change FBT, an unexpected content FBT and an appearance reality FBT. Results showed that training on sentential complements and on FBTs, but not on relative clauses,

did have a positive effect on FBT performance. Among other aims, the study wanted to rule out some difficulties pointed out by the authors in the design of two extremely influential training studies, namely de Villiers and Pyers (1997), where the measure of sentential complements included the use of mental state verbs; this according to Hale and Tager-Flusberg (2003) was avoided in their study, where the focus was on the content of the sentences uttered by the characters, and not their mental states, avoiding in this way having to train children about deception and the falsity of mental representations and thus potentially invalidating the results.

Something very similar can be said about the training study presented in Lohmann and Tomasello (2003), in which three groups of trained preschool children were compared. Children were familiarized with objects which were deceptive, as they appeared to be one thing but were instead something else (e.g. a pen that looked like a flower). In the *full training* condition, children were trained with sentential complement constructions with mental state verbs and communication verbs, and also trained in noticing deceptive characteristics of objects; in the *discourse only* training, no sentential complements were used, but deceptive characteristics of objects were highlighted; in the *sentential complement only* training, sentential complements were used but no deceptive characteristic was highlighted; finally, no verbal descriptions were used in the *no language* training group, where the experimenter would only make comments like "Look!", "Oh" and so on. After and before training children were tested for sentential complement comprehension and in FBTs. All groups except for the *no language* one increased their performance in the FBT, with the full training group outperforming every other group, and the sentential complement condition providing better performance then the other conditions. Perspective-shifting discourse and the sentential complement condition were found to be independently facilitating FBT performance, which suggested according to the authors that syntax training was providing sufficient facilitation to aid false belief reasoning. Language was a necessary condition in the experiment for an increase in performance in FBT, whereas experiencing the deceptive objects was not sufficient. Training in sentential complements was also sufficient to increase performance, but the combination of sentential complements and discourse-shifting measures to highlight deceptive characteristics of the objects was the strongest facilitator. Lohmann and Tomasello (2003) stress how, while the contribution of syntax seems to be sufficient for an improvement, it might not be necessary, since a condition highlighting deceptive characteristics of objects was already sufficient to promote an improvement in the performance. Notice, furthermore, that Tomasello's studies are important because one objection to de Villiers' studies made by Slade and Ruffman (2010) is, on the one hand, that statistical analysis might not be accu-

rate, and on the other hand that semantic measures for comparison were absent; however, the results from Tomasello present good supporting evidence, on the one hand, that syntactic influence is indeed significant, and that it holds in comparison with semantic variables, albeit with limitations, on the other hand.

In combination with the results on semantics and general language abilities, these studies further reinforce the finding that the development of linguistic skills and the development of mentalizing skills proceed hand in hand, and they might be interrelated on a deeper level. While syntactic skills might not be the only component that is relevant, it appears that a facilitation is indeed present when syntactic mastery is achieved; moreover, while the longitudinal studies and the correlational studies only provide partial evidence for this link, training studies offer a way to observe the direct impact of linguistic training in false belief understanding. Note that this brings us back to an issue that was treated in the earlier chapters, i.e. that of the constitution relation of language in cognitive processes as opposed to that of an *enhancing* role for language. In other words, when considering the role of language in mentalizing skills, one can adopt two different perspectives: assume that mentalizing processes are at least partially *carried out* in natural language, and hence assume that linguistic coding is constitutive of mentalizing processes, or assume that language can influence the processes, possibly enhancing them. As it will be seen, there are reasons to lean towards the latter; this way is also already indicated by the fact that, as underlined in the first part of this chapter, mentalizing skills like false belief reasoning seem to have important precursors and to be realised well before the age at which children acquire significant linguistic means. However, further evidence regarding this will be presented in the following sections. Another thing to be noticed, which will become clearer with further data, is that training studies are good support for the idea that linguistic skills and false belief performance are not simply the emerging result of the same underlying mechanism; this can be said because statistical analysis found that linguistic skills predict false belief reasoning performance, but not the other way around. This kind of data supports the idea that correlation is not only the result of skills using the same underlying abilities, but a matter of how language contributes to false belief reasoning.

4.3.3 Pragmatics

The relationship between pragmatic abilities and false belief reasoning is usually investigated in a direction that is somehow the opposite of the scope of interest of this book; while my main focus is the role that linguistic processes have

with respect to development of non-linguistic processes, proposals have been made in the literature that we consider pragmatics as part of the "Theory of Mind" module (Wilson and Sperber, 2002). This means asking a partially different question, which concerns the extent to which processes in the realm of pragmatic interpretation are a part of more general mechanisms devoted to understanding others. However, notice that this entails not only a specific view of mentalizing, i.e. one which envisages *modular* ToM, but also a specific view of communication and pragmatics. The idea is Gricean in character, where language use is finalized to address and modify agent's mental states, which is, as mentioned in chapter 2, a non-neutral view. In this sense, then, this sort of inquiry is not pertinent to the fundamental questions and approach of this work. However, from this body of literature, interesting results can be discussed.

In a study by Bosco and Gabbatore (2017), for example, the relation between first-order and second-order false belief understanding was studied in relation with the ability of children to deal with sincerity, irony, and deception. Along with a battery of mentalizing tasks, children aged 3 to 8 also had to take part in pragmatics tasks, either comprehension-based ("In your opinion, what did the girl want to say to the boy?") or production-based ("The child does not want to be discovered. What could he say?"). Results showed that, unsurprisingly, age explained a good amount of variance in the performance in the pragmatic tasks for children, especially in the deceitful tasks. However, a relation with performance in false belief reasoning was also found. Controlling for age, a correlation between overall mentalizing task performance, both first- and second-order, and linguistic and extra-linguistic irony, and deceit tasks was found. Second-order false belief tasks correlated significantly only with linguistic deceit tasks, but first-order false belief performance correlated with linguistic and extra-linguistic irony and deceit. Moreover, first-order false belief was found to be predictive in performance in the comprehension and production of sincere and deceitful communicative acts.

The results are in line with those reported in Talwar and Lee (2008), where children's better performance in first-order false belief tasks predicted their tendency to lie in a task in which they were instructed *not* to peek at a toy: children showed a tendency to hide their own knowledge in general, but this was correlated with their performance in the false belief task. Interestingly, children who performed better in the second-order false belief task were also predictably better at hiding their own lie when questioned about the toy later, suggesting that second-order false belief tasks are related to more sophisticated pragmatic and social skills. According to the authors, development of first-order false belief understanding is also related to the emergence of so-called "secondary-lies", i.e. lies used by children to conceal their own transgressions. These seem to emerge

between 3 and 4 years of age, which as should now be clear is also the age at which children drastically improve their performance in explicit FBTs.

The complex relation between pragmatic abilities, false belief understanding and language development emerges even more importantly in studies like that in Angeleri and Airenti (2014), where receptive vocabulary scores were found to be predictive of performance in both irony comprehension tasks and false belief understanding tasks, suggesting in line with the other studies above, that language abilities develop at least in parallel, if not in alliance with, complex social skills.

As will be seen, evidence regarding pragmatic abilities and mentalizing comes from schizophrenia, and it will consequently be discussed in the section on clinical studies. However, notice that the studies above already suggest something that will be stressed again at later points, i.e. that language acquisition and other social behaviors are hard to disentangle and are likely to make the picture of what it is in language which contributes to the development of mindreading skills more dynamic and complex than what it seems.

4.3.4 Bilingualism

Some studies have investigated the relation between acquiring two languages and false belief reasoning. While these results are relevant as they cast new light on the issue of whether linguistic coding and information can interact with non-linguistic forms of cognition, the focus of the research is mostly on whether bilingualism grants speakers more inhibitory control in some tasks due to code-switching, for example, which is a different issue. However, it is worth focusing on some of the results that are relevant for the issue at hand. For instance, English monolinguals, Mandarin monolinguals and Mandarin-English bilinguals were tested by Goetz (2003) in an appearance-reality task, in level 2 perspective taking and FBT. In the appearance-reality task, children were shown deceptive objects and, once familiarized with their double nature, were asked questions about what the objects looked like and what they actually were. In the level 2 perspective task, children had to guess how the experimenter was seeing pictures of animals, taking into consideration the difference in perspective, and the FBTs were of the type unexpected content and location change. The standard age group effect was found, with children aged 4 performing considerably better than those aged 3, but the results were particularly interesting for the comparison between monolingual and bilingual children: while Mandarin and English monolinguals performed similarly in the false belief and perspective taking task, bilinguals showed an advantage in both. The study was taken at two different points in time, and interestingly, while monolingual children im-

proved over time, performing better at time 2, the same was not true for bilinguals. However, this was not true for the location change task. The authors speculate that different factors might cause the different performances. One possibility is that bilingual children possess higher inhibitory control, given code-switching; the second possibility is that children that speak two languages are more accustomed to perspective taking, since they are trained in having different labels for different objects; finally, bilinguals might be accustomed to adjusting their own expressive means to those of others, and so be socially trained to take into consideration other speakers' mental states as well. However, the author recognizes that no evaluation was made of the cognitive skills of the children, thus leaving questions open about other factors that might interfere. Similar results, in any case, were found for Hungarian Serbian bilinguals, compared with Hungarian monolinguals and showed better performance in a battery of mentalizing tasks in Javor (2016). Note however that in a study by Kyuchukov and de Villiers (2009), bilingual Roma children were not found to perform better in mentalizing tasks compared to the monolingual group, thus providing contrasting evidence compared to Goetz (2003).

Further evidence comes from Nguyen and Astington (2014), where English monolinguals, French monolinguals, and English-French monolinguals were tested on FBTs, working memory and verbal skills, explicitly selecting children with equivalent parental income and education. Interestingly, the study took into consideration the fact that bilinguals normally have worse language ability scores compared to monolinguals, even though this difference normally is canceled out with time (Bialystok et al., 2010; Mahon and Crutchley, 2006). While no difference between monolingual and bilingual speakers was found in the first analysis, bilinguals did significantly outperform monolingual speakers once age and verbal ability were partialled out. The same was true for working memory, where scores were higher for bilinguals when controlling for age and verbal ability. This suggests, according to the authors, that a compensatory mechanism is at play for bilingual children, since their advantage in working memory might be what accounts for normal performance in FBTs, despite their lower linguistic scores. However, the authors admit limitations in the study given that no control in general intelligence was given, and once again this might imply that some relevant factors were not considered in the analysis. Note that the same advantage for bilinguals after controlling for language proficiency was found by Diaz and Farrar (2018), who tested Spanish-English bilinguals and English monolinguals.

While the results seem to be somewhat controversial, the fact that bilinguals' performance in mentalizing follows a slightly different path compared to monolinguals seems to be a confirmed result. A meta-analysis was carried out by

Schroeder (2018), taking into consideration sixteen different studies. The analysis found a small bilingual advantage in a variety of mentalizing tasks, thus confirming that an effect is present, although not deciding between the various factors that might cause it. This is of course a fundamental issue, since the different interrelated cognitive developments resulting from acquiring two different languages as a child might be very complex. Since acquisition of language is slightly delayed in bilinguals, and yet their performance in false belief reasoning does not seem to be consistently delayed, the results seem to point to a role for language that is not constitutive of false belief reasoning. However, two things have to be noticed: analysis of the specific language skills have not been conducted, so the results are not as accurate as some of those achieved in respect to monolingual children; moreover, the studies presented above mostly did not include measures of cognitive development and non-verbal intelligence. In summary, the research on bilingualism is promising but needs further investigation and more detailed disentanglement of different factors. In the next chapter, I will come back to the various hypotheses that can be made about the role of language (as a socio-cultural means of interaction, or as a labelling perspective shifting device) mentioned in Goetz (2003), since they are interesting for the central topic of this book.

4.3.5 Cross-linguistic and cross-cultural studies

Cross-linguistic and cross-cultural studies also shed light on the relationship between false belief understanding and mentalizing and language.

Mandarin has been the subject of attention for the false belief understanding debate because of studies like the ones about sentential complementation described above. While de Villiers and colleagues have found a connection between sentential complements and FBT performance in several populations, in Mo et al. (2014) the question was addressed whether the conclusions about sentential complements and FBT performance in English can be generalized to include Mandarin which, as explained in chapter 3, shows different semantic and syntactic patterns for mental state verbs. The study replicated training that was similar to that performed in Hale and Tager-Flusberg (2003): Mandarin speaking children aged 3 to 4 were either trained with sentential complements or with false representations, or assigned to a control non-training group. The FB pre-test was in this case a location change task, whereas the FB post-tests were two location change and two unexpected content tasks. In the sentential complement training groups, children underwent two training sessions. In the first they were asked about the characteristics of an object using complement

structures (for example, "What do you think/say this is?"); in the second, they had to answer questions about what the child thought or said. The two groups differed in that for one communication verbs were used while for the other mental state verbs were used. In the false representation training, children were familiarized with speech bubbles that represented the thoughts of characters. Then they were presented with narratives including, for example, a girl seeing the picture of a horse and therefore having a picture representing a horse in her thought bubble. Once the girl turned, the picture was replaced, and children had to answer questions about the content of the thought bubble (which they could no longer see). Children who were trained with sentential complements with communication verbs showed greater improvements in the post-test FBT: this was explained by the experimenters by the fact that Mandarin-speaking children use sentential complements earlier and more frequently with communication verbs than with mental state verbs. Mo et al. (2014) argue that this might have been a more straightforward strategy for the children "to capitalize on whatever structural benefits acquiring the sentential complement construction has for reasoning about false beliefs." (Mo et al., 2014, p.58). False representation training also facilitated FBT performance, which was not surprising. Recall what was said for mental state verbs in Mandarin and Cantonese in chapter 3: the semantics of these verbs is not straightforwardly translatable to those of English, which means that the choice of verbs for the study might have influenced the performance. However, also notice that the verbs used in Mo et al. (2014) are not classified in Tardif and Wellman (2000), making the task of navigating the data about Chinese studies especially challenging.

Another set of evidence comes from the study of languages with different evidential systems, like Korean and Turkish, as was explained in chapter 3. In Papafragou and Li (2001), the relation between evidentials acquisition and Theory of Mind was investigated. In the first experiment, the ability of Korean children and English-speaking children to identify the source of their beliefs was tested. In the tasks, they had to identify how they got to know a certain piece of information (i.e. what was in a cabinet) or how a character got to know it. While a significant age difference was present for English children, with 3-year-olds performing consistently worse than 4-year-olds, the same was not true for Korean-speaking children, where the youngest group performed much better then their English peers (90% correct answers versus chance). However, these results have been put under doubt by the same authors, who, while not referring to their previous paper, do replicate the experiment and tasks (Papafragou et al., 2007), this time reporting the statistical analysis they use. In the experiments, the differences between Korean and English and the different timeline of acquisitions do not reflect different performances in the source mon-

itoring tasks. While the difference between age groups remained significant, Korean children and English children's performance in the source monitoring tasks was equivalent, which according to the authors suggests that language does not have an effect on the ability to classify different sources of evidence.

While Papafragou et al. (2007) does not directly address the issue of mentalizing, but focuses on the slightly different domain of source monitoring, the data by Mo et al. (2014) is interesting because it focuses on individuating effects similar to those found in English-speaking populations by de Villiers (2007). While the positive results are encouraging, caution should be exercised until new data can be provided and more solid conclusions can be drawn. As will be seen below, a lot of studies about deafness consider cross-linguistic data and different populations, in particular non-hearing children; since they are aimed at addressing a different issue, i.e. that of early exposure to language input. They will be reported below in the dedicated section.

Another set of evidence comes from cross-cultural studies, which are very interesting because they shed light, not necessarily on language in a strict sense, but on the influence of cultural practices. Naito and Koyama (2006) report that Japanese children seem to pass standard false belief tasks considerably later than English children, reporting that the tasks are passed only when children are around 6 years old. Interestingly, Japanese children's explanations of the behavior of the characters in the tasks tended to refer not to mentalistic factors, but to interpersonal rules ("they promised that" and "they said that"). According to the authors this is a matter of cultural differences; Japanese culture stresses the role of interpersonal situations more than individual intrinsic motivation.

Even more variation is to be expected in cultures where folk-psychology narratives are even more different, and this seems to be definitely the case in the Quechua community in Peru (Vinden, 1996). The language lacks mental state terms, and Vinden reports that, while tested children performed perfectly in line with expectations in appearance reality tasks, there was considerable difficulty in the false belief tasks, where children between 4 and 8 found it difficult to reply to questions which asked them to predict behavior based on mental states, including instances in which the aim was to identify where the character in the story would look for an object. Similarly, Vinden (1999) found that explanations of emotions in terms of beliefs are not to be considered a universal phenomenon: comparing Western children with Tainae, Tolai and Mofu children in Papua New Guinea showed that, while performance in false belief tasks seems to level up with time despite variations in the age for all the groups, there were evident difficulties in making sense of the questions directed to identify the *beliefs* of the characters, suggesting that the practice of using beliefs to explain emotions and behavior might simply not be as universal as it is sometimes assumed.

More recently, an incredible effort has been made to study Theory of Mind and mentalizing abilities with an ethnographic mind in Wassman et al. (2013), in which five different sets of studies are presented where children belonging to Micronesian cultures are tested with plausible adaptations of Theory of Mind tasks, finding that often false belief reasoning performance does not follow the patterns of development that have been registered in Western cultures. Moreover, anthropological data in the book suggests that explanations in terms of mental states is just not as common in the studied cultures, where interpersonal relations seem to play a more important role (analogously to the case of Japanese culture above), making the case as a consequence for care when it comes to claims about innate mechanisms and universal tendencies. The same is true for the results in Mayer and Truble (2012), which also confirm a strong cultural relativity when it comes to the classic ToM acquisition timeline, with children belonging to the Samoa culture passing the standard false belief task (change of location) much later (in some cases, not until 13 years of age) than Western children. This is connected by the authors to what has been referred to as the "doctrine of the opacity of other minds" (Robbins and Rumsey, 2008), i. e. the fact that in some cultures, including many in the Pacific, explanations involving mental state attributions are not considered central because of the difficulties that are inherent in accessing other people's minds.

The cross-cultural data is limited in quantity, and of course not decisive in many ways, given how many complex factors might be interacting. In this sense, more research is definitely needed. However, cross-cultural studies do call attention to the specificity of what is often taken for granted, i. e. Western folk-psychology narratives, and they are in this sense very valuable. Not only the linguistic expression of mental states, but also cultural and social norms can heavily influence the development of mentalizing abilities, and there is a sense in which some of these can be culturally specific, as they develop in adaptation to the demands of the culture the child lives in. In the rest of the book, the importance of narratives and their cultural relativity will be stressed again, with particular attention to narrative practices in the next chapter, and to data regarding narratives and storytelling in the next session.

4.3.6 Storytelling, narratives and play

A good starting point for analyzing the evolution and development of language in children is observing their spontaneous behavior; while many of the studies above create rather artificial settings and situations to test children's comprehension and production, some studies have investigated the use of mental state

terms in play, pretense, and story-telling situations and have related it to mentalizing skills. On the one hand, interaction with the caregiver is essential, and many studies focused on the frequency and use of mental state terms by caregivers and their impact on the developing child. On the other hand, children's own use can also be valuable evidence.

In Ruffman et al. (2003), as seen above, different mentalizing skills were tested. What was not mentioned before, however, is that caregivers' interactions were also considered. Children and caregivers were tested at three different points in time; caregivers[29] had to look at a series of pictures with their children, and their interaction was recorded and coded to identify cognition-related language. Children were then tested as in 4.2. Caregivers' use of mental state language correlated with later Theory of Mind skills shown by the children, after partialling-out the children's performance at time 1, and the caregivers' mental state utterance variable was predictive of the results in the mentalizing tasks. A strong correlation between the child's linguistic ability and their later mentalizing performance was also found, thus confirming the data from other studies. Interestingly, caregivers' mental state utterances accounted more consistently for variance in mentalizing skills compared to children's linguistic measures. Note that the experiment also confirmed another fact about English-speaking children's development, i. e. that desire-related expressions are produced and understood by children before think-related ones.

Symons et al. (2010) presents results related to mental state discourse in storytelling and during book reading as well. Canadian children and parents took part in the first experiment. Children had to perform in a series of FBTs, ranging from unexpected identity to content task and change of location, plus an emotion false belief measure. In this task, children witness a story: Pingu is tricked by his friend Mickey, who substitutes the contents of a can of Coke, which Pingu loves, with milk, which Pingu hates. Children had to guess what Pingu thinks is in the can and how he feels when he thinks he is about to drink his favorite drink. Children and parents also read a story together about the first day of school (this was particularly relevant for the children, since they were aged between 5 and 6), and the language used by the caregivers to

29 Like many other studies, this study only refers to "mothers" as the tested caregivers. The reason why I will constantly refer to the parents in a gender-neutral way is that, while it is not possible to change past studies, I do believe that the use of gender-neutral language will, in the long run, contribute to defeating stereotypes about parenting and child care. While this book cannot change experimental psychology's practices, since mothers are indeed possible caregivers, I will use the more neutral language throughout the book, unless forced to do otherwise, e. g. when gender differences are taken into consideration.

talk about the book was coded by individuating utterances about mental, behavioral and physical states; eventual elaboration of the content of the book was also coded. Results showed a significant correlation between mental state utterances about the characters in the story made by parents and children's performance in mentalizing tasks, and while children who had interactions focused on mentalizing language and explanations were more likely to perform better in mentalizing tasks, they were not more likely to be the ones *initiating* discourse about mental states. In other words, input from the parents seemed to be more relevant than whether or not the child was the one initiating the interaction and focusing on mental state utterances. In the following experiments, children's own production of mental state utterances was measured without caregivers, by having children reconstruct a story by looking at pictures; children's reference to mental and emotional states correlated with their mentalizing performance. Along the same lines, the ability to produce and use mental state-related vocabulary was found to be correlated with narration skills in Italian children in Gamannossi and Pinto (2014). Similarly, a link with storytelling was found in Fernandez (2011). In the study, Spanish children aged between 4.8 and 8.8 years performed in first- and second-order FBTs and engaged in storytelling activities, where their performance was coded according to how coherent and cohesive the stories were and how often children made reference to internal states. The results showed a covariance of children's skills in storytelling and their performance in the second-order FBT, thus giving some evidence of the relation between linguistic skills and sophisticated forms of mentalizing, linked by narrative practices.

A different approach was taken in Nielsen and Dissanayake (2010), where the relationship between mental state terms and mentalizing was explored in conjunction with pretense. The tasks used to assess false belief reasoning were of the kind unexpected content and location change. Children were also tested on their pantomime skills by being requested to mime certain objects, and were observed during play in a controlled environment for pretense play, coded in six different categories: objection substitution, imaginary play, attribution of animacy, role assignment, role play, and joint proposal. In this study, use of mental states was coded by distinguishing between referential use, conversational hedges and idiomatic use, thus keeping in mind the differences cited in the previous chapter about the different functions that mental state terms can have in conversation[30]. The referential use of mental state terms was associated both with false belief performance and some forms of pretense play, i. e. object substitution and role assignments, thus confirming a relation between verbal

30 See chapter 3, Shatz et al. (1983).

abilities and mental state verb production with FBTs, and the interrelation of those with pretense play skills.

Note that this set of studies brings attention to language's contribution to false belief reasoning and mentalizing, not as a format of representation, but as a conversational and social means of interaction. The fact that pretense play and interaction with parents accounted for better performance brings focus and attention to the social dimension of language more than the grammatical one. While the two things are not incompatible, it is sensible to keep in mind that they can be disentangled, as will be especially relevant when assessing the theories that try to account for the presented studies.

4.3.7 Clinical evidence

4.3.7.1 ASD

The relation between autism and mentalizing has been investigated extensively in the literature, since impairment in social cognition is considered one of the hallmarks of autistic spectrum disorder. A few considerations are in order: firstly, the reason why we talk about ASD and not autism in the first place is that there is a variety of conditions and symptoms comprised in the ASD, and classification of the phenomenon, along with a clear mapping of the various subcategories, is far from being achieved. This is of particular relevance for this book because it's not only social cognition symptoms that vary along the spectrum; linguistic impairment in ASD varies a lot too, thus making the condition an extremely interesting but extremely complex case.

Famously, Baron-Cohen et al. (1985) discovered a substantial impairment for autistic children in solving the standard location change FBT discussed in the previous part of the chapter. The tests were carried on using mental-aged-matched children affected by Down's syndrome who, despite reduced cognitive abilities, performed significantly better than those with ASD. The result has been replicated and is very solid (see the analysis of a large part of the data in Happ (1995)), at least as far as the location change and content change FBTs are concerned. This result is often discussed in the literature alongside another finding that was mentioned above, i. e. that the poorer performance in FBTs is not reflected in their performance in the false photography task (Leekam and Perner, 1991; Peterson and Siegal, 1998; Leslie and Thaiss, 1992); moreover, "easier" forms of mentalizing are also apparently not as challenging for ASD children, with average results in attributing simple emotions (Baron-Cohen, 1991). Crucially, the amount to which visual perspective taking is intact in ASD is debated. In an extensive review of the available studies, Pearson et al. (2013) and colleagues point

out that, while results differ, several difficulties arise. Firstly, it is possible to distinguish between Visual Perspective Taking level 1 (VPT1) and Visual Perspective Taking level 2 (VPT2). While VPT1 is intended to measure the ability to understand that other people have a different line of sight, VPT2 implies understanding that different spectators can have different perspectives on the same object, and hence requires one to take into consideration *how* a certain object appears to different viewers. While the difference seems to be clear cut, the authors point out that it is hard to disentangle exactly which cognitive requirements are necessary to perform the task, and if different strategies are available that could compensate for the lack of other abilities. Of the thirteen studies examined in their review, three assessed VPT2, and in two of these it was recorded that autistic children performed worse. For example, Hamilton et al. (2009) found that ASD children performed worse, not only on a task in which they had to guess what the point of view of a doll would have been on an object, but on a battery of FBTs, compared to typically developing children; however, they showed no such difficulty with a mental rotation task, suggesting the difficulties in VPT2 were not due to a difficulty in spatial cognition. Moreover, the performance in VPT was significantly predicted by the score in FBTs. In a similar fashion, Yirmiya et al. (1994) found that ASD children performed significantly worse than their age-matched peers when they had to rotate a turntable in order to make the experimenter's viewpoint match with their own, and vice versa. Tan and Harris (1991), on the other hand, found that ASD children performed like their peers when they had to answer questions like "Which object would John say was "in front"?", thus suggesting VPT2 was on the contrary intact. However, as Pearson et al. (2013) notices, there might be a change in strategy: this last task might be solvable with spatial clues and not "social clues" that require one to understand *how* another agent sees the object in question.

Visual perspective taking is not the only interesting variable in ASD, and verbal mental age seems to be predictive of performance in FBTs both for typically developing and ASD children, and the fact that ASD children often succeed in solving FBTs a few years later than their typically developing peers, has been used to suggest that the modularist picture of ToM is not correct (Garfield et al., 2001). As will be specified below, while this view is challenged by results pertaining to other clinical conditions like William's Syndrome, the apparent evolution of a strategy to deal with FBT requirements for ASD children is indeed a strong argument in favor of a module for ToM understood in the most traditional sense.

The partial and slow success of ASD children in FBTs is not only of interest in terms of modularity disputes, but also fundamental for the issue of whether language has anything to do with the social cognition impairments in ASD.

Among the various questions that the ASD literature raises is the question of *what* enables *some* ASD children and adults to pass FBTs and mentalizing tasks, since a considerable part of the population do not. Among the core diagnostic features of ASD, communication difficulties are fundamental (Vicker, 2009; Mody and Bellievau, 2013): adults and children often exhibit impairments in conversational discourse, restricted speech-act use, difficulties with narratives and pragmatic issues with non-literal meaning. In the analysis from Happ (1995) cited above, an age-independent relationship was found between scores in the British Picture Vocabulary scale and performance in FBTs for both typically developing and ASD children, suggesting that a relationship with vocabulary and language processing might be predictive of the failure in FBTs. The results have been replicated in terms of vocabulary scale in Tager-Flusberg and Sullivan (1994a); Dahlgren and Trillingsgaard (1996); Sparrewohn and Howie (1995), but in Tager-Flusberg and Sullivan (1994a) a stronger relation was found for syntactic measures as well. In the study, syntactic knowledge and competence were tested with the Clinical Evaluation of Language Fundamentals test for syntax, aimed at testing sentence structures and complex multi-clause sentences in particular; syntactic skills proved to be predictive of performance in FBTs. One of the aims of the series of experiments conducted by the Tager-Flusberg group was further exploration of the hypothetical relation between sentence complementation and FBT performance that was described above in the experiments conducted by de Villiers and colleagues. In a series of studies reported in Tager-Flusberg (2000), they investigated the relationship between complement understanding and ASD, finding that comprehension of mental state verb complementation was predictive of poor performance in FBTs, in ASD children. This was the case in a variety of tasks, including a task where children needed to report what a character said, replying to a *wh*-question, after hearing a story about it, which according to the authors suggests a relationship between difficulties reporting statements and performance in FBTs that is unique to autism. Interestingly, the group tested for referential opacity as well, modelling their study on the task in de Villiers and Fitneva (de Villiers and Fitneva) above, finding that children passing the sentential complement task were also the ones succeeding in the FBT. Crucially, there was a stronger relation between understanding and use of *communication verbs* compared to *mental state verbs* for ASD children, a pattern that was reversed in mentally challenged children tested in the study, which suggests according to Tager-Flusberg (2000) that ASD children exploit the complement structures used in communication verbs to acquire competence in solving FBTs, being less sensitive to cognition verbs.

Tager-Flusberg's idea is that ASD children exploit sentential complements more than typically developing ones, in order to form their folk-psychology the-

ories and structures. In Lind and Bowler (2009), ASD children, children with a general learning disability and typically developed children performed in a memory for complement task, a Sally Anne task and a smarties (unexpected content) task. Complement task performance correlated with the Sally Anne task more strongly for the ASD children group, and the correlation was only significant for the performance in the Sally Anne task but not for the smarties task. Note that, in this study, children were considerably older than in de Villiers' study, being between 8 years old and 10 years old. The results suggest, according to the authors, that children might make use of a strategy based on complement structures scaffolding in *some situations* during development, but possibly abandon these strategies later on.

Note that some studies found interesting results not only for the complement use of mental state verbs, but also specifically for the *evidentiality* involved in the use of mental state verbs. In Ziatas et al. (1998), performance of ASD children, typically developing children, children with Asperger's and children affected by SLI (specific language impairment) were measured for FBTs and other tasks testing comprehension of *know*, *think* and *guess*. In the *belief comprehension task*, children had to correctly interpret the clues given by puppets about the location of smarties candy, where the clues were given by two different puppets using the three verbs, with different evidentiality. In the production task, children had to control the puppets themselves and help the experimenter find the smarties. The results showed not only that ASD children performed considerably worse than the other groups, but their performance in the FBT (an explicit version of the location change task) was correlated. The evidence is not limited to syntax and evidentials either. Baron-Cohen et al. (1994) showed that ASD children produced fewer mental state verbs and terms referring to the cognitive domain when describing scenarios, thus providing further evidence of the importance of mental state term production in narratives; a poor performance in the recognition of mental state terms for ASD children was found in Baron-Cohen et al. (1994), and similar results have been presented in Tager-Flusberg (1995), studying narratives produced by ASD children in response to pictorial stories. In the case of Tager-Flusberg and Sullivan (1994b), the "poor" performance of ASD children in storytelling (measured in terms of length, number of mental state terms, ascription of emotions, connectives) correlated with their poor performance in mentalizing tasks.

In summary, the literature on ASD seems to bring together a variety of effects that were found in the previously reported studies: poor performance in FBTs by ASD children seems to be correlated with their syntactic, semantic, and narrative abilities, thus suggesting that a combination of these elements is at least partially responsible for the impairment, and that this impairment is also connected

with some difficulties arising with perspective shifting tasks that are less "verbal" than FBTs. As mentioned, ASD presents a large variety of symptoms, it is a complicated condition, and includes subjects that present a wide range of different impairments in various degrees. However, the fact that social impairments, linguistic impairments and mentalizing impairments seem to be converging in this condition is surely a relevant matter for the issue at hand.

4.3.7.2 Williams Syndrome

As will be seen in the next chapter, William's Syndrome (WS) presents a specific case that is of great value for specific theories of the interaction between language and mentalizing skill. The reason is that William's syndrome presents an interesting case against the idea that social cognitive abilities like false belief reasoning depend on general intelligence. William's syndrome is a gene disorder that seriously affects various cognitive skills, including decision making, logical skills, and spatio-mathematical skills. Despite these complex and diffuse deficits, however, William's syndrome children perform fairly well in FBTs. In a study by Karmiloff-Smith et al. (1995), children affected by WS performed in a variety of mentalizing tasks, starting from inferring the content of mental states based on direction of gaze, two first-order FBTs, two second-order mentalizing tasks, and metaphorical comprehension. Children performed close to perfect in all tasks, including the second-order task involving true belief, but excluding the more complex second-order FBT. Crucially, as noted by Garfield et al. (2001) William's syndrome children have fairly normally developing linguistic skills, with the exception of pragmatic and particularly abstract metaphor tracking, in a way that is similar, as will be seen, to schizophrenic patients. However, people affected by William's syndrome do not show the same more general impairments in language. What makes WS particularly interesting is that it presents a case that is, at least superficially, quite the opposite of some of the patterns seen in cases of ASD where general intelligence is well preserved but mentalizing skills are seriously affected. This has been used as evidence for a mindreading *module*[31]. The fact that linguistic skills are fairly preserved in WS supports the notion that language aquisition plays a role in making false belief reasoning possible, especially given the more general impairment in cognition.

31 Further elaboration will be given in next chapter when discussing Garfield et al. (2001).

4.3.7.3 SLI

Specific language impairment is a condition that sees linguistic abilities impaired despite normally developing cognitive abilities. The disorder was not included in the DSM-5 (the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders) because of specific reasons raised by the American Speech Language Hearing Association (ASHA, 2012), highlighting how the label, while largely used in empirical research, does not denote a clearly identifiable disorder (Reilly et al., 2014).

For these reasons, the condition should be handled with care; however, it is worth reporting results that address the performance in FBTs, since children that are considered part of SLI groups in these studies do present a language impairment.

The language impairments in SLI are also very heterogeneous, hence making it difficult to pinpoint whether the grammatical deficits are also always accompanied by lexical impairments in SLI, and vice versa.

It is firstly relevant to underline how SLI children were used in control groups in studies previously mentioned, like Ziatas et al. (1998). In these cases, while SLI children's performance was similar to typically developing control groups, their age was usually higher, thus making their better performance possibly the result of other components also and pointing to the necessity of considering SLI performance per se.

Miller (2010) investigated the relation between language and false belief reasoning in SLI. SLI children were compared with age-matched, typically developing children and language ability matched typically developing children and had to perform in a one sentence complement task. The FBTs were highly verbal and quite complex: in the "think" task, children had to answer the direct question "Where does the puppet think the toy is?". In the "show" condition, the child had to show the experimenter what the puppet would do, and in the "pretend" condition the child had to answer a complex question: "What does the puppet think we are pretending the toy is?". In the "less verbal" condition, children had to watch a video and indicate which one of the presented pictures was the natural continuation of the situation. The results showed that, controlling for chronological age, performance in the sentence complement task accounted for variance in performance in the FBTs, and children that passed FBT s were the same children who had more mastery of sentential complements for the *think* condition. In the think and show condition, SLI children performed similarly to the age-matched group, whereas all groups performed poorly in the pretense condition, which was probably too complex. In general, the results show that a relation with complementation was present, but also that SLI children were able to pass the FBT as well as their age-matched peers. Notice that we are talking in

this case of children around the age of 4, i.e. an age at which children already normally pass FBTs. Importantly, the fact that passing sentential complement tasks is a predictor for false belief reasoning performance was also confirmed in a study by de Villiers et al. (2003), reinforcing evidence for the role of syntax even in SLI children, albeit probably with some limitations.

Interestingly, possibly clearer evidence for SLI comes in Farrant et al. (2006), which reports unpublished studies like Tucker (2004), for example, in which a delay of 12 to 18 months was found for children with SLI when it came to developing the skills involved in mentalizing. Farrant et al. (2006) focuses on VPT2, investigating the relation between visual perspective taking abilities, false belief reasoning and language. Children had to perform in ten mentalizing tasks, which ranged from the easiest emotion attribution task to an unexpected content version of the FBT. Children also had to perform in a VPT1 task, where it was necessary to understand that another agent, Mr. Jones, sitting opposite the child, was seeing the opposite side of a card held vertically between them. In the VPT2 task, children had to correctly identify the perspective another agent had on an object (e.g. seeing a picture upside down). In an *array* VPT task, children had to infer how a series of objects would have looked from other spatial locations. SLI children performed significantly worse than the control group in the mentalizing tasks, and the same was true for the VPT tasks as well. More detailed analysis showed that the difference was due to the VPT2 tasks and the FBT, while SLI children performed similar to their peers in the VPT1 task and the emotion recognition mentalizing tasks. Interestingly, most children passing the VPT2 task passed the FBTs as well; the performance in the two tasks was indeed statistically correlated. Since the linguistic instructions of the VPT1 and VPT2 tasks are very similar in terms of complexity, the authors interpret the finding as showing a delay in the cognitive development of skills that underlie mentalizing, and not as underlining the verbal difficulties of the FBT. Another study by Farrant et al. (2012) seems to support these results: children aged between 46 and 76 months were tested in a series of tasks for memory for complement structures and cognitive flexibility, and the relation to their caregivers' language input was also investigated. Linguistic input from the caregiver predicted both performance in cognitive flexibility and false belief understanding.

Finally, more solid confirmation comes from studies like that in Spanoudis (2016), where children (average age 9) with specific language impairment performed in a battery of tests including a relative clause task and a time clause task (sentences including "before" and "after", which in Greek imply variations in the word order depending on use), syntactic measures, a semantic ability test, a lexical semantic task, and pragmatics tasks including understanding of implicatures and metaphors. In addition, children had to perform in a "ToM under-

standing task", where they had to answer six questions regarding a social situation described in a story, identifying which character was not behaving according to social norms and being rude or insensitive to somebody, depending on their knowledge. Children with SLI were compared to two control groups, i.e. age-matched children and language skill matched children. While SLI children did not perform differently to the language-skill matched children on the language tasks, they were outperformed in the mentalizing task. In this case, no strong correlation between pragmatic abilities and performance was found. Similar results were obtained by Farmer (2000), in which three out of four studies registered a significant delay in ToM performance for SLI children.

Data about SLI is not straightforward and still somewhat contradictory, and the fact that the condition itself is not easy to assess nor identify surely adds to the complexity of the analysis. However, the fact that a delay is generally registered for children with language impairment is compatible with a role for language in mentalizing; this, unless it is assumed that mentalizing is itself linguistic and/or that linguistic skills are responsible for the development of mentalizing skills as a whole, and not a contributing factor interacting with other skills as well. The fact that SLI children eventually come to master FBTs might be a good reason to assume that, while language has an impact on mentalizing abilities, other mechanisms can at least compensate for its lack. As mentioned, however, the fact that SLI is not a straightforwardly identifiable condition according to the latest research should always be kept in mind.

4.3.7.4 Schizophrenia

As mentioned, interesting results related to pragmatic development come from schizophrenia. Schizophrenia is a chronic condition involving a wide range of symptoms, including delusions, hallucinations, and disorganized speech, impacting executive function, memory, and language. Subjects with schizophrenia have impairments in mentalizing skills that are well-documented[32]. for a review of the evidence in this respect.) The problems which schizophrenia presents for mentalizing seem to be reliable, with the meta-analysis in Sprong et al. (2007) showing that variables like IQ, gender, and age do not affect the size of the effect when considering failures in mentalizing skills for schizophrenic patients. Additionally, schizophrenic patients seem to fail at both verbal and non-verbal mentalizing tasks. It should however be kept in mind that most of the mindreading tasks analysed were explicit: the sample included explicit first- and second-order

32 See Frith (2004); Brne (2005); Harrington et al. (2005); Sprong et al. (2007).

FBTs, intention inferring tasks in which the participants had to guess the intention of a character from a short story and understanding of irony.

The impairment in schizophrenia seems to include a large variety of non-literal speech phenomena, including proverbs (Kiang et al., 2007; Rapp et al., 2014), idiomatic expressions (Schettino et al., 2010), and metaphors (Mossaheb et al., 2014). The deficit seems then to be related mostly to language comprehension and production skills at the pragmatic level (Mazza et al., 2008), and might therefore be related to more general social-cognitive impairment. However, some data regarding semantics and syntax are also available. In a study by Tavano et al. (2008), Italian schizophrenic patients' skills in pragmatic tasks and syntactic tasks were analyzed, showing that lower syntactic variety in speech production is present, along with deficits at the receptive level for syntax: as in other studies that have been presented in this chapter, patients had to match sentences like "The plate is on the table" or "The girl pushes the boy who is kicking the ball" with the correct picture. (Distinguishing between the picture in which the girl pushes the boy from the one in which the boy pushes the girl is in this case considered a matter of mastering syntactic skills, since it involves understanding how word order structures the sentence's components.) Access to syntactic structures seems then to be a core impairment in schizophrenia, which is in line with other results in the literature, such as Covington et al. (2005), who mostly reports a simplification in syntax for schizophrenic patients, but also with Kircher et al. (2005), where anomalies in the neural correlates of syntax production in schizophrenia are reported, with a lack of activation in the right posterior temporal cortex and the left superior frontal cortex. At the same time, Rodriguez-Ferrera et al. (2001) found semantic impairments in schizophrenic patients, suggesting an even more complex picture of interrelating linguistic and non-linguistic difficulties in schizophrenia.

Schizophrenia is an incredibly complex condition that cannot be reduced to a linguistic or pragmatic deficit. However, the relation between linguistic and mentalizing deficits in schizophrenic patients, similarly to the data on autism, seems to suggest that these abilities are interlocked and interrelated to at least some extent, bringing results to the table for semantics, pragmatics, and syntax.

4.3.7.5 Aphasia

The study of aphasia can be particularly interesting for the topic at hand because of the language impairments that this condition brings with it. Aphasia implies substantial grammatical impairments; however, even in this case, the symptoms are complex and not always homogeneous, with different kinds of aphasia of several degrees of severity being involved. Moreover, the deficit in language

might sometimes be the result of domain-general networks. Once again, then, it is hard to interpret the deficit clearly.

Studies investigating the relationship between aphasia and false belief reasoning bear interesting and yet contrasting results. Varley and Siegal (2000) studied a patient with severe damage to left hemisphere language centers. The patient was able to perform regularly in a large range of neuropsychological tests about abstract thinking and causal reasoning, and was able to carry out complex activities like planning finances and play chess. However, despite retaining a large vocabulary, he had severe problems understanding verbs. Yet, his false belief reasoning skills seemed to be left intact as they were assessed by a battery of false belief tasks. This has been used by Varley and Siegal (2000) and Siegal et al. (2001) to argue that grammatical information is not necessary for false belief reasoning.

While the results of this investigations are definitely interesting, several things have to be considered. Firstly, the patient is a single case, which is why replication of the results is obviously challenging. Secondly, while this case is a valid piece of evidence against a conception of language as *constitutive* of false belief reasoning, it does not completely rule out the possibility that language has some role in the *development* of false belief reasoning skills, contrary to what is argued in Varley and Siegal (2000) and Siegal et al. (2001). Finally, it has been argued by Baldo et al. (2005), for example, that the impairments in this cases were not as severe and "deep" as they might have appeared, since the patient was, after all, able to understand the verbal instructions in the false belief tests. On the other hand, Siegal et al. (2001) argue that training with mental state verbs was provided before the tasks to allow the patient to follow: this of course might actually be used by the opponents to claim that some verbal priming was after all necessary to help the performance, but at this point the issue becomes speculative, for there is no way to know what the strategy was for the patient to solve the tasks.

An attempt to disentangle these factors is made in Apperly et al. (2006), in which another patient with aphasia was tested on a battery of linguistic tasks, dedicated to confirming his impairments with respect to grammatical processing, and non-verbal false belief tasks of first- (location) and second-order (in which a belief about a belief had to be attributed). Moreover, the patient had to perform in a "ToM semantics" task in which he had to be able to understand the different degrees of certainty expressed in mental state terms by a clue-giver in order to identify the right box. Interestingly, while performance was poor on language tasks, the scores were considerably above chance for the non-verbal ToM tasks and the "ToM semantics" task as well, suggesting, according to the authors, that while sentential complement constructions were not involved in false belief

understanding, semantics of mental state verbs might indeed be a relevant factor.

However, I believe two conclusions can be drawn. First, it is good to keep data like this in mind when accounts of false belief reasoning are given which predict that false belief reasoning is carried out in a linguistic format. The fact that adult patients with severe language deficit seem to be able to carry out reasoning about false belief seems to suggest that false belief reasoning is not to be considered a linguistic operation. Second, lesion studies should be treated with care in that they can provide unique insight as well as suggestive but not conclusive results.

4.3.8 Deafness

Research on deafness and false belief reasoning naturally focuses on the role of language input and the forms that it takes. This is because it is generally possible to effectively distinguish between deaf children with hearing parents, who get exposed to sophisticated sign language in school, and deaf children with signing parents, who acquire language roughly at the same rate as hearing children.

Firstly, it is relevant to stress that late-signing children are documented to have false belief reasoning skills that are indeed delayed compared to those of their peers, whether early-signers or hearing children. In Russell et al. (1998), late-signing children's performance was compared to their peers in a variety of FBTs; the pool included children between 4 and 16 years old, and the performance seems to be correlated with age, showing that while children eventually reached skill mastery, they did so with considerable delay compared to the typical developmental pathway. In a similar study, Courtin and Melot (1998) found that late-signing children were delayed compared to early-signing children and hearing children in three different explicit verbal FBTs. Results are not limited to verbal FBTs only: Peterson and Siegal (1998) compared performance of autistic children, typically developing preschoolers, and late-signers who had to perform in false photography tasks as well as FBTs. The results showed the same pattern for ASD children and late-signers, with better performance in the false photography task, but impaired performance in the FBT. This was also true in a study in Peterson and Siegal (1999) that required a non-verbal response mode, thus making the results significant in addressing the question of whether the verbal prediction involved in standard FBTs might be the problematic feature for late-signing children. Poor results in non-verbal false belief tasks are also reported in Figuras-Costa and Harris (2001).

In an extensive review, Peterson and Siegal (2002) summarize results from eleven different studies on deaf children and ASD children, where different populations have been tested: Australian (Peterson and Siegal, 1998, 1999, 2002), American (de Villiers et al., 1997), Scottish (Russell et al., 1998), British (Steeds et al., 1997), and French (Deleau, 1996; Courtin, 2000). While most of these populations are English-speaking, the fact that the same delays were reported in these studies for non-hearing children is surely encouraging in pointing to a solid result in this sense.

In line with the studies they conducted on typically developing children (de Villiers and Pyers, 2002), Schick et al. (2007); de Villiers and de Villiers (2011) found that passing sentential complement tasks like the ones described above was a strong predictor of FBT performance for deaf children as well. de Villiers and de Villiers (2011), in particular, found that both syntax production and vocabulary comprehension were significant predictors of performance for both verbal and non-verbal FBTs. In Schick et al. (2007) study, American Sign Language signing children and oral signing deaf children's performance was compared. In the *low verbal* false belief task, children had to find stickers hidden in different boxes following the clues given by two different experimenters, one of which could see where the stickers were hidden, while the other could not. In the *(un)surprised face* task, children had to attach a sticker with a surprised or non-surprised face to a character in a story, thus predicting the state of mind of agents who had or lacked reasons to expect the change o location of an object. The children were tested with vocabulary measures, general English syntax, and comprehension of false complement clauses. The results showed that deaf children from hearing families were consistently delayed in performance for both verbal and low verbal tasks. Deaf children with signing parents, on the other hand, performed like hearing children. The authors point out that, while oral deaf children's performance in de Villiers et al. (1997) was predicted by language measures, and especially cognitive state verbs and complementation, their performance in attribution of simple emotional states was predicted by non-verbal IQ and age, thus suggesting that false belief reasoning does indeed involve different sets of cognitive skills than other forms of mentalizing.

de Villiers and de Villiers (2011) investigated the relation between deception, mentalizing skills and deafness, comparing oral deaf children with hearing children. Children performed in a location change FBT, a speech-bubble similar to that described above, and another (un)surprised face task. In the deception tasks, children had to either trick a puppet into choosing the hand that did not contain the stickers he was looking for, or trick a wolf trying to hide Mickey Mouse's cupcake. Language skills were measured with a standard syntax test, a memory for complements test and a vocabulary test. In an effort to explore the

relationship with Executive Function (EF), children also performed in a battery of tests testing executive function abilities: in a *Day-Night* stroop task, for example, they had to say "night" every time they saw a picture of the sun and "day" every time they saw stars in a night sky. Deaf children, as predicted and as confirming the results reviewed so far, had a lower score in both verbal and low verbal FBTs, whereas no group effect was found for the deception tasks and the EF tasks.[33] Importantly, for deaf children a correlation between FBT performance and vocabulary and syntax skills was found, but no relationship between EC and FBT performance; on the other hand, EC scores correlated with deception scores.

It is noteworthy to underline that the interrelation between false belief reasoning and language has been reported with Nicaraguan Sign Language as well. NSL presents an incredibly interesting opportunity of study for language development because of its recent emergence: the sign language appeared in the 1970s when deaf children were introduced in special schools, and was expanded by a second cohort of children that now possesses a much larger and more sophisticated form of the language, making possible a clear distinction between the first and second cohort of speakers (Coppola and Senghas, 2001; Senghas et al., 2004; Polich, 2005). In Pyers and Senghas (2009), NSL speakers were tested in low verbal FBTs and the frequency of mental state verb production was tested at two different times, three years apart. While NSL speakers from the first cohort barely produced mental state verbs in 2001, they did produce desire verbs; the second cohort's cognitive vocabulary was, in contrast, richer. This pattern changed in 2003, when not only did the production of mental state terms increase for the first cohort, but their performance in false belief reasoning also significantly improved, narrowing the difference between the two groups by several points. This was taken by the authors as good evidence of the fact that mental state-related language was indeed significantly helpful when it came to false belief performance, even though the task was non-verbal.

While most of the cited studies focus on deaf children who are older than 3 years, Meristo et al. (2012) investigated the performance of deaf infants aged 17 to 26 months in looking time paradigm tasks, which, as was explained in the first part of the chapter, is often used to assess implicit forms of mindreading in young children. In this study, a variation of the task used in Surian and Geraci

33 With the exception of a lower score for older deaf children; this might be due, according to the authors, to the fact that deaf children tend to be placed in traditional schools after a certain age, while children with more evident cognitive difficulties stay in specialized schools like the one these children were recruited from. The fact that no age group difference was found for the delayed performance in FBTs is encouraging in pointing to the fact that this did not impact the overall results of the study.

(2012) was used. In the true condition, the child witnessed Jerry hiding from Tom after exiting a tube and changing his location at a moment when Tom could clearly see the movement. In the false belief condition, Tom was absent when Jerry was changing his hiding spot. While performance was equivalent for the two groups in the true belief condition, with the children's anticipatory looks showing they expected Tom to look in the right box (assuming that this measure is indeed effective in explaining belief attribution, see paragraph 4.2.4), the same was not true for the false belief condition. While hearing infants looked preferentially to the right box (i.e. the box where Jerry was not hiding), the same was not true for any of the deaf children. This was considered strong evidence that language, especially rich communication with parents, is of fundamental importance for the development of false belief reasoning even in its most implicit form. These results were confirmed by a later study (Meristo et al., 2016) in which older children were tested with a eye-tracking task for false belief understanding. Spontaneous predictions were once again delayed and impaired for non-hearing children, who in this study were 4- to 8-year-olds. This is in line with other studies that explore the relationship between maternal input and false belief performance in deaf children, like Moeller and Schick (2006), where a correlation was found between cognitive terms used in caregivers' speech and FBT performance. Hearing caregivers of signing children were found to make use of significantly less cognition-related vocabulary, and this was found to correlate with their signing abilities, since more skilled signers produced more mental state vocabulary. Deaf children's false belief understanding was also correlated with the use of mental state vocabulary, thus once again revealing a fundamental effect of language use and interaction for the facilitation of false belief understanding. The same is true for the results reported in Woolfe et al. (2002), where syntax ability, mental age, and executive functioning were controlled when assessing the difference in mentalizing abilities between late-signers and younger early-signers in a simple FBT in which noverbal response was necessary, but children could place in the speech-bubble the item that they thought represented the character's mental state. Early-signers, despite being younger, outperformed late-signers in this task. Interestingly, the BSL (British Sign Language) test used to assess the linguistic skills of the children revealed that syntactic skills alone did not explain the difference in performance; however, standardized scores on a measure of syntax and morphology together did associate with FBT performance. Moreover, false photography task performance in the second experiment did not vary between early and late-signers.

Data about deafness is extremely revealing as far as the role of language for mentalizing tasks is concerned. Not only does early access to language seem to be predictive of FBT performance, but this relation seems to be present early on,

as reflected in the studies that tap implicit false belief understanding, and to have long-lasting consequences, as in the studies that investigate FBT performance in children who are significantly older than typically developing children who pass FBTs for the first time. As many of the researchers involved in the presented studies also underline, it is of fundamental importance to keep in mind that late-signers are not only deprived of rich linguistic stimuli, but often of richer interactions with siblings and hearing parents as a consequence. This once again reveals how the interaction of social and linguistic input probably underlies this differences in performance. At the same time, early access to language seems, somewhat surprisingly, to be related to very implicit false belief task performance in deaf children. While this might be due to task design, that language exposure has a role in mentalizing is largely supported by this data.

4.4 Conclusions

This chapter reported a wide variety of data from different research areas, with the aim of showing how empirical evidence points in the direction of language having a role in the development of mentalizing skills. Firstly, I delineated the issue of mentalizing in social cognition, defining what mentalizing is and mentioning how some issues in the debate make its own nature a theoretical matter. After marking a distinction between Theory Theory and Simulation Theory, I presented the basic literature regarding the most classic task assessing mentalizing skills, i.e. the false belief task, showing how the developmental trajectory of mentalizing skills is not only a matter of further dispute but also entrenched within issues regarding several cognitive skills, like perspective taking and gaze following. Subsequently, I presented data regarding the comparison between the false belief task and other false representation tasks, arguing that, while comparing performance in these tasks adds some relevant evidence to the debate, conclusions are hard to generalize, given that social skills and development are likely to be intertwined with the meta-representational skills involved in the FBT.

Data about the role of language in mentalizing tasks occupied the rest of this chapter. As I also argued in Berio (2020a), evidence is particularly encouraging as far as general linguistic abilities and semantics are concerned, since studies reliably show a correlation between linguistic development and performance in various mentalizing tasks. The evidence is also compatible with a role for syntax that, as will be seen, is central for some theories of false belief reasoning, given that training studies and longitudinal studies show how frequently sentential complement abilities and false belief reasoning are interrelated. At the same

Table 4.1: Populations and performance, summary.

	Unaffected	Affected
ASD		Perspective taking 2, Explicit FBT, Language
WS	Language, Explicit FBT, Implicit FBT	Second-order
SLI	FBT (partially)	FBT (partially), Language
Schizophrenia		First order and second-order FBT, Language
Deafness		Implicit, low verbal and verbal FBT, Language
Aphasia	low verbal and verbal FBT	Language

time, results suggest that it is somewhat hard to disentangle the various linguistic components that are likely to have an influence on mentalizing skills. This is also what research into pragmatic skills suggests, since both false belief reasoning and socially oriented linguistic skills are involved in interaction and dialogue, suggesting that syntactic and vocabulary measures are possibly not the only variables at play as far as mentalizing skills are concerned. The same is true for bilingualism, where contrasting evidence seems to suggest that other skills might be closely interacting with language acquisition. Cross-linguistic studies suggest that training in sentential complements has positive effects on false belief reasoning both in typically developing and atypically developing populations. Results concerning storytelling and interaction with caregivers, on the one hand, confirm that mental state vocabulary has an effect on social cognition skills; on the other hand, they also stress how interaction with caregivers is of fundamental importance, and how false belief reasoning is interrelated with pretense and play as well. Clinical evidence also brings new reasons not to underestimate the role of language acquisition in the development of mentalizing skills to the table: ASD children's performance is influenced by their semantic, syntactic, and pragmatic skills, and this is extremely relevant given that the ASD population is often at the center of the mindreading debate. William's syndrome data suggests that general intelligence is not solely responsible for false belief reasoning, while evidence about SLI is more difficult to interpret, as is evidence about aphasia. In both cases, considerations about the complexity of the condition and the limited data available recommend treating the issue with care, but also point in the direction of a non-constitutive view of language with respect to mentalizing skills. Finally, data on deafness is of considerable interest,

since it strongly suggests that early exposure to complex linguistic input and mental-related dialogue is crucial for the typical development of mentalizing skills. Looking at the table in 4.4, it is clear that conditions that imply a delay in language acquisition are often accompanied by difficulties solving FBTs: obviously, the table strongly simplifies the matter, as it does not take into consideration the variables presented during this chapter, like age, executive function, and whether the data is correlational or comes from training studies. Given that all of these variables have been explained in the previous sections, however, the table provides a quick overview that makes it clear that, if conditions like aphasia where language is only affected at a late stage seem to rule out a linguistic solving of FBTs, language difficulties in development might indeed be intertwined with difficulties in some basic mentalizing.

In summary, the results of this empirical review strongly suggest that language has more than one relevant role in mentalizing (Berio, 2020a). While this role does not seem to amount to any sort of constitutive relation, the impact of vocabulary acquisition, syntactic skills, and semantic knowledge on mentalizing skills seems to be confirmed by data on a variety of populations and by studies of different kinds. In the next chapter, these studies will be used as the background against which theories addressing the relation between language and mindreading are to be evaluated.

Part III: **Bridging the gap: theories about the interaction between language and mentalizing**

Chapter 5:
Psychological and philosophical theories of the influence of language on mentalizing

5.1 Introduction

This chapter will analyze theories of social cognition in a way that is substantially different from what is normally done in the literature, since it will focus on those theories that explicitly address the role of language in mindreading. While I have analyzed some of these theories in Berio (2020a), I will go here in much more detail. It is customary to treat the debate about mentalizing by marking a strong distinction between Theory Theory and Simulation Theory[34], as explained in the previous chapter. This categorization often goes along with listing *Modularity* Theories as an option (for instance in Goldman (2006)). While all of these distinctions make good theoretical sense, they would obscure this enterprise's main question, i.e. the question of the role of language in the development of mentalizing skills. The focus of this chapter is on theories that either account for the data presented in chapter 4, or that address the question of how to make room for such data. This will produce an overview that naturally gives more space to some theories that are overlooked in the literature, and that gives less relevance to often-discussed theories that are *language free*, in that they implicitly or explicitly exclude language. The exclusion of some of these theories comes naturally given the empirical review in the previous chapter; as argued, evidence that language acquisition has some role in mentalizing is solid enough to make the case for excluding the theories of mentalizing that sees it as a module-like, encapsulated, non-acquired ability. It is in this spirit that strong modularist theories like Leslie (1994b, 2000); Leslie et al. (2004) are excluded from this chapter. Leslie, simply put, considers language as just a manifestation of the ability to conceptualize mental states.

However, there will be other "big absents" in this review, one notable example being Nichols and Stich's Theory of Mind proposal (a good example of which is developed in Nichols and Stich (2003)). The reason for this is twofold. Firstly, the debate around Nichols and Stich's proposal mostly revolves around the difference between first-person and third-person mindreading, which, while certainly essential for social cognition, constitutes a separate issue that cannot be dealt with at length here. In a sense, this applies equally to the view developed by

34 See Hutto (2008b), which harshly criticizes both approaches.

https://doi.org/10.1515/9783110748475-006

Goldman (Goldman, 2006), which is in open opposition to Nichols and Stich in terms of how the access to one's own mental states is regulated, and how "privileged" this access is. Notice that both approaches are widely recognized as *hybrid* approaches that lie at the intersection of ST and TT. Secondly, Nichols and Stich's theory has very little to say about the acquisition of language and its relation to mentalizing, and therefore can be classified as a theory that simply does not address the issues presented here. In particular, Nichols and Stich consider language to be valuable evidence for getting to know about other's mental states in social situations: in this sense, it serves as informational input and gives more insight into other people's mental states. While this is certainly a fundamental part of the story, it is once again not exactly the role of language investigated in this book. The evidence presented in the previous chapter has made clear that it is necessary to have an account of how the development of linguistic abilities interacts with the nascent mentalizing skills present in infants. It is not to be excluded that some traditional approaches to mentalizing, for example Nichols and Stich's approach, can be further developed to include such an account. However, the fact that the question about the role of language is left aside is enough to motivate partially discarding these approaches.[35]

Another theory that is definitely worth mentioning despite the fact it will not be treated at length here is that pushed by Alison Gopnik (Gopnik and Goldman, 1993; Gopnik et al., 1999), which is regarded as possibly the most prominent Theory-Theory account available in the literature. Gopnik's pioneering work has been vital in shaping the controversy between ST and TT, and a large amount of empirical work has been carried out by her and her colleagues in determining the emergence of social cognitive skills in infants. Moreover, an essential part of her account is that the conceptual abilities implicated in mentalizing do evolve and change over time, which is an assumption that is in line with this book and with the approach that is encouraged in this work in general.

This makes Gopnik's version of TT fundamentally more compatible with the scope of this book, which adopts a non-nativist approach, compared to accounts like Leslie's. As should be clear from the brief treatment of her theory of conceptual acquisition in the chapter dedicated to language and thought, language plays a marginal role in her theory, but it is still counted as relevant evidence. The main assumption underneath Gopnik's Theory of Mind account is that children learn by formulating hypotheses, testing their predictions and consequently

35 I include in this sense Goldman (2006) among the theories that do not openly engage with the linguistic literature and are therefore not analyzed at length in the chapter. I present Gordon's view (Gordon, 1986) instead because it is on the simulationist spectrum but attributes a more central role to language.

modifying their concepts on the basis of the evidence they possess. This approach has been complemented in more recent work by a Bayesian take on learning, which has been integrated in Gopnik's view (Weisberg and Gopnik, 2013; Bonawitz et al., 2014; Gopnik et al., 2015; Gopnik and Bonawitz, 2015). In her framework, children rely on all sorts of evidence to advance their Theory of Mind, which starts from a "like-me" assumption: in other words, children tend to read other people's behavior in light of their own (Weisberg and Gopnik, 2013) until, with age and experience, they gather enough evidence to build more sophisticated theories about behavior and psychology. Notably, she argues that some of this evidence is heavily cultural and linguistic: for example, the fact that producing explanations of behavior in terms of individual psychological traits is a widely diffused practice in Western society seems to have direct consequences for the kind of explanations given by 6-year-olds (Gopnik and Seiver, 2009), compared to 4-year-olds. This is evidence, it is argued, that linguistic and cultural information is going to be of fundamental importance for what kind of theory gets formed about other agents' behavior, and how this interacts with previously held assumptions about how behavior can be explained. This general approach is perfectly in line with the kind of enterprise of this book but it does not explicitly engage with the impact of specifically linguistic skills, nor incorporate an account of how to address the data about syntactic acquisition, the development of pragmatic skills, vocabulary expansion and so on, presented in the previous chapter.

This is the reason why no extensive treatment of the theory will be given in this chapter.[36]

Another relevant factor that will be considered in this review is how these theories relate to the fundamental questions regarding cognition architecture presented in chapter 2. One of the purposes of this work, as a matter of fact, is

36 Note that another theory that I do not focus on here is that developed by Carruthers, for several reasons. Firstly, Carruther's account of Theory of Mind changes slightly between Carruthers (2002) and Carruthers (2016). Secondly, as has been seen in Chapter 2, Carruthers' view on language and cognition is not one that I consider in line with most of the empirical data on the relation between language and thought, or the most theoretically solid, and his view on Theory of Mind is tied to his more general theory and comes with heavy modular assumptions. Theory of Mind is a module, for Carruthers, and language is mostly what is needed to *express* the content produced by the module. In this sense, *natural language* has a secondary role. In Carruthers (2016), he argues that language contributes to solving false belief tasks in the sense that language abilities are required to interact with the ToM module to produce the adequate output, and as said this is not the kind of relation between mentalizing and language I am after in this work.

that of individuating how a theory of mentalizing skills can be inserted in the more general picture of how language interacts with non-linguistic processes. As a consequence, this also constitutes the background against which the theories mentioned below will be evaluated. One last note is essential; in the next chapter, while many of these theories will be further mentioned and discussed, I will also introduce two theories that, while they do not directly engage in the discussion about the role of language, leave room for integrating an account of the influence of language in mentalizing. These are the double system theory proposed in Apperly and Butterfill (2009) and developed in Apperly (2011), and the double system account in De Bruin and Newen (2012) and De Bruin and Newen (2014). Contrary to the theories cited above, there is clear room in these two frameworks for integrating an account of the role of language; the reasons why this is true will become clear after other theoretical proposals have been analyzed and my own account has been presented in chapter 6.

5.2 De Villiers and the bootstrapping hypothesis

The first theory that will be presented originated in psychology and gives a particular role to syntactic information. The fundamental idea of theories like that of de Villiers is that syntactic structure acquisition can provide a new representational means for attributing mental states to another agent. The idea has been developed in de Villiers (2004, 2005); de Villiers and de Villiers (2009); de Villiers et al. (1997); de Villiers and de Villiers (2012); in a nutshell, syntactic structures are seen in the theory as linking the semantic property *point of view*, which is marked in language, to the actual attribution of a point of view (from now on, POV) related to a determined representation. The theory relies on the specific features of mental state verbs that have been explained in chapter 3; as a reminder, let us keep in mind that these verbs have the peculiarity of accepting complement clauses, accepting complement clauses, and so too the embedding of another sentence that constitutes the argument of the verb. Consider the example below, where the complement of "knows" is a sentence with a finite verb:

(1) Galileo knows that mum is in the kitchen.

Mental state verbs can be further divided into different categories. For the current purpose, it is only important to focus on a specific feature of mental state verbs, namely the fact that they can be *non-factive*, i.e. they can have as sentential complement a false sentence, without it impacting the truth value of the main clause, like in (2):

(2) Vega thinks there is a unicorn in the garden.

While the choice of verb in (1) suggests that, if mum is not in the kitchen, (1) is false, the same is not true for (2), if there is no unicorn in the garden.

As explained, the acquisition curve is highly important when talking about mental state verbs: while the first uses of mental state verbs emerge around age 2 and a half (Limber, 1973), they are usually used as conversational edges (Shatz et al., 1983) or to express certainty or uncertainty (Diessel and Tomasello, 2001), and ultimately only used to express mental states around the age of 4. This is also the age at which children learn to distinguish between *know* and *think*. The fundamental intuition beyond the syntactic bootstrapping hypothesis, then, is that the acquisition of complement structures like those in (1) and (2) and false belief performance is causally connected. According to the hypothesis, once children learn to master these structures, they also acquire a new representational format for attributing mental states to an external actor whose behavior they interpret.

As should be clear, de Villiers' hypothesis is supported by the evidence regarding sentential complement acquisition and false belief task performance presented in the previous chapter. As seen, while performance in FBTs has been shown to correlate with the development of language abilities in meta-analysis (Milligan et al., 2007), more specific data is also available: training with use of mental state verbs and with syntactic structures in which they typically appear has been shown to improve the performance of children in FBTs (Lohmann and Tomasello, 2003); passing "sentential complement tasks" has been proven to be a predictor of false belief performance in both typically developing children (de Villiers and Pyers, 2002) and deaf children (de Villiers and de Villiers, 2011). In this sense, Tomasello and Rakoczy (2003) make a claim similar to the bootstrapping hypothesis, arguing that the linguistic expression of propositional attitudes can be the leading ability that brings about the ascription of propositional attitudes to the child and to others.

Fundamentally, the syntactic structure is thought to have a pivotal role in the formation of a new representational structure that allows for the attribution of PoV-*y* over the event/situation *X*. This is possible because a link is established between the view on a situation and the situation itself: in other words, a representation is involved, which is not attributed to the speaker. Arguing that language provides a PoV, then, means for de Villiers attributing to the syntactic structure the means for having a representation "*Subject P oV* [event/situation]". In the case (2) above, that means attributing to *Vega*, and not the speakers, a *PoV* over the "state of affairs" *that a unicorn is in the garden*. The influence of the grammatical structure is supposed to be not only limited to cases of explicit at-

tribution of mental states, but also to the very specific cases in which an explicit decision has to be made regarding their attribution. In other words, we are only dealing here with a very specific kind of ToM practice, and the claim should not be generalized.

Her view is often named "syntax bootstrapping": the idea is that linguistic structures that typically embed mental verbs give children the ability "*to reason about the content of others minds, a new format, if you like, for thinking about these abstract events*" (de Villiers and de Villiers, 2014); in this sense, acquiring mental verbs gives children "a point of view".

An essential remark that needs to be made before further explaining the account regards the scope: de Villiers and colleagues assume that their theory works for false belief tasks that entail a decision:

> We proceed under the assumption that for false belief tasks that entail a *decision*, which may or may not be the only tasks that require representation, the evidence is strong for linguistic complementation being a powerful predictor. (de Villiers et al., 2014, p.226)

What de Villiers means by "decision" is, simply put, the fact that a behavioral choice is based on the evaluation: the claim is then that when children have to act on the information they are gathering by observing adults, language becomes essential. Now, this raises, from a philosophical point of view, several questions, related to the kind of representation that is entailed here. Before proceeding with a more philosophical analysis of the argument, however, let us complete the analysis of the theory.

Rakoczy (2003) argues contra syntax bootstrapping views, showing in his results that training with false complements in general does not enhance false belief understanding. De Villiers' view (de Villiers et al., 2003) offers an interesting argument in favor of discarding syntactic views on this evidence: her theory is not a theory of syntax in general as scaffolding false belief conceptual understanding and Theory of Mind, but is rather a theory that gives a very specific role to a particular kind of syntactic construction, namely the non-factive mental verbs that take *realis* and *irrealis* object; in this sense, since her analysis clearly differentiates between *want* and *pretend* versus *believe* and *know*, it actually predicts Rakoczy's result.

In de Villiers' theory, *say* and *think* basically assign a PoV to the complement they take. See de Villiers' example:

(3) Giorgio thinks [a unicorn is dancing in the garden].

where the idea is that the complement [a unicorn is dancing in the garden] is assigned a PoV, namely Giorgio's. Note that the main clause is actually assigned

a PoV as well, namely the speaker's. Recall how memory for complements was one of the criteria for testing whether or not children were familiar enough with the sentential complement constructions in the studies cited in 4.

De Villiers' starting point is an analogy with other linguistic features that seem to have patterns of acquisition similar to those of mental verbs; this is possible because, as said, she focuses on syntactic structures more than lexical entries. In one of her studies, she tested children's comprehension of stories like the following:

(4) The Mom said she bought apples, but look, she really bought oranges. What did the Mom say she bought?

3-year-old children replied, incorrectly, "Oranges", but 4-year-old children replied, correctly, "Apples". Now, this kind of test is designed with the specific aim of testing mastery of embedded sentences (and, in this case, of wh-questions), and not mental verbs. What is interesting to notice, however, is that according to a study by de Villiers and Pyers (2002), standard false belief tasks are passed after requiring the complement structures with communication verbs, and that comprehension of this structure is actually predictive of false belief score in regression analysis. In de Villiers (2005), a possible developmental sequence for the acquisition of the meaning of mental state verbs is spelled out: first, syntactic evidence allows the child to classify verbs like *think* in the same way that verbs like *say* are classified. In other words, children firstly familiarize themselves with communication verbs, which can take false sentences as their complements. This happens relatively soon after the child has started mastering the structures underlying the sentences because the child can come across instances where the uttered sentence is in contrast with reality. This is then transferred to verbs like *believe*, *think*, and so on, which share the syntactic structure but do not have such an easy relationship to observed reality. Such an account clearly implies a strong role for acts of mischief-making, lies, and pretense: it is the experience of this difference between what is said and what is happening that plays an important developmental causal role. Moreover, the account presented in de Villiers relies on specific grammatical representations: as argued in de Villiers (2018), a PoV shift occurs only when tensed finite complements are included, as is easily seen if you consider the following examples:

(5) She said to throw out the food in her fridge over there yesterday.

(6) She said he threw out the food in her fridge over there yesterday.

In (5) there is no determination of truth: from the hearer and speaker's point of view, the event is *irrealis*, and there is no possibility of determining whether the

event in the clause happened or not (i. e. there is no determination of whether or not the food was thrown out yesterday). On the contrary, in (6), there is a determination of truth, in the sense that the embedded sentence is either true or false and the truth of the embedded sentence is, de Villiers argues, "subject-oriented". The same occurs for the tense mark: it is indetermined in (5) in relation to the speaker's utterance, but it is marked as clearly precedent to the speaker's utterance in (6), in which de Villiers considers it as part of the subject's matrix. This way, she argues, marking of *tense* and *truth* pattern together, with "the tensed variety explicitly provid[ing] a different perspective, namely that of the matrix subject".

In line with this, de Villiers notices that three fundamental factors have to be understood by the child to get to a "full representation" of a verb like *think* :

> The crucial point is that the child must put three important pieces together to get the full representation of a verb like think. The first is its lexical meaning, referring to some hidden activity or state of the mind but offering no potential clues as to its prepositional nature: it could be a mirroring of reality, as Perner (1991) proposed. The second is its syntactic structure, assessed across a range of different contexts, which provides clues that "think" takes propositions as its content. The third is discovering that these embedded propositions can be false compared to the world. (de Villiers, 2005, p.196)

A fundamental assumption in this description should immediately jump out at the observers, i. e. that de Villiers theory is committed to a representationalist stance in two ways: (1) at the linguistic level and (2) at the conceptual level. (2) is an assumption about mentalizing, or in this case ToM: it is assumed in this account that reasoning about false belief implies operating with the concept of belief. (1) is an assumption about psychological representations of grammar and linguistic information: it is assumed here that to comprehend and use the verb *think*, one has to have a (potentially implicit) representation of the verb in question. However, de Villiers does disentangle her proposal from a "conceptual" proposal, as she names it: the idea is that conceptual development does not come before linguistic acquisition, but instead linguistic acquisition precedes conceptual development in an important way. In de Villiers' account, false belief reasoning in the explicit form is not possible without linguistic acquisition.

In one of her most recent versions (de Villiers et al., 2014), de Villiers argues that *recursion* is the factoring change not only in the understanding, but also in the *representation* of second-order false belief. This, along with the idea that reasoning about false beliefs involves complementations and that embedding is essential for it, places some emphasis on the role of recursion in cognition in gen-

eral. Given that the focus of the book is first-order false belief reasoning as part of mentalizing skills, this issue will now be put aside.

De Villiers (de Villiers, 2014) leaves a number of possible interpretations of her account open, some of which are in line with LoT, and some of which are not:

1. Language *triggers* the corresponding representations in cognition: in this version, linguistic structures connect, when acquired, with already existing structures in a language of thought;
2. The representations of false belief reasoning are parasitic on the results of specialized linguistic means;
3. False belief reasoning is itself linguistic;
4. Language provides an enhanced mnemonic device for avoiding a load on processing capacity;
5. Language provides new computational power: for example, it might provide symbols that map together different situational input and different occurrences.

Let us analyze these options in order. In 1, the possibility that is indicated is that acquisition of natural language representations interact with LoT representations."Dormant" representations are then activated thanks to linguistic input. The following question naturally arises: if the available representations are already existent, why do they need language to be present, especially considering that other forms of intention reading and mentalizing seem to be present before linguistic input? In this sense, it sounds like either language contributes actively to the new representational means, or it constitutes the representation itself, as in 3. This, however, seems to be at least partially excluded by the data presented in chapter 4; data from bilingualism, aphasia, and SLI do not support a constitutive relation. Also, it would imply that at least some part of thinking, even when not conscious, occurs in language, which would imply taking one of the two following options: either this is an exception, and most other cognitive processes do not require language, or thinking strongly relies on linguistic input. While 2 was excluded in chapter 2, 1 seems to be at least implausible, because of the fact that other data presented in chapter 4 seem to suggest that at least some precursors of mentalizing skills are non-linguistic, which would make false-belief reasoning a very notable exception in this sense. However, the option is worth considering. Option 2 obviously implies steering away from a conception of LoT that is entirely Fodorian, and possibly assuming something like that which is argued for in Gopnik (2001) and Carey (2009): language contributes to cognition by providing new structures. Alternatively, this position is possibly in line with Tillas' (2015a) idea, that thought can piggyback on language for

some structures to be at least arranged in a way that is novel, thanks to associations between linguistic and conceptual entries. Notably, this is also compatible with Clark's proposal, that language provides for the enhancement of cognitive tools. The same can be said about 4 and 5: these options resonate with many accounts presented in chapter 2, including LASSO, LFH, and Carruthers' proposals.

As I argue in Berio (2020a), the Bootstrapping Hypothesis presents a variety of advantages and disadvantages. An obvious strength of this account is that it provides rather specific means for tracking the role of syntactic information. However, there are also equally clear problems: as formulated, the syntactic bootstrapping hypothesis relies heavily on the idea of internal means of representation which have grammatical features. While in some of de Villiers' work (de Villiers, 2014) it is argued, albeit not in relation with false belief reasoning, that an I-language is necessary for the acquisition of determined concepts, it is not clear that I-language as conceived by Chomsky is indeed representational. In this sense, de Villiers seems to be relying more heavily on Hinzen's (2006) minimalist proposal when arguing that I-language has to be the format in which some reasoning occurs. The main assumptions are made following Hinzen (2006) and also in line with the generativist tradition, according to which we can effectively distinguish between E-language, the spoken language that the Chomskyan tradition sees as not suitable as an object of inquiry, and I-language, the ultimately structured format that allows for the mental representation of complex thought.

> All and only humans have the capacity for I-language, that is, syntactic structures that permit the representation of structured meanings at sentence-level complexity and above. (de Villiers, 2014)

What seems to be fundamental is the claim that an I-language, so a level of representation that cannot be identified with E-language nor with non-linguistic thinking, is needed in order for some propositional representational format to happen. While the above quote is not explicitly addressed to the issue of the attribution of mental states, it is argued in de Villiers (2014) that I-language can provide a representational format that handles representations of a propositional nature, not possible without language, and therefore adds a powerful new representational tool to the mind of the child. In this case, the bootstrapping mechanism would be possible because acquiring E-language shapes I-language, that is, the means through which we represent structured meanings. What is then advocated is a Chomskyan distinction between the two different levels, and in particular the prevalence and pre-existence of I-language compared to E-language. Moreover, consider the proposal made in de Villiers (2004):

> By what means does the syntactic mastery enable the reasoning? It seems unlikely that the children formulate complete sentences in their native language in the process of answering the false belief questions. But could they generate enough of the underlying representation to support the reasoning without formulating phonological forms, that is, could they use LF to reason? (de Villiers, 2004, p.12)

Here de Villiers seems to rely on the notion of logical form as suggested by Carruthers and exposed in chapter 2: this, however, implies committing to a very specific architectural vision of cognition, and one that is not particularly neutral. In Carruthers' own view, moreover, language is mostly a means for *expressing* mentalizing skills, or at least this is considered the more plausible option when compared to the idea that competent language use is needed to have a concept of belief (Carruthers, 2002).

Finally, a general concern with this view is that there is no intuitive way to conceive of this mechanism as in continuity with any of the pre-4-year-old skills that seem to be present in the almost-mindreading child. In other words, if false belief reasoning emerges completely as the result of syntactic acquisition, it is hard to see what the relation with the other mentalizing skills, that nevertheless seem to be precursors to false belief reasoning, could be. While this is not a conclusive objection, coupled with the fact that the bootstrapping hypothesis relies so heavily on a representational format that is sensitive to linguistic features and given the very strong predictions it makes, i.e. that syntax is both sufficient and necessary for understanding of false belief reasoning, it gives good grounds to be cautious. To show more limitations of the account in more detail, I will start by presenting some critiques that have been made of it, while presenting an alternative in the following section.

5.3 Linguistic practice: the role of pragmatics

In Van Cleave and Gauker (2013), several points are raised that contradict the analysis presented above, and they are worth mentioning not only to better comprehend the debate, but also to highlight some necessary points that are central for the current discussion.

Firstly, it is worth emphasizing that the most popular objection to de Villiers' account comes from Perner et al. (2005), where he presents a fundamental challenge to the sentential complement account of false belief understanding. The challenge can be easily comprehended by recalling the data presented in chapter 3, and it has to do with cross-linguistic data. Let us have a look at the example construction in German again:

1. Die Mutter will dass Marta ins Bett geht;
   *The Mother wants that Marta goes to bed.

2. Die Mutter glaubt dass Marta ins Bett geht;
   The Mother thinks that Marta goes to bed.

In the constructions above, *will* and *glaubt*, respectively *wants* and *thinks*, behave similarly in being able to take a sentential complement, allowing the complementizer phrase (CP) to have exactly the same characteristics spelled out above by de Villiers for communication and mental state verbs. In other words, words for desire take a tensed sentential complement in German when the subject of the main clause and that of the subordinate clause are not the same; at the same time, and this is what really raises problems for de Villiers' account, sentential complement tasks with the verb *want* are passed by German 3-year-olds, which is obviously earlier than children pass FBTs (Perner et al., 2005).

De Villiers does reply to the challenge by pointing out, on the one hand, that a hidden syntactic marker could be the reason why the two constructions are, in the final analysis, not the same. However, recognizing that the postulation that children are sensitive to a deep grammatical feature in such a strong sense might be hard to swallow, de Villiers also brings into the discussion another important concept, i.e. the distinction between *realis* and *irrealis* (de Villiers, 2005). As explained in chapter 3, the "direction of fit" is different for desires and beliefs: a desire is a mind-to-world state, while a belief is a world-to-mind state. While Van Cleave and Gauker (2013) deems the explanation not necessarily clear, I do think that it raises a good point: the attitude of wanting something seems to be different from that of believing something, in several ways, one of which is definitely the fact that desire verbs express a potential change in the configuration of states of affairs, whereas *know* and *believe* are not "proactive" in this sense, as they are supposedly positions on states of affairs that are (or are not) already the case. The relation between the embedded sentence and reality is different, and I think this might be visible when looking at the tense taken by the embedded verb too. While it is indeed acceptable to have (7), it is not straightforwardly acceptable to have (8), nor (9)[37]:

(6) Die Mutter will dass Alex ein Lied singt.
   *The mother wants that Alex sings a song.

37 The native speakers I consulted agreed that the constructions sound strange; however, I have not carried out empirical research on this point.

(7) *Die Mutter will dass Alex ein Lied singen wird.
*The mother wants that Alex will sing a song.

(8) *Die Mutter will dass Alex ein Lied gesungen hat.
*The mother wants that Alex has sung a song.

But the same does not apply to *think*. All the following constructions are fine:

(9) Die Mutter glaubt dass Alex ein Lied singt.
The mother thinks that Alex sings a song.

(10) Die Mutter glaubt dass Alex ein Lied singen wird.
The mother thinks that Alex will sing a song.

(11) Die Mutter glaubt dass Alex ein Lied gesungen hat.
The mother thinks that Alex has gone to bed.

Now, the fact is that this syntactic difference might indeed be a relevant distinction: as a matter of fact, constructions like the one in (12) are especially powerful in underlining a distinction between what is thought to be the case and what is the case, since using the past points to falsity in this case. However, this is not the kind of syntactic marker de Villiers talks about, and Van Cleave and Gauker (2013) are right in underlining that the *realis-irrealis* distinction is not a superficial, easily identifiable distinction, which begs the question whether a hidden syntactic marker can do the job de Villiers attributes to the sentential complement structures. It is worth noticing, at this point, that it is recognized by de Villiers herself that the distinction might be semantic (de Villiers and de Villiers, 2009). In the case of a semantic distinction, however, it is not clear if the original proposal, i.e. that the syntactic information contained in the sentence is doing the explanatory work necessary to explain the change in performance, still holds. In other words, if the difference boils down to a semantic distinction concerning the meaning of wanting something against the meaning of thinking that something is the case, we are left wondering how this distinction gets acquired in the first place: this is after all the reason why so many have argued in favor of *conceptual* change in infants (Gopnik and Astington, 1988; Gopnik and Goldman, 1993; Gopnik et al., 1999; Perner et al., 1998; Perner and Ruffman, 2005; Perner et al., 2005), which brings us partially back to the points raised in chapter 3.

As an alternative to de Villiers' account, Van Cleave and Gauker (2013) propose to shift the focus from syntactic features to other relevant aspects of language, more precisely to linguistic practice. Their proposal is to overcome the tendency to think about the function of attributing beliefs and desires as that of *explaining* behavior: the assumption, they claim, that explanations of what other people do are the main reason why we express ourselves through mental

state attributions is far-fetched. As an alternative, they propose to consider a progression of different functions that mental state verbs and desire verbs have. For desire verbs, they distinguish the following variety of functions:

1. *Command-conveying attributions of desire*. These, despite being listed as third in the list proposed by Van Cleave and Gauker (2013), are according to the account the functions from which all the others derive. These attributions are thought to be in place when, for example, a child recognizes a command given from a caregiver: e. g. "Dad wants me to pick up my toys". This requires the child to understand the function of a command, and the authority that comes with it.
2. *Attributions of expressions of desire*. These are basically reports on what another person has expressed; if a child says to an older sibling "I want to draw", the sibling might report to the parents "She wants to draw". In this case, what is necessary is to understand the expression of desire in the first place, and the ability to report on it.
3. *Expressive attributions of desire*. In this case, first-person utterances are the vehicle for expressing desires; the authors stress how this does not necessarily amount to attributing the concept of desire, since the child might just express "I want to go out" as an instrumental action to get what she desires, i. e. to go out, but without having a concept of what his willingness entails.
4. *Need-conveying attributions of desire*. This happens in the case of an overarching goal: if I want to carve a pumpkin and I am opening all of my kitchen drawers, a speaker might utter "She wants to find the carving knife" in describing my behavior. This attribution of desire, Van Cleave and Gauker (2013) argue, entails the ability to grasp stages and outcomes of a plan, and to recognize that an agent is performing the action. In this case, the function might be that of communicating the need to a third person, e. g. to elicit some response that help me to find the knife.
5. *Explanatory attributions*. The same act can be used as an explanation of my looking-in-the-drawers behavior.
6. *Predictive attribution of desire*. In this case, judgements are made about what a person will do in light of their desires; Van Cleave and Gauker (2013) consider the case a limiting, mostly marginal one.

Van Cleave and Gauker (2013) claim that, while 1 is logically prior to the other functions, 2 and 3 are developmentally prior to the rest of the functions. In other words, children will start expressing their own desires before being able to characterize them as desires, and before they attribute desires to others; however, the verb *to want* is used to express desires in this way mostly because children learn it in the context of the other functions, primarily 2 and 3. While the

discussion about beliefs is tangential to the core topic of this chapter, it is useful to bring attention to it because of the way that functions of belief are listed, i. e. following the same criteria:

1. *Expressive attributions of belief*. As in the case of desires, these do not entail the attribution of a belief to oneself, but express a belief independently of the comprehension of it.
2. *Repetitions of expressive attributions to belief*. These are simply repetitions of what is expressed by another speaker, and are in this sense mechanic and automatic.
3. *Indirect-discourse*. In this case, attributions of beliefs are used as a way to quote somebody else's expression in speech without quoting them literally. This requires a high level of linguistic mastery, the authors claim.
4. *Paraphrasing attributions of belief*. This attribution consists in inferring, from behavior and from what people say, what their position on a particular subject is. Van Cleave and Gauker (2013) phrase this as not a completely explanatory "insight" about other people's states of mind, because it can be based on "linguistic commitments" more than on internal states. (Van Cleave and Gauker, 2013, p.315)
5. *Explanatory attributions of beliefs*. These are often made in the case of excusing a mistake, and are essentially repairing strategies.
6. *Predictive attribution of belief*. Van Cleave and Gauker are not convinced that this kind of function ever occurs: it is possible that, when the child says that the puppet will look for the marble in the right box, for example, and explains it by thinking that the puppet *thinks* it is in the right box, they use the fact that the puppet will look in the box as the *means* for identifying the mental state, and not vice versa. In this sense, the predicted behavior is used to individuate the belief.

Several points have to be discussed about the entire division. Even in the case of beliefs, type 3 attributions are supposed to be the "original" ones: it is the fact that indirect discourse attributions of beliefs are part of what the child is exposed to, that makes it likely that "I think" can be used as a hedge in expressive attributions of thought (Van Cleave and Gauker, 2013, p.316). The distinction between the different functions, or steps, is, to me, at best arbitrary. When paraphrasing attributions of belief, for example, we are likely to be *explaining* what has been said in light of what we think the mental state is; the fact that our evidence is linguistic commitments, instead of looking behavior, or movement, does not seem to be a relevant factor in trying to identify the communicative function. If one considers statements and linguistic evidence as part of behavior, as one definitely should, there is no real conflict or difference between

inferring somebody's mental state from what they say or do. Clearly, this does not mean excluding the fact that the child might only be reporting other's direct statements: but in the case of paraphrasing, explaining a series of statements in light of one attitude is *exactly* the kind of mentalistic explanation mentalizing theories are looking for. Consider further the case of the predictive attribution of belief, with respect to which Gauker and Van Cleave argue that a child might be using the behavior to explain the mental state and not vice versa: this can definitely be the case, when the behavior is clearer to the child than the mental state. However, it is not clear how the child is supposed to come by possession of information about behavior that has not yet been displayed, when instead it has the material sufficient to attribute a mental state to the subject (given that the child, for example, knows where the puppet has put the marble). In other words, the relationship between mental state attribution and behavior is supposed to be such that, since the two things are interrelated in folk-psychology explanations, what is known or inferable explains what is less known or less inferable. The fact that a mental states can be used to explain behavior, in other terms, does not exclude the possibility that behavior can be used to explain mental states.

Let us go back to the account of false belief understanding in Van Cleave and Gauker (2013), which they partially base on Harris (2005): it is conversational practice, they argue, that predicts the performance in FBTs. In the container task, it is maintained, the basis for the child's understanding that the agent will expect there to be smarties in the box, and not pencils, is based on the fact that people's assertions about the content of a container are based on the typical content of the container:

> But in order to have taken an interest in and discovered such relations between a person's observations and his or her assertions, the child will have to have had experience with discussions in which he or she had to decide whether to accept and act on another persons assertions. (Van Cleave and Gauker, 2013, p. 316)

What is not completely clear in this explanation is why the relation between a person's observations and their assertion is not part of a mentalistic understanding of the mind: after all, understanding that somebody says something because of what they have seen before is somehow working on the assumption that a non-necessarily verbalized experience (possibly identifiable with a mental state?) is the explanation of verbal behavior.

However, let us focus on the explanation given for the change-of-location FBT: in this case, it is argued,

> If we take success on the change-in-location task as a test of the childs understanding of the concept of belief (on the grounds that success on this kind of false-belief task is well correlated with success on others), then we must be supposing that the child in effect says to itself "Sally thinks the marble is in the basket". In that case, by the present account of indirect discourse attributions of belief, when the experimenter asks the child, "Where will Sally look for the marble?", the child understands the experimenter as in effect asking the child for Sallys contribution to a conversation concerning the location of the marble. The child has to understand that in a conversation about the location of the marble, Sally would assert, "The marble is in the basket" [...]. (Van Cleave and Gauker, 2013, p.317)

Unpacking this claim, it seems that the explanation of success in the task depends more on the child's understanding of a hypothetical conversation with Sally, than on attributing to Sally mental states about the location of the toy. While this sounds like a simulation ability, or as though some hypothetical scenario is going to play a role, one might wonder why this should be considered an easier explanation than mental state attribution, since it seems to imply that the child has to engage in hypothetical reasoning that is not strictly necessarily. Also, it seems to presuppose that a mental-state-based explanation includes some inner speech about Sally's beliefs; this, however, is hardly what is commonly assumed in the literature.

The explanation given by the authors about the emergence of desire attribution earlier than belief attribution is also not completely clear: in the case of desire, they argue, understanding commands seems to be easier than understanding the joint activities involved in indirect discourse attribution of belief. While this is probably true, they build the entire case of indirect discourse attribution on a fictional example in which two children have to get ready for a party and have to accept the fact that another character says the party is starting later than they think. In this case, one child might report on the third person's belief by uttering "Markie thinks that the party starts at 4 o'clock"; doing this, it is argued, requires understanding planning and collecting information pertinent to the plan. This is clearly not the only case in which such an indirect attribution might happen, and a plan might not always be involved.

Of special interest, on the other hand, is the explanation given by the authors as to the role of sentential complements in facilitating false belief task performance. The reason, they claim, is related to the fact that children develop critical thinking and come to the ability to critically accept or reject orders, and what they are told it is the case: sentential complements facilitate this because critical evaluation of what they are told to believe and do is only possible via linguistic means. This means assuming that this kind of reflection, about contrasting desires not necessarily fitting with their view, is only possible either openly verbally or in some kind of inner-speech format, so only possible with language. There

is a strong sense in which this is an empirical question, and while the authors simply replicate the conclusions reached by de Villiers and colleagues, i.e. that linguistic acquisition is essential for false belief reasoning of that kind, they go even further, maintaining that critical reflection on conflicting desires and knowledge can only be obtained linguistically. Depending on what exactly is meant by "reflection" in this case, different consequences can be drawn. If this reflection on other people's expectations and their desires/commands includes cases in which the child has to understand that two agents might have conflicting desires, it should be reflected in a lack of this skill in non-verbal animals, but this is of course hard to test. It also does not seem to fit perfectly with the data presented in chapter 4, where it was pointed out that linguistic impairment does not necessarily seem to imply false belief reasoning impairment. Finally, what is puzzling is that the proposal of "reflection over assertions" does not seem to relate, after all, to the false belief task performance at all: it is not clear how the critical reflection over assertions should bring to understand Sally's behavior when looking for her toy in the wrong box, or at least not straightforwardly so.

Although not convincing, the account given by Van Cleave and Gauker (2013) underlines the fact that the pragmatics of mental state verbs possibly has an important role to play as far as the relation between linguistic skills and mentalizing goes. For them, it is pragmatic mastery of the terms that results in mentalizing skills; while this might not be the complete story, it does bring some attention to some relevant facts, among which is the fact that desire attribution seems to be involved in rather different communication practices. Moreover, the fact that results conflicting, or difficult to integrate, with de Villiers' approach have to be considered is of essential importance. In this sense, while the approach presents several difficulties, it also has the advantage of being able to account, for example, for the relation between the development of pragmatic skills and mentalizing skills presented in chapter 4, and to propose an alternative to the syntactic bootstrapping mechanism that deviates from a strong role for syntax. This particular idea will be fundamental in the next chapter. As will be clear from the following chapters, moreover, the role of conversation is stressed by a good number of alternative accounts.

5.4 Language and social input: Garfield's proposal

Not all of the theories analyzed here are full-blown theories. The theory advanced in Garfield et al. (2001) is a case in point. However, it is worth mentioning because the authors rely heavily on the kind of data that has been previously

analyzed, and because they show a general tendency to combine features of other accounts.

In Garfield et al. (2001), the authors propose a model according to which language acquisition and social cognition factors are *jointly sufficient* and *individually necessary* for the development of ToM. Garfield et al. (2001) proposal is (weakly) modular and tries to reconcile some aspects of innatism, which makes it an extremely interesting case for the current analysis. As has already been explained, strong modularism is taken to be strongly falsified by data showing that the interaction between cognitive processes and mechanisms seems to be far more general and omnipresent than a rigid modularism would imply. However, modules do not need to be conceived of as innately determined, which lies at the heart of the proposal by Garfield et al. (2001). Garfield et al. (2001) draws on evidence mostly from William's syndrome, deafness, and ASD in proposing a framework in which adequate social and linguistic skills are jointly causally sufficient and individually causally necessary for producing ToM. The fact that ASD children are impaired in both mentalizing and language, that William's syndrome children seem to lack second-order false belief understanding but to perform very well in standard FBTs, along with their social skills, and that deprivation of language input seems to bear important consequences for social cognition for deaf children, they argue, all point in the direction of an interrelation between social input and language input in creating a complex dynamical picture of Theory of Mind.

In this sense, the hypothesis of a single innate ToM module gets discarded, but it gets substituted by the interrelation of general purpose innate language acquisition mechanisms and social intelligence modules. The modular features maintained in the account, then, are *fast* and *mandatory* processing; this last feature, according to the authors, is not disputed in the literature. However, notice that this is not such a strong assumption in the literature any more: as will be exemplified in the next chapters in the course of explaining Apperly and Butterfill (2009)'s synthesis of the mentalizing debate, there is some consensus that some false belief reasoning might not be as fast and mandatory as envisaged by Garfield et al. (2001). Interestingly, the authors address de Villiers' proposal suggesting that, while it accounts for the necessity of language to develop false-belief reasoning, she does not take into consideration sufficiency conditions: this objection, which might appear vague at first sight, does stand. While de Villiers' studies and theories do prove a connection between the development of syntactic skills and mentalizing, they do not give much space to the larger question of (1) whether the mastery of sentential complement structures is the only way to acquire mentalizing abilities and (2) what it entails for a larger picture of cognition, as already mentioned above.

The framework in Garfield et al. (2001), on the other hand, relies on the fact that humans are wired for social communication and language; given the interrelation between these two "settings", we learn to use mental states (and the concepts thereof) in explanations of behavior. In this sense, they argue, acquiring language means developing a set of representational and interpretative practices, and these are the foundation of the mentalizing skills developing consequently. In this picture, development sees a change when language enters the scene. ToM piggybacks on language also with respect to the kind of structures that are used in mentalizing, it is argued: in other words, the fact that we give explanations of attitude ascriptions is what guides our folk psychology to rely on representations that have a propositional format. It is familiarizing oneself with propositional ascriptions as a way of talking about the world, then, that leads to the formation of internal representations that are structured in terms of propositional attitudes. Note that an especially interesting testing ground for this claim would be the exploration of those cultures and languages that do not seem to use mental state ascriptions as much as Western languages, as documented in chapter 3.

Considering language as an integrated part of a system that includes basic low-level skills, they adopt a Vygotskian perspective (explained in chapter 2, Vygotsky (1962)), according to which language, firstly a means of communication, ends up through a scaffolding relation to be integrated in thinking, and enhances in this way the possibilities of cognition. Only at this late stage does ToM develop fully, and this happens in the context of a *social matrix*; in other words, the interaction with others is what embeds this use of language, which becomes the privileged way of interacting with some cognitive demands. *Concepts* of mental states, in this picture, are not necessary for children to engage in social cognition activities; this assumption is particularly relevant, because it frees space for an explanation of the intention reading abilities demonstrated by implicit false belief studies, embedding them in a developmental story that goes from basic mindreading to advanced, language-dependent social cognition.

The framework does not specify which kind of features of language make which kind of mentalizing work easier, only assuming that part of de Villiers' story is correct. Most importantly, it lacks specification of how this new representational format introduced by language interacts with other forms of thinking. There is also a sense in which the proposal makes an unnecessary assumption, i. e. that of the modularity of linguistic and social skills; the issue, however, can be set aside. What is relevant is that, while the proposal does not address some necessary details, it does address the co-existence of social and linguistic factors as intertwined in the development of social skills, and it accommodates (albeit partially, for lack of specificity) most of the empirical literature described in

chapter 4. In the next chapters, it will be argued that some of these shortcomings can be overcome by proposing a specific way in which social practices and language input come together to enhance mentalizing abilities. Before that, other theories will be analyzed, including proposals that stress the role of social input without making modular assumptions, as Nelson's *community of minds* proposal does.

5.5 A community of minds: Nelson's proposal

Nelson (2005)'s idea is based on the partial rejection of the label "Theory of Mind", for reasons that are partially in line with simulationist accounts outlined above, but also differ slightly. While opposing the theorist view on mentalizing per se, as a conception of hypothesis testing that is highly debated in the mentalizing literature, she also proposes to take a direction that is more centred on the social aspect of mentalizing practices. In this sense, the view goes in the same direction as Garfield et al. (2001)'s proposal, in stressing the social and interaction component.

The attempt is to put standard Theory of Mind tests in context considering that, while the laboratory situations cast some light on the processes underlying mentalizing skills, most of our social cognition is based on interactions within a community, where complex nets of beliefs are entrenched with and related to social nets of beliefs that include many individuals, their relations, conventional and moral rules, and so on. In this sense, mindreading activities are to be understood as processes including a *community of minds*, so a community described and interacted with in terms of mental states. The community develops through narratives used in fiction, imagination, moral rules, and so on. Note that this point is in many ways similar to that of Hutto (2008b), but is also in line with proposals about metacognitive language being essential for the development of ToM.

According to Nelson (2005), the first access to knowledge is "private", so mainly deriving from perception and motor information: without language, there is no way to communicate and learn about one's own mental states and mental world. Language is what provides perspective; while listening to her parents, the child slowly acquires information about other perspectives, different from her own. Once again, this shows some thematic connections with de Villiers' work and connects fairly well with what emerges from the previous empirical review: perspective shifting is central in false belief understanding and mentalizing, and language contributes to the shift in taking into consideration other's perspective.

This first "private" phase, which also includes joint attention activities, shared goals, complex interactions, and imitation, is not considered by Nelson (2005) to be a representational phase; once again, then, linguistic knowledge introduces a shift in representational abilities. With language, a mechanism of abstraction is necessary, since talking about mental states is an abstraction used for entering the community's language about other's actions. When children start using mental terms, according to Nelson (2005), it is still too early to attribute to them *concepts* of mental states, as these emerge gradually, are particularly complex, and allow for looser use in previous stages of learning. Note that this assumption, which is in open contrast with how traditional ToM theories characterize the developmental evolution of mentalizing, is not on the other hand incompatible with the data presented in the previous chapters, assuming that an understanding of mental state concepts that is only partial is definitely in accordance with the idea that development goes from intention and goal directness to more sophisticated representation of beliefs. Also, it is definitely also compatible with the fact that linguistic development sees children use mental state terms with an increasing variety of complexity and elaboration.

A central role is given to narratives like stories that are used in the education and formation of the child; these require representational means that are fairly sophisticated, in needing to keep track of different mental states, sets of beliefs and desires, and so on. The developmental pathway proposed, then, is as follows:

> [...] through relevant conversational experiences, the child is exposed to increasingly complex and extended uses of representational language and comes to master the skills, involving short-term memory and semantic interpretation – that are necessary for the comprehension of such linguistically formulated messages. Next, the child becomes capable of repeating to self or others what has been heard on the same or a later occasion. (This is reflected in tests of receptive language. It also appears in the repetition of stories or of other people's experiential reports.) Then the child may begin to use verbal representations both to compose stories or reminiscences (reflected in expressive language) and to serve as internal cognitive representations, enabling the duality of mental representations. (Nelson, 2005, p.44)

Note that language is used as an *internal* representation system that allows for entertaining conflicting perspectives, keeping track of relevant mental states, and representing alternative possibilities, two different states of the world, and so on. Representational language allows these kinds of representations and is then an entry point into the community of minds.

Nelson's position fits the landscape described so far in various ways. The idea, once again, is that language provides new representational means,

which bring on a qualitative change in the interactions; in this sense, it not only enhances mentalizing skills, but brings new possibilities to actively engage in the social practices that characterize the community of minds. This is in line with approaches like Carey's Quinean bootstrapping view as well as Clark's predictive coding take, which are two approaches at opposite sides of the spectrum in many ways. In a sense, Carruthers' inner speech hypothesis also qualifies as one that allows language to be a specialized representational means, but the rest of the assumptions of his cognitive model do not make the option especially compatible with such a social-centered view. All in all, Nelson's story does not make specific predictions, so it is hard to evaluate against the background of specific claims about the role of language, despite its clearly fitting with the language data. Something that is valuable about the proposal, however, is that it indicates how it is possible to take into consideration a socially-oriented approach that does not dismiss the role of language as a cognitive enhancer. As will be see, this theme is central in the rest of the chapter.

5.6 Pragmatic roles: Montgomery (2005)

Montgomery (2005) presents an account of the influence of language on mentalizing that is heavily entrenched within a view of *meaning* itself, and which proposes seeing the role of language as providing pragmatic roles comprehension more than concept formation. In this sense, the proposal can be seen as emerging from a pragmatic approach to word meaning, where the central assumption is that the meaning of mental state terms derives from the pragmatic ends for which these are used, and not their semantic meaning. In this sense, the proposal is to be read as a reaction to the *ostensive* paradigm of word learning, which assumes that meaning derives from correctness in the relation between the denoting term and the corresponding state.

Montgomery, in this way, rejects the idea that mental state terms are referentially related to psychological states; in this sense, the proposal puts into a different perspective the entire idea of mentalizing, casting doubt on the idea that it involves actively attributing a psychological intentional state to a subject. This, as said, is a view shared by Nelson (2005); after all, Montgomery argues, mental state terms can be used by a child to express internal experiences without necessarily having a concept of the internal experience itself. In other words, saying "I want ice-cream" might not be equivalent to possessing the concept *desire*, at least not in the sense in which such a concept is supposed to be a *theory* of what desires are. On the contrary, uttering "I want ice-cream" can be a way to participate in the linguistically developing game of expressing our tendencies in a spe-

cific socially regulated interaction; notice that this emphasis on the role of communicative intentions is exactly what results from Van Cleave and Gauker (2013)'s analysis as well. This means that the child asking for ice-cream has learned the linguistic rules that govern the game and how to obtain ice-cream, but does not necessarily have an idea of what the concept *desire* is. In Montgomery's example, saying that something "hurts" is an expression of the mentalistic experience of pain, and it can be meaningful without the child possessing the concept of pain at all.

The intuition that learning mental state terms and their use is less a matter of learning by ostension and more a matter of learning roles is not new in the literature, and heavily relies on Wittgenstein's private language argument in many ways (Wittgenstein, 1953): following a "rule" for the application of a mental state term is problematic, because of the impossible distinction between thinking one is following a rule and actually following a rule in the case of a private referent. Rules for the use of language in this sense must be public (and hence "language games"), Wittgenstein argues, because of the sociality of language as an interaction device, and because this is the only way to make sense of the rule-system that governs language use. Montgomery draws on this argument from Wittgenstein to argue exactly this: mental state terms' meaning must be defined publicly. Children's use of language is analogous to that of adults in that they use it to accomplish particular ends and tasks, and their meaning is associated with how they are used in these social practices. In this sense, and connecting to what was said above on the meaning of terms, the child expressing herself via

(13) I know that birds fly.

expresses the content of the belief, but does not describe the nature of believing. The meaning of the expression emerges from the interplay between utterance and response, in the dialogue that predicts a reaction to the child's assertion, and does not have to do with theoretical knowledge of beliefs. Family routines and communication practices, then, are at the basis of the emerging meaning.

Montgomery applies this to false belief tasks as well, arguing that the evidence in favour of the child having a fair understanding of expectations and behavior before the task is passed is a good indication that the social practices and routines underneath mentalizing abilities are already evolving; the intrinsic difficulty in the FBT, then, is that pragmatic clues suggest to the child the mental state terms are used to express something true. As explained at length in chapter 3, initially mental state terms are used to: express true statements; hedge conversationally; or describe successful actions (Bartsch and Wellman, 1995). In the

case of FBTs, children then have to override the conversational assumption, inhibiting it, and assuming that the verbs can be used to express something false.

Montgomery's proposal, while rightly emphasizing the role of social experience and pragmatics, does not address other issues in the literature. FBT performance is not only related to inhibitory control[38], and hence the sole explanation of children's failure might not be best cashed out by appealing to the inhibition of pragmatic assumptions; it is also the case that the evidence in favor of a role for vocabulary acquisition and especially syntactic development in mentalizing performance is somehow left out from the analysis. While Montgomery's emphasis on pragmatic competences address the fact that those seem to be connected to false belief task performance, it does not put this into a bigger picture painted by the data. Similarly to de Villiers, moreover, it is limited to explicit folk-psychology practices, and therefore does not connect fully with a complete story of how, if at all, this has its origins in other forms of mentalizing. Later in this book, a proposal to solve this problem will be presented.

5.7 Baldwin and Saylor: structural alignment

Baldwin and Saylor (2012), while addressing the role of language as a social tool to improve mentalizing, present a different idea. The proposal is that abstraction, as the necessary process for forming concepts of mental states, is achieved by analogy and structural alignment of different situations aided by language. More specifically, language promotes the ability to compare different behaviors of different people across distinct scenarios, providing a way to unify different situations and highlight their commonalities. As is already clear, this position has a lot in common with theories presented in chapter 2 that stress the role of language in "glueing" different scenarios and representations, for example Clark's position.

Baldwin and Saylor (2012) argue especially on the basis of Gentner's work on analogical reasoning (Gentner and Gunn, 2001). "Structural mapping" is the procedure through which alignments are discovered between objects and situations; once the alignments are discovered, they can be "encoded as relations in their own right, meaning that a more abstract level of representation is achieved." (Baldwin and Saylor, 2012, p.125). Comparison of structural mapping, can give rise to iterated and increased levels of abstraction: once an abstracted representation is formed because of structural alignments discovered, the new abstract

38 For instance, Devine and Hughes (2014).

level can make it possible to see other analogies with other representations that were not visible before, and so on. A "source domain" constitutes the already understood cluster of information, in light of which the new cluster of information (the "target domain") is understood, through the research of structural alignments.

The central claim, then, is that language is a powerful tool for aiding structural alignment and abstraction in developing children. In support of their claim, they cite studies like that by Desjardins and Baldwin (1992), for example, where children were significantly more likely to play with a toy in the same way they saw somebody playing with another (only moderately similar) toy if the same labels for the two toys were used. The effect was found with 20- to 22-month-olds, suggesting that labels provide a way to map together only relatively similar stimuli and that infants are already able to do this from a very early age. This kind of aided structural alignment[39], is then applied to the issue of mental state terms: the concrete proposal is that the construction of mentalistic concepts is aided by the use that parents make of the same labels to describe behavior. In particular, the idea is that this goes through the way children learn to refer to absent entities; while mental state terms denote something potentially abstract in nature, dialog with children and child-directed speech are rich with reference to non-abstract, concrete absent entities, like food, toys, covers and people. Baldwin and Saylor (2012)'s proposal is that, when making sense of reference to something that is physically absent, structural mapping intervenes, making room for the idea of an intended referent – in this case an internal focus of attention. Unable to retrieve the pattern *word "dog" – perception of DOG* previously associated with the experience of a dog, for example, the child will use her memory as the *source domain* to understand the absent-referent scenario, i.e. the *target domain* and language will provide a link between the two, aiding the process that will lead to the relating of the word "dog" to dogs, whether absent or present. This can only happen through an inference:

> One such inference that would do the trick is that /dog/ utterances and dog directed gaze are separate reflections of an unseen, internal dog-related thrust or focus (i.e., intention/attention). Under this inference, /dog/ utterances can reflect dog-related focus even when dogs are absent. This inference would also help the language learner to make sense of there being some gaze switching with dog-present dog references but gaze directed primarily at self with dog-absent dog references. If dog-directed gaze reflects a dog-related focus, gaze switching between self and dog then reflects coordinated focus on self and dog. And then structural alignment would lead to the inference in the target domain that /dog/

39 That is reported in adults as well, see for example Gick and Holyoak (1983).

> utterances (reflecting dog-related focus) accompanied by gaze at self (self directed focus) also reflects coordinated focus on self and dog. (Baldwin and Saylor, 2012, p.135)

This brings us to the possibility of the infant's concentrating on a "internal focus of attention" of the interlocutor, that is different from the "external" focus of attention that is signaled, normally, by gaze. Such a mechanism would bring, firstly, to the appreciation that single words and utterances can be explained in light of an internal focus of attention, and later to the understanding that language itself is about promoting communication in terms of sharing attentional focus and influencing it; as is possibly clear already, such a view is highly Gricean, as communication is understood as an attempt to influence each other's mental states. In general, this kind of approach is envisaged to understand how children come to the conclusion that internal states, like mental states, desires, intentions and so on, are to be attributable to other agents, and hence to the formation of mental state concepts. Note that this is partially in conflict with the idea proposed by Montgomery (2005) that an ostensive paradigm should not be applied to mental state terms: ultimately, structural alignment counterbalances the lack of external referent that is present for concrete referents.

While this proposal might not be very detailed, the authors recognize that it might constitute only part of the story, and it does integrate relatively easily in a general picture of cognitive development that does take into consideration all the data presented above, since it is compatible with the fact that language, in various forms, promotes a richer understanding of mentalizing. At the same time, Baldwin and Saylor's proposal is in line with the idea that linguistic information enhances cognitive abilities and provides new conceptual structures, thus being substantially compatible with Clark's proposal as well as Carey's bootstrapping proposal, that new concepts are acquired in light of pre-existing abilities. While the idea of applying structural alignment learning and mechanisms to mentalizing is particularly appealing, the view proposed by Baldwin and Saylor seems to assume that it leads to the formation of *concepts* of mental states, deriving from the inferences that the child makes about "internal focus of attention": while this might be true, it does rely a lot on a view of communication that is strongly Gricean, for a start, and also assumes that a concept of "inner state" or "mental state" is formed relatively early in development, thus bringing us back to the idea that false belief reasoning necessarily implies conceptual use. Moreover, it does not clarify how the influence of syntactic and pragmatic information interacts with the vocabulary component that is central in the structural alignment hypothesis. The idea of structural alignment will be reconsidered in the next chapter; before that, I will analyze theories that make an even stronger claim with respect to language.

5.8 Substantive, explicit mindreading: Bermúdez (2009)

An account that relies a lot on linguistic knowledge for mindreading is that presented in Bermúdez (2009), where the focus is placed again on propositional structure rather than on social input. On this view, language is essential for mindreading abilities of at least one kind; as a matter of fact, Bermúdez distinguishes between *minimal* and *substantive* mindreading, in a way that recalls other divisions present in the literature[40]. Minimal mindreading implies systematic covariation: in other words, it implies that variation in the psychological states occurring in the mind of the social actors produces a variation in the "mindreader"'s behavior. On the other hand, *substantive* mindreading occurs when there is actual attribution of mental states.

"Minimal mindreading" (MM), Bermúdez specifies, is a descriptive characterization, not meant to provide an insight into *how* the covariance occurs. On the other hand, *substantive mindreading* (SM) is intended as an explanatory label, since it involves the idea that the behavior of the mindreader depends on her representations of the psychological states of other agents. Furthermore, more levels of substantive mindreading can be individuated.

The core idea is that SM requires relating psychological states with the background of behavior. In other words, to engage in SM, we need to relate the psychological states that are being represented with the related behavior of the holder of these psychological states, in order to make the SM a fruitful activity. The complexity of this relation, between behavior and psychological states, varies substantially according to the situation: for example, Bermúdez argues, it is without much insight about individual psychology that we can predict that leaving a lot of money on the sidewalk will result in somebody who is watching taking the money. In such a case, general rules and heuristics about human behavior will suffice, and no complex psychological profile will need to be given. Moreover, no explicit representation is needed for this operation, making this a case of substantive mindreading that does not need linguistic form. However, the more complex and variable the details about background psychological profile of the agent become, the more explicit and accurate the representations have to be. Interestingly, Bermúdez argues that, while complexity does not necessarily lead to explicitness, variability does entail more explicit representations: if there are many ways in which the behavior of an agent can be linked to behavior, then it becomes more important to have explicit representations. This kind of explicit

40 For instance, the distinction between low-level and high-level mindreading in Goldman (2006).

SM, then, is what is entailed in explanation and prediction at a personal level, and it is an activity that is either conscious or consciously accessible.

When propositional attitude attribution is involved, it is argued, explicit representation is essential: this is because desires, beliefs, hopes, and so on do not feed into behavior in a direct, unambiguous way, but rather they need to be considered against the background of the psychology of the agent involved and they have to be fed into mechanisms dedicated to reasoning about how the different attitudes interact with the psychological profile considered, and used to formulate explicit predictions. In contrast to "perceptual mindreading", which does not need to involve specific representation about the other's mental states (i.e. it can be based on rules and assumptions of another kind, like the fact that seeing a source of food, or a predator, might elicit certain behavior), *propositional attitude* mindreading involves the explicit representation of the attitudes. At this point, it is fundamental to specify that, if on the one hand Bermúdez assumes that mentalizing of this kind is propositional and representational, he also considerably downplays the role that such mentalizing has in everyday life, giving a bigger role to *frames*, *stereotypes*, and *routines* that are used in most common interactions; the idea is that many social situations have predefinite roles and routines, where agents fulfill specific roles and specific parts, and therefore not much explicit mentalizing is present. The idea is developed at length in Bermdez (2003); Bermúdez (2005), where it is argued that these social routines are pervasive in human activities but, as will be seen, a similar claim is made by Hutto (2008b) and Apperly (2011). In any case, the kind of *frames* Bermúdez refers to are stereotypical social situations (like being in a bakery) that help to solve otherwise computationally demanding problems, i.e. keeping in mind a great load of detailed information about people present in the bakery, how to interact with them, and coupling it with effective processing and working memory load on how to make as smooth an interaction as possible. He relies on the definition of *frames* given by Minsky (1975); as it will be seen in the concluding section of this book, a bit more about frames, understood at another level, can be said. In any case, the idea is that stereotypical frames will have parameters, called by Minsky terminals, that will be available to be filled by different kind of perceptual information; the frames also have attached information about how to behave in certain circumstances, how to use the frame as a data structure, what to expect from such a situation, and so on. As will be seen, these frames can be thought of as situational models.

What is most relevant is that, when frames like these are usable, it is not necessary to have propositional attitude ascription; however, when these are present, they need to be in a format of representation, according to Bermúdez, that makes them:

1. Subject to truth conditions.
2. Usable in inferential reasoning about how they lead to action, i.e. they can be evaluated against the psychological background information present for a given agent.
3. Consciously accessible, i.e. they must be integrated with the rest of the propositional attitudes held by the mindreader.
4. Exemplary of the structure of the represented representation of the state of affairs (proposition) attributed to the agent.

The union of these requirements is what makes it the case, for Bermúdez, that this kind of mentalizing has to be carried out in natural language sentences. LOT does not do the job, it is argued, because LOT is supposed to be a hypothesis about sub-personal codes of thinking, and this cannot be so readily consciously available. The kind of conscious availability that is entailed in propositional attitude mindreading, then, requires a format that is, on the one hand, propositional in structure, and on the other available for conscious processes. It is obvious that the only candidate for such a role, then, is natural language, at least according to Bermúdez.

Note that this view is *not* as radical as it seems, in so far as the role for language is basically essential, but only in respect to a very specific form of mentalizing, i.e. the one that involves the attribution of propositional attitudes in a propositional form, is integrated with decision-making processes, explicitly represented, fully conscious, and deals with strictly truth-evaluable propositions. This includes, naturally, second-order mentalizing of a sophisticated form. What is not clear, however, is what else it includes, in the sense that, for example, it might include standard FBTs like the Anne and Sally paradigm, but it definitely does not include cases like that of Onishi and Baillargeon (2005) and Scott et al. (2011), or Scott and Baillargeon (2009). In a sense, the weight of Bermúdez's claim heavily depends on whether or not cases like FBT attribution in the Anne and Sally case has to be taken at face value as the explicit attribution in propositional form of consciously accessible sophisticated representations, or if it can explained away by another kind of phenomenon, for example, if it can be taken as part of a frame or a social routine like the ones indicated by Bermúdez himself, or if it can be considered an example of another kind of phenomenon that does not involve this very complex attribution. In general, note that, as will be seen in the next session, Sally and Anne's task can be seen as revealing an ability to deal with associations, for example, which makes it a task that can be explained away through less explicit forms of mentalizing. This, however, leaves open, in an account like that of Bermúdez's, why performance in mentalizing tasks like FBTs might be correlated with the acquisition of language and

development of linguistic skills. In this sense, Bermúdez's view does not seem to address the evidence gathered around the role of language acquisition in mentalizing tasks that has been presented in chapter 4.

Bermúdez's claim can be summarized as follows: if representation of mental states has to happen through some propositional truth-evaluable format and in a consciously accessible way, then language is necessarily involved. This is a position taken by many accounts that do not endorse LOT, as we have seen in the case of Tillas (2015a) for example, where the solution to the propositional thinking problem implies appealing to language as the format for propositional thinking. Naturally, it is also possible to think about propositional thinking as independent from language and at the same time not realized in a LOT of the Fodorian kind. What is most relevant though is that Bermúdez's claim is limited to explicit propositional false belief mentalizing, assuming that it is the case that this is distinct from other forms of mindreading. At the same time, the scope of this book is a bit broader in trying to understand if the development of more general cognitive abilities is related to mentalizing, and Bermúdez's proposal itself is openly only dedicated to a very specific, limiting case of mentalizing, which is not routinely used. While his view focuses on propositional structures (in this sense, similarly to de Villiers' syntactic proposal), we will turn now to an approach that sees in a relatively simpler feature of language, i.e. labels, a resource for mentalizing.

5.9 Mental files and mentalizing

An account of the changes that underlie children's development of false belief abilities and takes into consideration mental representations is that proposed in Perner and Leahy (2016); Perner/Haemer/Leahy (2015). It should be specified that Perner's account of mentalizing abilities has been evolving throughout the many years that the author has dedicated to the topic, but has maintained a few fixed points that are worth mentioning, because they are reflected in the mental files approach endorsed now and because they situate his position in the larger debate. Firstly, Perner has for a long time argued that results like those of de Villiers can be reinterpreted: not only was he the one who raised the objection concerning German and the verb *want* employed in sentential complement structures (Perner et al., 2005), he also generally argued that mentalizing is a matter of a general mechanism dedicated to metarepresentation, more than a ToM-specific ability. (Recently presented arguments can be found in Leekam et al. (2008) and Perner and Leekam (2008).) Secondly, Perner has long maintained that several elements, including being able to entertain the perspective of others

(Perner and Leekam, 2008; Perner et al., 2011, 2002) and interaction with siblings (Perner et al., 1994) contribute to the formation of ToM. While the claim that labels and in particular sortals are functional in learning to assign different *perspectives* is already present in Perner et al. (2002), I will focus on newer versions of this claim, which pay special attention to how *mental files* play a fundamental role in Theory of Mind development, since it follows the same line in attributing a role to labels and presents the most recent formulation of the "perspective" view that Perner endorses.

The mental files explanation introduced by Perner (Perner and Leahy, 2016; Huemer et al., 2018) is somehow more heavily reliant on language mechanisms than other accounts he has been proposing, but there is a sense in which this is only a superficial difference: the fact that ToM is in constant evolution, and that social and linguistic information play a role in its growth has been a constant in the theories proposed by his group. In this case, however, applying a mechanism like mental files, which directly stems from the philosophy of language, brings new light to the role of labels in this developmental trajectory, thus making it of fundamental interest for this book.

The idea of mental files stems from a theory famously presented in Recanati (2012) and subsequently developed in literature in various ways. The fundamental and most attractive feature of mental files, is their ability to accommodate the predicative structure of language. A mental file is "anchored" in the referent and contains information about the referent: this is supposed to mark the difference between talking *about* something and *what* is said about it. The idea is that an *epistemic rewarding relation* links a mental file to the object it is originated for; mental files also keep track of information that is gathered about the referent of the mental file; the epistemic rewarding relation can also be of different kinds, as for example I can know about an object because I see it or because I have heard about it. In any case, the *anchoring* relation between the referent and the mental file is of fundamental importance, because it is what makes the file *about* the object; the content of the mental file, on the other hand, is the predicative information.

Any new conceptual perspective on a referent is supposed to open a new mental file; this means that mental files can encode perspectives on objects in separate ways. However, files that refer to the same object are also linked "horizontally". The horizontal marking is supposedly what allows for *flow* of information between the two files, which outside of metaphors means that the information in the two files can be conceived as referring to the same object.

Given these basic elements, let us see how the theory deals with false belief and Theory of Mind performance. In Perner and Leahy (2016), dual naming games are described: children are asked to participate in a game where the ex-

perimenter uses a label for an object, and the child is required to use a label that also applies to the object but does not correspond to the one used by the experimenter. The classic example, used in Perner et al. (2002), is that of "rabbit" and "bunny". The experiment is set up so that children have to help a puppet take part in the dual naming game: the puppet wants to name objects, like the experimenter, but does not want to use the same name. After being tested for their knowledge of different names and labels (e. g. after making sure that they can use both "rabbit", "bunny", and "animal" to name the same thing), children engage in a task where, when the experiment says "rabbit", they have to suggest that the puppet say "bunny", or vice versa. Interestingly, 3-year-olds fail this task, whereas 4-year-olds pass it easily. Crucially, the same problems are absent when children have to perform in the same task but the object in question is a double-color object; when asked to name the color the experimenter does not name, 3-year-olds have no problem performing this task. This pattern of performance is strikingly related to that of FBT, and the connection for Perner and his group is far from casual: the idea is that the same mechanism lies at the heart of both false belief tasks and alternative naming tasks.

One central idea, picked up in Perner and Leahy (2016), Perner et al. (2015), and Huemer et al. (2018), is that different mental files can be opened for the same object, depending on how the object is described, on the assumption that individuating an object implies individuating an object *as* something, and that the same object can be individuated *as* different things. So when I describe a rabbit as a "rabbit" and as a "bunny", my labels are related to two different mental files that are anchored to the same object: in this sense, mental files are perspective-relative, because every mental file encodes information about an object gathered according to a specific perspective on it. At this point, a central point emerging in Perner and Leahy (2016) is that the individuating power that labels have only applies to sortal names, because they are the only linguistic mechanism that allows to uniquely identity one *"as-x" perspective* on the object indicated. In other words, when pointing to an object and uttering "This is a rabbit", I use a sortal that provides unique information about which kind of perspective is given on the relevant object. When uttering "This is an animal" for the same object, it is maintained, I might be opening a different file: this is because I am identifying the object in another perspective, i. e. as an animal and not as a rabbit, which might make different properties salient. However, the same cannot happen with adjectives, since uttering the sentences "This is white" and "This is green" while referring to the same white and green object does not open two different mental files, because labelling the object as white does not uniquely make clear what I am referring to: I could be referring to the rabbit, the spots on the fur, the fur itself, and so on. In this sense, sortal

nouns are thought to be a privileged way to individuate a perspective on an object, which is why two different sortal nouns for an object are likely to be forming two different mental files, while highlighting properties of the object with two different colors does not. This presents some problems in general, mostly due to the role attributed to sortal nouns as uniquely referring to something; to see why, it is sufficient to consider the idea of having different rabbits in front of you. In that case, the label "rabbit" does not uniquely pick out a given rabbit; but exactly the opposite case can be made when thinking about a group of objects in which only one is white, where using the label "white" is a perfectly suitable way to identify the object uniquely. This kind of problem is beyond the scope of the book, but does suggest some doubts about the general approach.

Going back to the alternative naming task, the interpretation given by the authors is relatively simple: while children aged 3 can already create two different mental files, one labeled "rabbit" and one labeled "bunny", for the same object, they cannot link the two files together; since the object has been individuated *as rabbit*, it is re-individuated by the child under the same description, or under the same perspective. When the child turns 4, she has the ability, on the other hand, to perform what has been named in the mental files literature as *horizontal* linking, i. e. to connect two files in a way that allows information to flow between the them. The linking operation grants sameness of reference, and the information in the two files can then be used by the child interchangeably.

Let us assume for the sake of argument that this is what happens in the case of alternative naming; in the case of the false belief task, it is also a matter of linking: in this case, however, the assumption is that other people's beliefs are represented as mental files indexed to other people, or "vicarious" files. A vicarious file is anchored to the same object, but is tagged as belonging to another agent: in the case of Maxi's chocolate bar, for example, the child is expected to have a chocolate file for herself and a chocolate file for Maxi: while the chocolate file of the child contains information about the current location, the chocolate file for Maxi contains information about the location at which Maxi saw the chocolate. In this case, to be able to relate the two perspectives, the child also needs to be able to link the two files; however, this operation cannot be understood as *orizontal linking*, where free flow of information is possible, or this would entail that there would be constant confusion between information in the child's file and that in the vicarious file. The notion introduced, then, is that of *vertical* linking, i. e. linking between two files that does not allow constant flow, but that connects the files in such a way that they, once again, have sameness of reference. This poses the interesting question of how the information can "travel" from one file to the other, and a mechanism that is (quickly) mentioned in Perner and Leahy (2016) is that of ascension which, as will be seen, is intro-

duced by Gordon (2007). This approach to false belief, it is argued, is in line with the simulationist intuition that understanding others entails understanding their perspective, and that to bear on their perspectives we need to employ our own representational resources and reasoning abilities and relate them to other agents (Perner and Leahy, 2016, p.500). To sum up, while alternative naming is a process that is made possible by horizontal linking, vertical linking (i. e. linking where the flow of information is not free, but constrained) is what explains false belief performance. The linking operation is necessary, in both cases, because the two files, while anchored to the same thing, have to be connected (one might say, because the ability to link them is equivalent to the ability to treat them as referring to the same thing). The "ascent" operation will be described in the section dedicated to Gordon, and it is very quickly mentioned in Perner and Leahy (2016), but the idea would be that information can travel from one file to the other by way of marking the content as *believed* by the other agent. This seems to *imply* more than *explain* false belief reasoning, of course, but the story does rely on the fact that learning to perform vertical linking is indeed learning to mentalize at some level.

The whole argument, then, is based on the fact that linking different mental files, which is the basis for both the alternative naming task and the false belief task, is a core ability which develops in children around 4 year of age, and that this ability develops over the third year of life. To better understand the theory, it is useful to consider another task used in a series of studies (Apperly and Robinson, 2001, 2003; Sprung et al., 2007; Perner/Haemer/Leahy, 2015), the Heinz task. In the task, children are shown a rubber die, and then they are shown that the die is also an eraser. The die is then put in a box, and a normal eraser is put in a different box. The puppet Heinz then enters the scene, and children are asked whether Heinz knows that the die is also an eraser; subsequently, children are asked where Heinz will look for an eraser. Children at the age when they pass the first-order belief task are able to reply correctly to the first question (that Heinz does not know that the die is an eraser), but they are not able to reply correctly to the second question, as they point at the location of the die/eraser as often as they point at the location of the normal eraser. According to the studies in (Sprung et al., 2007; Perner/Haemer/Leahy, 2015), this happens at the same developmental age when children are able to pass first-order FBTs but not, crucially, second-order FBTs (at the age, in other words, at which children are not able to reply to questions such as "What will Max say if we ask whether he knows where the chocolate is?"). The explanation for this phenomenon, then, is that the child is able to link vicarious files (as established, this is the prerequisite for generating the right answer in false belief tasks); however, the child generates *too many* vicarious files.

This excessive deployment is due to several factors; on the one hand, the child needs to develop a sensitivity to *which* conceptual perspectives are available to Heinz, and this is, it is argued, a matter of second-order belief attribution; they have to be able to disentangle the fact, on the one hand, that the die is visually available to Heinz, but on the other, that the fact that the die is also an eraser is not visually available for Heinz.

In conclusion, managing mental files explains, according to the author, both first-order and second-order belief reasoning.[41]

It should be clear that the mental files approach only tangentially touches the issue of language, but still makes a prediction about how different labels invite taking different perspectives on different objects, which is in line with the data presented in previous chapters about how general language abilities, vocabulary, and interactions with siblings and caregivers all facilitate false belief reasoning. The theory, on the other hand, makes no specific predictions about other aspects of language; it does not address the issue of pragmatic components, nor does it provide an idea of how the ability to link different mental files (and to create vicarious files at all) is related to a larger developmental story. As the authors recognize, it does not provide a story about *how* mentalizing tasks are solved, rather it gives a common description of phenomena that might entail the same mechanism, i. e. alternative naming and mentalizing tasks. What is lacking is not only a developmental paradigm that goes behind the mental files explanation and its metaphorical taste, but also a description of *how* and *why* this ability to link files, horizontally or vertically, develops.

To address another issue with the story, it is useful to refer to the idea of ascent routines mentioned in Perner and Leahy (2016). In the mental files framework, ascent routines might play the role of marking information present in vicarious mental files. The notion of ascent routines will thus explained in the next section.

41 Note that in Huemer et al. (2018) interesting results concerning the kinds of limits that can be imposed on the excessive deployment of mental files are discussed. The authors argue that the way the information is presented to the child, in a temporal or verbal manner, can prevent excessive deployment. However, these are fundamentally practical details concerned with adapting mental files theory's assumptions and explaining this excessive deployment in higher-order mentalizing and are not really relevant for the topic at hand.

5.10 Ascent routines: Gordon's (2007) proposal

The following proposal has to be understood in light of early simulation theory; as a matter of fact, Gordon was historically one of the first proponents, along with Heal (Heal, 1986) and Goldman (Goldman, 1986), of a simulationist account of Theory of Mind. The reason why his account is here reported is that, compared to rival conceptions (especially Goldman's), it attributes to language an exceptionally important role.

In order to understand the account, it is necessary to briefly mention the problem of self-ascription. Nichols and Stich (Nichols and Stich, 2003, 2000), as well as Goldman and Gordon (Goldman, 1986, 2006; Gordon, 1986, 1992, 2005) have for a long time defended opposing views concerned with whether or not the ascription of mental states to oneself is preliminary and/or privileged compared to the ascription of mental states to others, i.e. mentalizing. While, as mentioned already, the debate deserves an entire book, it is worth understanding that the question about the role of language cross-cuts that between TT and ST, and this is especially evident in the case of the differences between Goldman and Gordon. While Goldman's view is based on introspection, and hence privileged access to one's own mental states, Gordon's view is radically different; and introspection does not have the same role.

Gordon's view sees mindreading abilities as language-dependent, and especially dependent on ascent routines:

> [...] the way in which adults ordinarily determine whether or not they believe that p is simply to ask themselves the question whether or not p." (Gordon, 1986, p.16)

In this sense, our access to our mental states is linguistic in form and, Gordon further claims, language is essential to the employment of mental state concepts. An ascent routine is a mechanism for self-ascribing a propositional attitude by redeploying what Gordon calls a "lower level" utterance: the idea stems from the intuition that, to report our beliefs on for example the weather, what we have to do is to "go down" a linguistic level, and not by appealing to a theory of my own mental states nor, as Gordon insists, on the basis of a feeling or mark that denotes my knowledge state. The routine, however, is thought to be "ascent" in the sense that, to formulate ascriptions of mental states, one has to go from the "lower" level assertion "The weather is bad" to the level "I believe that the weather is bad". This kind of mechanism, which is openly linguistic, is what guides self-ascription. While expressing an attitude is not the same thing as self-ascribing an attitude, Gordon notes, the two operations can (and are) carried out in the same linguistic form. While the main focus is on belief ascriptions,

Gordon claims that the same can be said about any attitude. In this sense, getting to know your own mental states is not a privileged routine and it does, crucially, only happen in the presence of a language.

While the core of the idea is evident from the characterization of ascent routines for *self-ascription* given so far, it also holds for ascription to others, since the first step in mentalizing development is, for Gordon, a linguistic matter:

> If I am right, to attribute a belief to another person is to make an assertion, to state something as a fact, within the context of practical simulation. Acquisition of the capacity to attribute beliefs is acquisition of the capacity to make assertions in such a context. (Gordon, 1986, p.170)

Simulation is here intended not in mirror neuron terms, but as simulation of practical reasoning: Gordon assumes, and so does Goldman, that one's own practical reasoning processes and abilities are employed in the ascription of mental states in a hypothetical scenario, which has its developmental origins in pretend play. This applies to prediction of one's own behavior, which implies simulating a different situation than the one one's in, and prediction of other's behavior as well, engaging in what is ultimately hypothetical practical reasoning. As mentioned earlier, this kind of approach stems from the necessity to reject one of the fundamental assumptions of theory-theory accounts of mentalizing, i.e. that it implies a theory-like approach to other people's behavior. However, in the case of Gordon, the operation of ascription is fundamentally linguistic.

Going back to the mental files proposal for a moment, it is worth noticing that including ascent routines to mark information in vicarious mental files, while connecting the account presented by Perner with data regarding the importance of syntactic information, would present other problems. If it is the case that the vicarious files get characterized by propositional ascent routines, and this is what allows for the marking of the file as belonging to another agent, it looks like the ascent routine, and so the ability to inscribe the descriptive content of the file in a propositional attitude structure, would already do the heavy lifting necessary for mentalizing and understanding the false belief task, making the mental files story almost superfluous. In other words, if the child is already able to mark the descriptive content of the mental file as belonging to another agent, it seems like the ascription of a mental state or perspective is possible, and no linking operation is necessary. However, there might be some less obvious, more useful way to implement the ascent routine in the mental files story.

Let us focus on the ascent routine proposal once again. Language plays a pivotal role, for Gordon, in the possibility of ascribing mental states to oneself and others, and this is fundamentally reflected in two facts: on the one hand, practical reasoning in the form he describes is supposed to be carried out in linguistic form; on the other hand, the ascent routine mechanism is a mechanism that goes from utterances to utterances, i.e. fundamentally in natural language. Moreover, Gordon argues that not only does linguistic expression of selfattribution precede the mastery of mental state concepts, but that the use of verbal expressions is what fundamentally could provide the bootstrapping necessary to achieve attribution. The first uses of mental state verbs, it is argued, do not indicate a use of mental state concepts because they are merely non-assertive, conversational uses of mental state verbs, as explained in chapter 3: however, they introduce the child to the use of these expressions, which are then central for developing mentalizing abilities themselves. Since the abilities in mentalizing are those of practical reasoning in linguistic form, this makes mentalizing ability fundamentally just an emergent result of the combination of other general domain mechanisms with the ability to express propositions.

Gordon (2007), then, claims that it is through the expression of sentences like "I (mental state verb) that p" that one gets to the mastery of ascription: while this is only one developmental stage, Gordon argues that it is through the expression of propositional attitudes, which is *followed* by mastery of the deriving mental state concepts, that mentalizing is possible. While it is not specified how this mechanism works, i.e. how one passes from the linguistic mastery to the conceptual mastery, Gordon suggests that the use of the linguistic routines in practical reasoning brings upon the formation of mental state concepts.

Gordon's predictions, while in line with lots of the data presented in the previous chapter, are clearly very strong: conceiving of mentalizing as a fundamentally linguistic ability comes at a price, i.e. discarding anything preceding linguistic competence as non-related to mentalizing. This is related to another concern, which is not developmental: the form of mentalizing described by Gordon appears to be over-intellectualized in assuming that most of our mindreading abilities (and even self-ascription) happens in an explicit linguistic form. This obviously entails problems in respect to the relatively unimpaired or only delayed performance for SLI patients and aphasic patients, and does leave the question open of how this highly explicit, highly sophisticated activity relates to other skills related to intention reading and goal-directed action. In the next section, an equally bold proposal will be presented that goes in a somewhat similar direction, attributing to language a big part of what mentalizing activities entail, albeit in another form.

5.11 Folk-psychology routines

The last proposal that will be analysed completes the picture by showing that proposals on the interaction between language and mentalizing skills come from theory-theory approaches, simulation approaches, and finally from approaches that reject both TT and ST.

Hutto's take on folk psychology can be best understood in light of his commitment to an embodied approach to cognition, i.e. the idea that the best way to understand the mind is not via computational assumptions like the one traditionally endorsed, but rather by conceiving of the mind as embodied and embedded in social practices that expand the borders of cognition to cover not only individual psychology but also interaction with the environment, including other agents. This is the reason why Hutto (2008b) proposes a Narrative Practice Hypothesis, (NPH), that embeds social cognition in the narratives used in everyday practice (in this sense, very similarly to Nelson) and that focuses on *second-person* mindreading, i.e. the fact that most of our mentalizing experiences happen *online*, when interacting with another agent, and not while taking a "spectatorial stance" towards mentalizing. The standard approaches to mentalizing, it is argued, focus too much on either artificial situations like third-person attribution as a spectator or the standard FBT, or on the self-attribution of mental states, as is the case for Gordon and the debate surrounding Goldman's view. However, mentalizing skills are especially in play when we encounter and interact with people in specific situations and during communication, and not just as silent viewers of situations, or when alone.

Folk psychology is, in this view, a specific kind of narrative practice; and it is through engaging in narrative practices of a particular kind, where agents' behavior is explained in light of their characteristics and in light of external circumstances, that mentalizing abilities develop. In other words, any narrative, whether fictional or based on facts, whether heard or read in books and other stories, provides substantial ways to "make sense" of the behavior of agents. While a standard ST or TT approach assumes that the skills that are necessary to engage with these narratives are already mentalizing skills, Hutto's claim is that the socially embedded nature of these narratives is what provides the skills necessary for the practices of mentalizing; in other words, first come the narratives, and later the practice of second-order mentalizing in dialogue, followed only then by the spectator-like mentalizing tested by standard accounts.

Hutto's view is best understood while holding in mind his view on propositional content: to be able to have content-involving propositional attitudes, one must be able to understand sentences, since sentences in natural language possess the semantic and syntactic characteristics that are necessary to represent

complex states of affairs, to evaluate them in truth-conditional terms, and to subject them to complex reasoning. To have contentful states, for Hutto, is to be able to master natural language. Infant's social cognition is to be interpreted not as content-handling, then, but as substantially embodied and grounded in non-theoretical, motor-based principles; what the infant does, before mastering language, is to employ its own perceptual and motor resources to understand the bodily expressions of other agents, and to engage in what is ultimately not mentalizing nor mindreading, but a lower form of understanding. Appealing to mirror neurons, Hutto claims that these forms of understanding others are essentially not content-handling, not representational, and not concerned with propositional attitudes, ascription or anything similar. In Hutto (2017), the focus is on the representationalist stance taken by the majority of Theory of Mind positions and authors, which according to Hutto makes several methodological mistakes in assuming 1) the possibility of non-conceptual, content-handling thought and 2) the necessity of appealing to representations in many low-level cognitive processes. While this kind of debate is essential for a full-blown theory of cognition, it is not the focus of this chapter, nor of this book. However, Hutto makes predictions about folk psychology and language use that are (partially) independent on this stance on representations.

The NPH assumes that, in engaging with narratives, children become sensitive to generalities and variables that are employed in folk psychology; Hutto is, however, clear in affirming that these generalities are not to be understood as rules and principles, at least not as a mentally represented set of laws that causally explain folk psychology (Hutto and Kirchhoff, 2015, p.8). In an embodied view of cognition, skills can develop over time, becoming increasingly refined and sensitive to context rules, and this applies to physical skills like those employed in sports, for example, and to cognitive skills as well, including those involved in mentalizing. In other words, and somewhat simplifying the embodied view on cognition, folk psychology is an instance of *know-how*. What defines a folk-psychology narrative is that it provides details on the psychological states and attitudes of the people involved in the story: "narrative" denotes for Hutto the mode of presentation of the story line, and in this sense a narrative is a linguistic and representational artifact, a result of the human practice of story-telling in the particular format that is folk-psychological practices. Importantly, engaging with these practices does not mean acquiring a new set of representational rules, as the only representations involved are those of the narratives themselves, which, being linguistic, are bearers of propositional content.

Note that Hutto emphasizes how the NPH is to be embedded in the more general approach that can be named the Scaffolded Mind Hypothesis, i.e. the idea that the evolution of the mind depends at least in part on social-cognitive

scaffolding, and hence that cognitive resources evolve as a result of environmental resources. While the enactive stance taken by Hutto is very salient in his framework, note that a view of cognition as fundamentally scaffolded by social interaction and linguistic input is likewise supported by Vygotskyan approaches (Vygotsky, 1962), by Carey's bootstrapping idea (Carey, 2009) and by Clark's view (Clark, 2013). In any case, as far as folk psychology is concerned, it is argued by Hutto that humans construct and shape their niche, and the environmental constraints that result from the niche, at the same time, influence further cognitive development. This dynamic interaction explains how folk psychology has evolved to be a significant part of our interactions, since the construction of narratives that involve it has been shaping the cognitive development of children for generations[42].

This approach should not be mistaken, Hutto points out, with the idea that cognitive processes *internalize* the narratives to develop new representational means: most folk-psychological practices remain instances of know-how which do not need to be seen as internalized rules. However, there is at the same time a strong emphasis in Hutto (2008b) on the content-bearing power of language, which makes folk psychology, not based on representational rules, but embedded in a fundamentally contentful linguistic pattern. In this sense, mindreading is a practice that involves no theory-making and no speculative inference (Hutto, 2011); in cases in which the routines are not working, on the other hand, it is possible that we do employ theory-like formulas, but those are not to be understood as a ToM, and instead are going to rely on general behavioral rules based on what we know about humans in general, for example. "Theory-like" mindreading is, then, a limiting case. Cases of intention reading are characterized by Hutto as being grounded mostly in intentionally directed responsiveness (Hutto, 2009, 2008a) to the intentional attitudes of others. At the same time, the process of mastering mentalizing narratives is thought to be much longer than the first 4 years of life, being an extended complex process that comes together with the growth of linguistic abilities and narrative skills, which extends obviously longer. In this sense, NPH relies on language, heavily, because it is only with the development of linguistic abilities that mentalizing abilities develop; this is also reflected in the fact that learning to use narratives implies learning how to use mental state terms: "[...] children do not learn that beliefs and desires exist they learn how to make sense of others using such terms." (Hutto, 2009, p.34).

42 This idea of niche construction derives from philosophy of biology, and Hutto particularly relies on Sterelny's work (Sterelny, 2006).

It is of fundamental importance to specify that an assumption made by NPH is that concepts of mental states are already in place before the lifting of social narratives is made:

> To avoid misunderstanding, it is important to stress that the NPH supposes that FP narratives do crucially important but nonetheless limited work. They are not responsible for introducing an understanding of mental concepts, such as desire and belief for the first time, rather, being complex linguistic representations of particular events, they put on show how these attitudes can integrate with one another (and also how they fit with other mental states and stand with respect to other contextual factors). The NPH assumes that kids already have a practical grasp on what it is to have a desire or belief before learning how to integrate their discrete understanding of these concepts in making sense of actions in terms of reasons. FP narratives enable this by showing how these core attitudes and other mental states behave in situ. (Hutto, 2008c, p.178)

This is, in my opinion, the most confusing part of the claim: while it is clear that the objective is not to make narratives the source of children's conceptual understanding of beliefs, it is not clear where this practical conceptual understanding comes from. While one option would be that of innate concepts of mental states, this option would be odd in an enactivist account of cognition like that assumed by Hutto. At the same time, it is not clear what the function of mental state concepts would be, since low-level forms of mindreading (or mind-minding, as Hutto terms it) are not considered to employ such conceptual representations, but are cashed out as enactive practical engagement with social narratives. At the same time, "high-level" mentalizing is supposedly only in place when narrative practices are actually learned, since the NPH is what really provides a picture of how mental state talking relates to behavior in the sense of explanation and description: this leaves us with the question of what exactly is the practical understanding Hutto attributes to infants *before* they engage in narrative practice, and, crucially, why this should be in place before the supposedly best application of these concepts, which is explanation of behavior, which only comes with social narrative practice. In other words, not only does assuming that some practical understanding of mental states is in place before narratives are present seem to be pointless, it also seems to undermine the role of narratives themselves, and seems not to be in line with the idea that no mental state ascription is present in lower forms of mindreading. One way to save the hypothesis, at this point, would be to maintain that narratives do provide that basic understanding of mental state concepts for the first time; this, however, might imply that there is indeed some internalization process at play, and that abstract generalizations can be derived from the practice of mentalizing or narrative psychology. This is however a claim that Hutto wants to avoid at all costs,

which makes this way unavailable for a proponent of NPH. In the next chapter, the idea of narratives as central for the development of mentalizing will be reconsidered in light of these claims, in an attempt to make clearer what the relation between narratives, conceptual development and linguistic information could be.

5.12 Concluding remarks

This chapter has analyzed several theories in light of the evidence presented in previous chapters. While many of these theories accommodate part of the data available, none of them reaches a level of specification sufficient to inscribe it in a more general view of cognition and at the same time is fully compatible with the data presented in previous chapters.

While addressing specific data regarding the role of syntax, the syntactic bootstrapping hypothesis (de Villiers, 2005) does not hold against some of the objections raised by its opponents. It fails to identify purely syntactic means of representations that explain cross-linguistic data, and leaves a lot to desire in terms of providing a mechanism that can be generalized and inscribed in a coherent developmental story of mentalizing. Frameworks like Garfield's (Garfield et al., 2001), while going in the right direction, remain underspecified in many ways, failing to describe in detail mechanisms at the core of the development of the interrelation between social cognition and linguistic acquisition. Something very similar can be said about the account proposed by Nelson (2005), where specific predictions are somehow lacking, even if the account potentially is well connected with data presented in the previous chapter on the importance of narratives and social interaction. The proposal in Montgomery (2005) is also somewhat underspecified and not situated with respect to a more general picture. Gauker and Van Cleave's (Van Cleave and Gauker, 2013) hypothesis falls short of explaining mentalizing skills in the details of the account, which presupposes a lot of specific skills, including possibly hypothetical reasoning, and relies on an unclear distinction between pragmatic roles for mental state verbs. While Bermúdez (2009) and Gordon (2007) make reasonable (perhaps trivial) claims for the essential role of language in explicit, propositional mindreading, it remains to be determined how this fits into a larger picture, and how non-essentially linguistic social skills are influenced by language acquisition in this picture. A promising story is that proposed by Perner (Perner/Haemer/Leahy, 2015; Perner and Leahy, 2016) in terms of mental files; however, as underlined, it does not yet provide some of the details that would make it the preferable account, while still leaving many questions open, both empirically and theoretical-

ly. Finally, Hutto (2008b) emphasis on narratives, albeit essential, does not solve the more general problem of assessing the interplay between conceptual development and linguistic acquisition.

There is a lot to learn and take away from each of the accounts presented here, however, as they present an array of solutions that nicely point out some of the most important factors that emerged in the empirical review in the previous chapter as well; while a role for linguistic input has to be recognized, it cannot be considered in isolation, but only as part of a story that considers social interactions and the development of social skills as well. Nelson's account and Baldwin and Saylor's framework provide valuable insight regarding specific factors playing a relevant role in folk psychology: the power of labels as expanding comparing abilities on the one hand, and the importance of mentalizing as a public, social activity on the other. The same holds for Hutto's emphasis on folk psychology. Nelson's idea, namely, that we consider language as providing internal representational means, distinguishes her proposal from the practical stance taken by Gauker and Van Clemens. Both accounts provide good reason not to forget the relevance of pragmatic development, despite their explanatory weaknesses highlighted in the respective sections. The syntactic bootstrapping hypothesis efficiently addresses a lot of data regarding the role of language and specific sentential complements, which has to be taken into consideration. Finally, Montgomery's proposal that we regard the first instances of mental state verbs as fundamentally not ostension-based, but shift the focus onto a kind of linguistic and non-linguistic behavior, is a valuable one, which will be mentioned again in the next chapter. In what follows, I will provide a story for the role of mentalizing that borrows many components from the theories presented so far, including mechanisms like structural alignment; I will also formulate an account that combines the importance of folk-psychology narratives and of linguistic skills development. As will be seen, the story will not be incompatible with some of the theories presented here; on the contrary, it will complement some of their features.

Chapter 6:
LALAS: Language Associations Labels Acquired Schemata

In this chapter, I will build on the analysis of the previous empirical and theoretical reviews and on Berio (2020a) to provide an account of how the acquisition of mental state terms, the constructions they appear in, and their use in conversation contribute to the formation of mentalizing mechanisms. The chapter will be structured as follows.

Firstly, I will present a summary of the analysis conducted so far, in order to provide the general framework for the proposal. Secondly, I will present LALAS (Language Associations Labels Acquired Schemata). LALAS is meant to be the description of a mechanism according to which language promotes the acquisition of schemata that relates agents and descriptions of situations. These schemata are used when solving false belief tasks, and similar folk-psychology situations. In my account, developed also in Berio (2020a), language is considered as a *communicative tool* first and foremost, and taken into consideration not only as a set of rules and structures, but also as a dynamic interactive device. I will argue in the rest of the chapter that this approach overcomes many of the limits that emerge from other theories, and that taking into account the communicative context and the narratives in which the linguistic input is embedded is essential to providing a comprehensive story about the contribution of language to mentalizing. After presenting the basic assumptions of LALAS, I will dedicate part of this chapter to situating the account within a larger picture of mentalizing, relying on two different double-mechanism frameworks of mindreading (De Bruin and Newen, 2014; Apperly, 2011); I will argue that a double-mechanism approach is not only more suitable for solving what has been called the "developmental paradox" (De Bruin and Newen, 2012) of mentalizing, but it is also most likely to adequately incorporate an account of how linguistic acquisition aids mentalizing. In this way, I will situate LALAS within a more general account of mindreading, which provides a specific advantage compared to many of the theories presented in the previous chapter. Finally, I will sketch how LALAS fits with the theories presented in chapter 2, giving an idea of how LALAS is cognitively plausible and can fit within a possible account of how language interacts with other cognitive resources. In the concluding chapter, I will draw conclusions and outline some possible further research aimed at gaining further empirical evidence.

https://doi.org/10.1515/9783110748475-007

6.1 Developing LALAS

6.1.1 The state of the art so far and some assumptions

The previous chapter's focus was theories that directly address the issue of the relation between mentalizing abilities, language acquisition, and development. As was pointed out at the end of the analysis in chapter 5, none of the pictures look complete, as they either only address a part of the issue, as is the case for the syntactic bootstrapping hypothesis, or they make more general underspecified claims, as is the case for Garfield et al. (2001). In general, there is fragmentation: specific accounts focus on specific aspects of language as providing a significant aid for mentalizing, but tend to overlook other factors. It is true for example of the hypothesis by Van Cleave and Gauker (2013), where the pragmatic aspects involved in mental state terms are discussed in relation to ToM, but other factors are mostly overlooked; and it is also the case with the hypothesis by de Villiers (2005), where the focus on the syntactic component tends to overshadow the contribution of social interaction. The social component is stressed by Nelson (2005) and Hutto (2008b), but no more concrete picture is given regarding acquisition of specific linguistic means.

In general, the evidence presented in chapter 4 is only partially accounted for by the theories described; even more importantly, most of the theories do not take into account how the role of language is integrated in the more general developmental course of mentalizing abilities. In what follows, I will address these issues, pointing to a possible solution to this problem.

Before diving into the account, a few words on how to collocate this proposal. As will soon be clear, my account pertains specifically to how specific linguistic input and specific mechanisms can aid false belief reasoning. In this sense, while the account is one of language acquisition and cognitive development, it has a rather restricted scope; in other words, I don't propose here a full-blown view of socio-cognitive development. At the same time, as will become clear to the reader, I do make some basic assumptions about cognitive development: these have to be taken as preliminaries to my view which I adopt from other research, and not points that I defend here. For example, like many other scholars, like Heyes (2018), but also the already cited Hutto (2008b)), I do assume that cumulative culture can have a relevant role in shaping cognitive development. I do not make assumptions here of the kind Heyes makes when discussing mindreading, i. e. I do not argue that it is a "cultural" cognitive gadget. I do however assume that cultural input can be fundamental in shaping children's abilities. At the same time, as will emerge especially in the last part of the chapter, I assume the child is well equipped with a range of cognitive abilities when 3/4 years of

age. For example, as stressed at several points in this book, I presuppose fundamental abilities in developing skills of joint attention (Tomasello, 2014) are already in place when the described mechanisms facilitating false belief reasoning intervene. Moreover, I assume that by the age of 3, children are endowed with the skills of object permanence and individuation such as those described in Rakoczy (2017), and that they are definitely endowed with the skills to understand perception and intentional action that Proft et al. (2019) attributes to infants. While I do not commit to theories regarding the onset of these abilities, I do consider them in place at the age that is relevant for LALAS, namely at around 4 years old.

6.1.2 Towards embedding language in cultural practices

One thing that characterizes approaches like de Villiers' is that language is considered to be, mostly, a set of rules, often assumed to be implicitly known and internalized. It is not unusual, within the generativist tradition, to assume "implicit markers" of different kinds, like the syntactic hidden marker proposed in the context of the syntactic bootstrapping hypothesis. There is a sense in which approaches like Nelson's, Hutto's, Van Cleave's, and Gauker's, while being different in many theoretical respects, all point to a different understanding of language, according to which language is a *tool* for communication, something which emerges from communicative practices among people, in a less abstract way. The shift is from an internalized, cognitive approach to language to a social view of language interaction.

Note that, obviously, studying language as a set of rules and patterns is *not* incompatible with studying the social, interactive dimension of language. What is advocated here is a shift in perspective, in the sense that, as is underlined by both the empirical results that stress the role of pragmatic development and interaction with siblings and caregivers, and by the theoretical analysis carried out by Nelson (2005), Hutto (2009), and Van Cleave and Gauker (2013) among others, both language and mentalizing are best understood as communal, social practices whose processes develop *online* in specific, concrete situations. While testing in a controlled environment and within very specified task-designs is essential for observing psychological phenomena in a rigorous way, this cannot lead us to overlook the fact that both explaining other people's behavior in terms of mental states and acquiring one's own language are processes that happen in the context of interaction with other agents, and as part of the cultural upbringing we experience as humans. In this sense, the idea of a "Community of Minds" (Nelson, 2005) is not far-fetched; it reminds us that the development of the skills

of mentalizing are part of how members of a given community interact with each other and explain each other's behavior. At the same time, Van Cleave and Gauker (2013) focus exclusively on pragmatics, and as a result exclusively take into consideration pragmatic use of mental state verbs, losing sight of the underlying question of what it is about them that makes folk-psychology possible or better.

Take as an example the debate regarding the sentential complement constructions between Perner and de Villiers (Perner et al., 2005; de Villiers, 2005): the existence of the German finite complementation for sentences with "wollen" presents an extreme danger for the syntactic bootstrapping hypothesis only if one considers the syntactic input to provide the representation itself. If one, for example, accepts the idea that semantic and pragmatic information interact in a significant way with the syntactic structure, and that it is the interface of these different informational sources that co-opts the representational resources that are necessary in mentalizing, one does not have to choose between discarding the role of syntax as a whole and relying exclusively on grammatical rules for explaining the role of language in mentalizing. More on this will be said when presenting LALAS; for the moment, however, it is necessary to stress how a wider look at language acquisition should take into consideration both syntactic components and the fact that they are socio-culturally embedded. A construction like "Mum wants you to pick up your toys", or the German equivalent, is normally heard in the context of a request or, in a certain sense, an order: the relevant communicative situation for the child is that in which somebody asks them to do something concretely, to apply a concrete behavior to a concrete situation. As observed by Van Cleave and Gauker (Van Cleave and Gauker, 2013), this is ultimately among the most common uses of desire expression that the child is exposed to: in this sense, the difference between *realis* and *irrealis* complements is surely evident, and de Villiers (de Villiers, 2005) has good reasons to point it out. However, the difference cannot be reduced to a purely syntactic one, or one loses this pragmatic and communicative difference. LALAS is designed to account for how syntactic information interacts with communicative situations in a way that partially reconciles these intuitions with the syntactic bootstrapping hypothesis.

6.1.3 The core idea: Language Associations Labels And Schemata

The main starting point for LALAS is, as stated, a view of language as a tool for navigating the world in concrete situations and a desire to account for the role of language learning and its impact on mentalizing in a way that successfully integrates the role of syntactic information, as suggested by de Villiers and Tomasel-

lo, with attention to the socio-communicative dimension of language learning, including the context of story-telling. LALAS is a proposal that is limited to the case of language learning and tasks like FBT. However, it has the potential to be expanded to other mentalizing tasks, and it provides a mechanism for associating agents and states of affairs on a more abstract level.

The idea, already partially developed in Berio (2020a), is to take seriously Montgomery's (2005) suggestion, that the pragmatic roles that are implied in mental state constructions are also a relevant part of the story. In this sense, as argued in Berio (2020a), I claim that learning the meaning of mental state terms is a matter, firstly, of forming abstraction schemata relating agents and descriptions, and only secondly of forming mental state *concepts*. In a nutshell, the LALAS proposal is that the role of language is to provide schemata that see a structural link between an agent and the description of a state of things, and the ability to label different attitudes denoting the relation between the two.

Let us go through the main claims of LALAS, which will be later spelled out:

- Performance in false belief tasks can be facilitated by being able to associate an agent and a state of affairs in a general enough way and being able to use this information in culturally meaningful ways.
- **Syntax** helps this process because:
  - It provides structures that systematically describe a relation between an agent and a state of affairs;
  - Emerging linguistic (semantic, syntactic, comprehension) skills play a pivotal role in individuating the common pattern in the way actions are described (and reasons for actions are given);
  - Individuating this pattern is facilitated because language promotes structural alignment, i.e. the individuation of similarities when comparing perceptually different situations;
  - Syntactic information plays a role in the acquisition of schemata, as it provides a way for the child to form the schema, i.e. to find an association pattern. However, syntactic information is *not* recruited *online* to solve FBTs.
- Upon individuation of this pattern used in explanation, the child forms a relatively abstract representation of a schema relating an agent and a description of a state; possessing this schema makes the storing of new associations easier, helps to locate new associations in given stimuli, and helps to retrieve them for new tasks.
- These schemata are **culturally embedded** in narratives (Berio, 2020a), which:
  - Provide a context for the relation between agents and states of affairs: descriptions of this relation appear often in Western culture in stories

and in narratives concerned with people's behavior and their explanation;
  - Train the child in the use of this information for explaining and predicting behavior.
- As far as **concepts** of mental states are concerned:
  - Labeling the relation between agents and states of affairs in different ways provides a pathway for distinguishing between different mental states;
  - *Concepts* of mental states in a full-blown sense likely emerge *after* sophisticated linguistic and cultural information is gathered for them;
  - Attribution of intentions, understanding goal-directed behavior, and first-level false belief reasoning do *not* require a sophisticated understanding of mental state concepts;
  - The ability to pass the false belief task is built on previously in place abilities: language does not foster a cognitive revolution, but enhances other existing mechanisms.

The above points represent the core argument behind the proposed model, and they will be developed in the rest of the chapter. The resulting argument will show that a role for language in mentalizing can be accommodated in a framework that does not make strong assumptions about logical form and that does not rely on syntactic information alone. The resultant view gives a more general account of how language, considered as a cognitive tool, on the one hand, and as a means of social interaction, on the other, enhances mentalizing skills.

6.1.4 Schemata: a brief history

Clearly, the idea of involving schemata in cognitive development is far from new. There is a sense in which what I mean by schema is just the result of an abstraction process over regularities picked up in language, which makes it a rather different idea from that which has been proposed in the literature. However, there are definitely connections, and it is worth keeping the idea of schema and schemata in other fields of psychology in mind, since the spirit is partially the same. While, as it will been seen, they are partially invoked by Apperly (2011) as well, the reader has probably realized that the idea of schemata as instruments for organizing knowledge can be traced back to Piaget (1936, 1952). Piaget described schemata as repeatable action sequences or a series of linked mental representations used by the learning child to respond to the environment. Most schemata develop with learning, and they fundamentally increase in complexity and num-

ber when the child goes through developmental stages. Crucially, new schemata evolve in response to environmental changes that cannot be understood through the current schemata, and are actively *built* by the child. Piaget's view also famously entailed a conflict with Noam Chomsky[43] regarding what is *innate* about language construction: if on the one hand Piaget's view contrasted with Vygotsky's insistence on the role of language and social interaction, the disagreement with Chomsky was based on whether or not linguistic principles were supposedly innate (Chomsky, 1980a; Cattell, 2007). As I argue in Berio (2020a) and as it will be clear in the rest of the chapter, the idea of schemata proposed in this case is slightly different because it implies generalization over language patterns, so the concept of schemata does not have to be considered strictly Piagetian. Also, as has already been stated and will be made clear again, the perspective is rather in line with Vygotsky's idea, that language acquisition and cognitive development develop closely, which is also not completely in line with the Piagetian approach.

Schemata have been used in other fields as well, starting from the ideas in Schank (1982); Schank and Abelson (1977) and Gilbert (1998) in social psychology, where the notion is that situational schemata are formed to deal with specific situations, like shopping in a store, that require one to operate according to a large variety of variables that can include large quantities of data. Action schemata can be used in dealing with complex actions that require complex strings of motor commands as well, as is the case in Norman and Shallice (1986), for instance. In general, what these different ideas of schemata from different fields of research have in common is that information can be stored in a rather abstract representational format that captures regularities. In the second part of this chapter, when analysing Apperly (2011) and the idea of the application of schemata to mentalizing therein, I will stress the ways in which my account fits in this tradition.

In the next sections, I will go into the details of how syntax, on the one hand, and narratives, on the other, play a role in my account. This will expand significantly my previous elaborations of this account (Berio, 2020a). First, I will dedicate time to explaining how syntactic information has the function of helping to organize existent information and to make it more reliable for sharing and communication. Subsequently, I will focus on how narratives provide the essential context of use for said information.

43 Even maintaining some common traits, see Tomasello (1996).

6.1.5 Syntax: patterns and generalizations

6.1.5.1 On the role of language to individuate patterns: structural alignment

As was explained in the previous chapter, Baldwin and Saylor (2012) propose that language aids *structural alignment*. In particular, the structural mapping hypothesis can be summarized in this way, following Gentner et al. (2011):

> Comparison entails a structural alignment process that promotes a focus on common relational structure. (Gentner et al., 2011, p.1175)

Language, then, gives means for structural alignment: it provides a way to unify different representations that have something more in common, i.e. the way they are described. As mentioned in the previous chapter, there seems to be evidence that using the same labels for different objects elicits similar interactions with the objects in young children (Desjardins and Baldwin, 1992); but also, as was already mentioned in chapter 2, there is evidence that language can be a valuable resource in categorization. Recall in this case the experiment carried out by Lupyan et al. (2007), where labelling "aliens" made it easier to classify them as either dangerous or not dangerous. In this case, the use of labels *not* directly linked to the classification purpose facilitated a sorting task. More to the point, however, *relational* language can be of fundamental importance. Consider the study reported in Gentner and Medina (1998). Children had to correctly guess the location of a sticker, which was to be found underneath an object included in the set (a triad) of objects they had. An experimenter had a triad of objects as well and would place one sticker underneath one of the objects; this would serve as a clue for the child to identify the sticker. The correct solution was based on relational similarities and not perceptual similarities: for example, if the experimenter put the sticker underneath the biggest pot in his triad, the child had to look underneath the biggest pot in her own triad, independently on whether or not the size of her object matched the size of the experimenter's object. (It was possible that the exact sized pot was present in the child's set too, but not being the *biggest*, it would not be the correct answer.)

3-year-olds improved their performance significantly, up to a 5-year-old level, when they were given relational language, i.e. when labels like "daddy", "mummy", and "baby" were used for the objects. In this way, using *relational language* facilitated performance. More recently, the impact of learning relational language for the formation of *relational categories* has been tested by Gentner et al. (2011): in this case, children had to form pairs of cards playing what was described by the experimenteer as a "learning meaning of funny words" game. They were initially given two cards with two different figures, e.g. a watermelon

and a knife. Subsequently, they were given an additional card, representing for example a piece of paper, and three cards to choose from to form a second, related pair: for example, a picture of a stack of pieces of paper (taxonomic relation), a picture of a pencil (thematic relation) and a picture of a pair of scissors (relational relation).

In this case, the correct answer implied choosing the pair of scissors. In one condition, children were given relational clues in form of language, e. g. "The first word is dax. The knife is the dax for the watermelon". In the condition without relational language, the first set of cards was introduced with sentences like "The knife goes with the watermelon.". Afterwards, the children had to reply to the questions "Now its your turn. Which one of these is the dax for the paper?" or "Now its your turn. Which one of these goes with the paper in the same way?". For 4- to 5-year-olds, relational language significantly improved the performance; for 6-year-olds, the "relational" choice was easier to make in *both* conditions, whereas 3-year-olds failed the task even when relational language was provided, thus indicating according to the authors that the ability to form relational categories evolves in time, and that other criteria are more salient for younger children. Interestingly, in the same study it was shown that training children with pairs of objects that were perceptually similar (watermelon and knife; apple and knife) and incrementally proceeding to more abstract relations (watermelon and knife, apple and knife, tree and ax, trunk and saw, paper and scissors) was facilitating 3 y.olds performance *when coupled with the use of relational language*. In this case, the experimenter was inviting comparison ("Now lets look at all of them. You see how these (gesturing across all four operators) are daxes for these; go with these in the same way (gesturing across all four entities")), thus giving evidence that experience with progressive alignment and the right kind of labels can significantly improve performance (Gentner et al., 2011, p.1182).

As seen, some of this data has been used in Baldwin and Saylor (2012) to argue that language aids the formation of the concept of mental states through structural alignment; the claim I am making is different because it does not entail that structural alignment is directly responsible for the formation of a *concept* of mental states. On the contrary, the idea is that what emerges first is an abstraction that should not be conceived as a full concept, rather as a fixed schema that allows for the easy storage, retrieval, and use of associations between state of affairs described by sentences and agents (Berio, 2020a). In this sense, structural alignment is thought to be a mechanism operating at a "lower level", not necessarily for the formation of conceptual structures. The idea is that using the same recurring linguistic means across a variety of situa-

tions will thus help to individuate the underlying pattern of association between agents and descriptions of state of affairs. So hearing sentences like:

(1) Galileo thinks the chocolate is still on the shelf.

(2) Sameer believes the toy train is in the living room.

(3) Nora thinks Ian is asleep.

(4) Kara knows that mum is in the kitchen.

will invite us to compare very different situations in which the same pattern, that of relating an agent to a description of a state of affairs, emerges.

This means, in a sense, reconciling the idea of *structural alignment* as fundamentally boosted by language without endorsing the ostension-paradigm that is maintained by Baldwin and Saylor (2012). Language which connects agents and descriptions, in this sense, does not directly find a referent for mental terms, but it helps in individuating a pattern. As will be explained later in the chapter, a fundamental part of this story is that how mental state verbs are used is also of fundamental importance, since the schemata are embedded in cultural practices.

6.1.5.2 LALAS and parsing

In this section, I will argue at some length that LALAS does not contradict prominent views regarding the acquisition of language parsing abilities, in order to give my proposal a concrete grounding. LALAS's view, that children can abstract schemata by being exposed to linguistic stimuli, is in line with some of the theoretical views that are present in the debate on syntactic and language acquisition, a relevant example being Tomasello's *usage*-based theory of syntactic acquisition. Acquiring a language, Tomasello argues (Tomasello, 2009b), includes fundamental skills, among them, firstly, *pattern finding* and secondly, the ability to use *schematization* and *analogy* in order to be able to abstract syntactic constructions starting from the concrete utterances one is exposed to. The theory has been proposed in open contrast to both associationist theories of syntax and generative syntax. According to Tomasello's take, children's ability to learn new syntactic constructions relies on the fact that they are able to generalize over concrete instances and work their way to abstraction schemata. This is a slow process that only terminates late in development, since children will firstly learn holophrases, then schemata that are still fundamentally concrete, and only later abstracted ones (Tomasello, 2009a). Tomasello himself recalls the idea of structural alignment for his argument, claiming that the ability to create analo-

gies and compare different input to find structural similarities is one of the elements that grants the child the ability to learn new syntactic structures. The position maintained by Tomasello is in open contrast to the classic "poverty of stimulus" argument (Chomsky, 1980b) according to which the stimuli in the environment are simply not enough for the child to ever arrive at a stage of competence with the grammar of their own language, unless one poses a Universal Grammar as a set of principles whose parameters get set thanks to experience with specific languages. This is a very general way to characterize a Chomskyan position, and there are many minimalist positions available[44]. However, what is important for current purposes is that the usage-based theory stresses the role of generalizing over patterns and schemata, as opposed to setting parameters for a Universal Grammar.

This debate on language acquisition is partially parallel with that in the realm of language processing, where a significant dichotomy is that between associationist views of language and mental phrase markers. Mental phrase markers are thought to be constructed representations of the syntactic structures of linguistic stimuli; these representations are used in parsing and have a strongly Fodorian foundation. As a good example of this kind of view, I will henceforth refer to Pereplyotchik (2017), as it is an up-to-date, well argued for position. Pereplyotchik makes use of the notion of Human Sentence Processing Mechanisms (HSPM), which has some traits of Fodorian language-dedicated modules; the HSPM construct mental phrase markers that provide a structure for the linguistic stimuli that has to be processed. This kind of account relies heavily on many assumptions, of which at least a few have to be pointed out.

Note that while this might be a bit of a stretch in terms of the choice of words ("mental phrase markers" might not be a common expression), it is not too farfetched in terms of the kind of argument that is often produced in neurolinguistics: it is true that a syntatic representation of some kind is usually assumed. On the other hand, the nature of this kind of structure is not very often specified.

An essential remark at this point is that some of the evidence that is offered by Peroplyiotchik as supporting his mental phrase markers hypothesis is connected to even stronger claims, as in the case of Ding et al. (2016), describing them as "[...]neural representations of abstract linguistic structures that are internally constructed on the basis of syntax alone." (Ding et al., 2016, p.163).

Note that such a statement has to be of primary concern for a theory committed to representational abilities, and that two different positions have to be disentangled:

44 See for example Hinzen (2006).

1. The claim that "grammar" has a psychological reality in the sense that grammatical structures are recognized and computed at a cognitive level in a way that is prior to semantics, for instance;
2. The claim that grammar-like representations have additional cognitive roles, apart from language parsing; in other words, that grammar provides a representational format for thinking.

These two claims are substantially different, which is why they should be analysed separately. In most neurolinguistics studies, as a matter of fact, "mental representation" should not be confused with "concept"; rather, "mental representation" refers to the representational means for online processing. Naturally, the two things should ideally be connected in a theory of cognition, but what is important to specify is that what is talked about is a level of representation that does not involve categories, but a format of linguistic processing. Obviously, though, this is still relevant when looking into the role that syntactic structures play in terms of semantic content and, even more importantly, in terms of what role syntactic processing can play in a more general scheme of cognition.

A view that goes in the opposite direction, in not postulating mental phrase markers and relying on associationist networks instead, is advocated in Devitt (2006). Devitt's associationist view is based on the idea that brutal causation is what allows for language processing; comprehension is based on associations between stimuli; responses are causally evoked by linguistic input, in a action-reaction fashion. Note that an important aim of Devitt's theory is to argue against the Chomskyian paradigm of the distinction between performance and competence.

One important difference identified by Pereplyotchik (2017) between an account relying on MPMs and an associationist picture so far is *incrementality*; language processing starts the moment the first morpheme is perceived, and it builds up (bottom up, as a matter of fact), a representation of the sentence structure in an incremental way. The fact that the HSPM makes use of structures incrementally built also implies that the "reaction" to the linguistic stimulus, i.e. the parsing, will be flexible enough to guarantee that what is called in this case "context" will also be considered in the processing of the sentence – e.g. the parser will be able to represent the linguistic context and therefore to choose, for instance, to which subsequent NPs to attribute the role of object of the transitive verb, and so on. In Devitt's account, the rule that attributes the role to the NP does not need to operate on a *representation*; the language comprehension system is directly responsive to the properties that, for instance, make an adjective an adjective. In Devitt's terms, this means that the parser will treat the stimulus as an adjective without representing it as such.

Now, there are two dichotomies in the literature that have been highlighted so far; on the one hand, there is the tension between usage-based and minimalist approaches. This entails the difference between assuming that innate, usage-independent principles are applied to specific natural language grammars during acquisition, and assuming that grammatical competence emerges as a result of usage. The second dichotomy is that between assuming that language parsing entails structured representations (or mental phrase markers), which can be abstract or embodied, and a brutal causal associationist view of how language is processed and comprehended.

While there is a sense in which LALAS finds clear affinity with an approach like that of Tomasello's (Tomasello, 2009b), it is worth underlying that the focus is not on how language gets acquired, but on how the acquisition of language promotes the enhancement of other cognitive abilities. This is reflected by the fact that I argue that the individuation of the schema used for false belief reasoning is possible *once* the child has already acquired a lot of the syntactic abilities that allow her to parse the sentences. This not only is in line with the data presented by de Villiers (de Villiers, 2005; de Villiers and Pyers, 2002), which gives evidence that children already master the complemental clause constructions to at least some extent, but it also implies that LALAS does not make a specific claim about how the syntactic skills themselves are acquired; rather, it makes a claim about how the syntactic skills in place help to individuate a pattern in the way mental states and behavior are talked about in social interactions. These specifications are important because Tomasello's approach, relying on structural alignment procedures and on analogies, clearly bears a natural affinity with my account. Moreover, there is a sense in which approaches that posit mental phrase markers can be particularly compatible with the idea that I propose, that an abstract schema emerges from parsing language; one could easily make a theoretical case for the view that the schema emerges as a result of the mental phrase markers involved in syntactic comprehension. This is a possible approach, and one that specifies a specific relationship between language and thought; the issue of how LALAS fits a more general view of the interaction between linguistic skills and cognitive development will be dealt with at the end of the chapter. For the moment, what is of fundamental importance is to see that each view of syntactic acquisition can be in principle compatible with the mechanism proposed by LALAS, since no specific claim is made about how the structures are acquired or how they develop. At the same time, one does not need to adopt a mental phrase marker approach to accept LALAS, since it is compatible with an associationist view of syntax processing to think that being exposed to linguistic input allows for the individuation of a pattern that gets exploited in the formation of a schema. Schemata are, in LALAS, not inherently linguistic, which

means that there is no commitment to grammar principles as essentially psychologically real.

What is fundamental, on the other hand, is that the syntactic skills are in place when the structural alignment mechanism LALAS proposes is supposed to operate. This is because some skills are necessary to individuate the patterns that lead to the formation of the abstract schemata that facilitate false belief reasoning. Consider:

(5) The apple is green.

(6) Galileo thinks that the apple is green.

In the first case, there is attribution of a property to an object; in the second case, there is a subject followed by a verb followed by an entire sentence, which includes another subject and the attribution of a property. In order for the child to make use of the syntactic clues that are implied in the structure used to communicate mental state attribution, there must be something that allows the child to distinguish the two sentences at a syntactic-semantic level, i.e. to represent (in a loose sense, whatever our theory of language parsing is) that (5) associates a property with an object, but (6) associates an agent with a full description, which entails the association of an object with a property. This is not to say, obviously, that the child is supposed to explicitly recognize this, or have a notion of what a full-blown sentence is; implicit mastery of loose grammatical principles is enough, and in this sense LALAS does not imply any revolutionary claim. Moreover, the developmental trajectory charted for syntactic skills of this kind fits with LALAS's predictions: recall what was exposed in chapter 3. Limber (1973) and Bloom and German (2000) found that in the earliest years, complementation mostly appears in formulaic uses (with the same limited amount of verbs, the same pronouns, etc.) and mostly did *not* express a propositional attitude. These results, combined with the empirical results presented in chapter 4 and chapter 3, suggest that mastery of sentential complements as a subject + verb + full sentence is indeed a progressive, slow process that considerably improves around the child's 4th birthday, making LALAS's prediction that it can be used to identify schemata only then particularly plausible.

Recall the debate mentioned in 3 about the conceptual change and informational change hypothesis, regarding the sorts of cues that allow for the acquisition of mental state verbs. LALAS, like the informational change hypothesis, relies on syntactic structure as a means through which linguistic stimuli are divided and used for further different tasks. In this sense, conceptual development of what a *belief*, for example, is, or what exactly differentiates *know* and *think*, are conceptual advancements that arrive later. At the same time, it should

be clear by now how the emergence of false belief reasoning skills is thought to be, in the framework proposed, the result of the dynamic interaction between social abilities, linguistic skills, and conceptual development. If parsing abilities are necessary for the child to individuate the patterns used in folk-psychology and consequently to use and retrieve the schemata that can be recycled in other mentalizing tasks, that does not exclude the fact that the evolution of conceptual abilities also plays a role on the side. Trivially, semantic knowledge and more general linguistic skills are necessary for the child to understand what happens in the narratives, what the communicative intentions of the caregivers are (arguably, even though surely many low-level skills take part in that to), and so on. At the same time, LALAS does take a stance in predicting that a sophisticated distinction between mental states at the conceptual level will be at least in part the result of the use of different labels to denote the different relations between agents and descriptions of state of affairs. More on this in the following sections.

6.1.5.3 Syntax: which role?

As should be clear from the previous chapter, and as it was partially elaborated in Berio (2020a), theories like the syntactic bootstrapping hypothesis, while explaining data on the relation between syntax and mentalizing, make very strong assumptions about cognitive architecture. In particular, there are at least three options considered by de Villiers when describing the mechanism at play in enabling new representational means for the solution of the false belief task:

1. False belief reasoning relies on logical form (de Villiers, 2004);
2. Specific reasoning and conceptual abilities rely on I–language (de Villiers, 2014);
3. Language provides new computational power to pre-existing resources (de Villiers and de Villiers, 2012).

While options 1 and 2 are developed in a bit more depth, option 3 is just presented as a possibility, and not really endorsed in de Villiers' work. The problems with options 1 and 2 have already been underlined in the previous chapter, at least partially. While 1 is bound to a very precise understanding of cognition that might not be the preferred way to go, option 2 is also problematic in more than one way. I–language is not representational, for Chomsky, which means that de Villiers' account must turn to Hinzen (2006) and his account of I–language as the representational means involved in some forms of thinking. Once again, this means relying on a very specific view of language acquisition. This is not a problem per se; however, as has been pointed out, options 1 and 2 both rely heavily on views according to which *linguistic* representations of

some sort are responsible for passing FBTs; this is, on the other hand, a problematic assumption for a series of reasons, most importantly the fact that it does not clarify what the relation between false belief reasoning of the explicit kind is and what more implicit false belief tasks entail. Additionally, the view implies problems when accounting for the fact that aphasic patients do not seem to have substantial impairment as far as false belief reasoning is concerned, as has been pointed out in chapter 4 and chapter 5 (Varley and Siegal, 2000; Siegal et al., 2001; Apperly et al., 2006).[45] Finally, option 3 is the one that most definitely can fit within the analysis of theories of language and thought interaction that was given at the beginning of this work (chapter 2). Option 3 is, then, the path that LALAS takes. At the end of this chapter, I will highlight in more detail the picture of cognition which is endorsed, or at least suggested, here. For now, it is fundamental to understand that, while syntactic information does play a role in the picture, the schemata are not thought to be *linguistic* per se; on the contrary, they are loose abstract representational schemas that are not sensitive to syntactic clues like tense, but are formed through abstraction from linguistic input. In this sense, LALAS differs radically from the syntactic bootstrapping hypothesis, while still trying to accommodate its fundamental idea that sentential complementation and false belief reasoning are connected in a way that goes beyond the trivial expression of false belief reasoning in language. Syntax helps by giving structure to the linguistic stimuli, and is of crucial importance in individuating the schemata that are used in folk-psychology narrativse. However, syntax does not need to be the representational format in question.

To see what the role of syntax is in this picture, it is useful to think about an analogy. Consider the dynamic, fluid, and varied movement that you perceive when observing ballet. You can either perceive a continuous flow of movement, or break it down into essential steps: Arabesque, Assemblé, Attitude, Développé, Échappé, etc. These movements can also be broken down into motor commands, of course; the level of fine-graininess will depend on the purpose. There is a sense in which language can be used to break down the movement to the desired level of complexity and detail: labels can be applied to specific parts. As the reader might have noticed, this is partially in line with the understanding of syntax as providing fundamental information about the semantics of verb meaning, which is characteristic of the informational change hypothesis (Gleitman, 2009) presented in chapter 3. Syntax is thought to have a similar role in LALAS: com-

45 As underlined in chapter 5, in de Villiers (2014) de Villiers makes claims about conceptual structures but not specifically false belief reasoning, so there is a sense in which I am only assuming that this might be extended by her to false belief reasoning, as it fits perfectly with the rest of her account.

plex explanations of behavior are given in narratives, and understanding the syntactic components in them provides the means for breaking down the input and identifying recurring patterns. This is the case for the schemata emerging from folk-psychology patterns: that of associating an agent with a description of a state of affairs. To sum up:

- Repetition of the association *agent* (denoted by a noun, usually a name) and *state/description of a state* (denoted by an embedded sentence) will create a recognizable pattern;
- This will form a *schema* that abstracts away from the specific contingent associations, and forms a more general relational pattern;
- The syntactic structure underlies the input, and collaborates on the formation of the *schema*;
- The schema encodes a dual relation between an agent and a description;
- There is no explicit, propositional or conscious iteration of the syntactic structure;
- Being able to master this schema makes the information readily available for use and storage, and for interaction and communication.

In the next sections, I will focus on the role of narratives in the account.

6.1.6 Narratives: contexts of use

6.1.6.1 Storytelling and folk-psychology narratives

Another claim on which LALAS rests is that the structure *agent – description of a state of things* is a recurring way in which folk-psychology narratives develop, and that exposition of these kinds of narratives lies at the origin of the understanding of false belief reasoning, and of the formation of the schemata used in folk psychology. As argued in Berio (2020a), storytelling makes significant use of descriptions of mental state vocabulary: explanation of agents' behavior is often given in mentalistic form. While the role of narratives in storytelling has been especially underlined by Hutto, data presented in chapter 4 seems to suggest that, independently from the enactivist account endorsed by Hutto, skills involved in narratives are to be considered importantly related to performance in mentalizing tasks. Not only does interaction with caregivers and siblings bring an increase in mentalizing abilities (Perner et al., 1994), but the specific use of mental state-related vocabulary seems to have a similar positive relation (Moeller and Schick, 2006), and a relation between mentalizing skills and narration skills also appear to be present (Gamannossi and Pinto, 2014; Fernndez, 2011). These findings appear to be in harmony with Hutto's claim that fundamen-

tal access to folk-psychology is given by the narration practices we engage in. Once again, notice how cross-cultural studies on how performance in false belief tasks and mentalizing of a specific kind might develop in cultures that make substantial minor use of these narratives, like those mentioned in chapter 4, would make the case stronger. For cultures that are more familiar to Westerners, Hutto's claim that a fundamental feature of folk-psychology lies in how widespread it is as a social practice, mediated by language and related to specific verbs, is definitely central. LALAS sees narratives as the practice ground and exemplar behavior for using abstract associations between agents and descriptions that encode their own perspectives. It is through narratives that children get acquainted not only with mental state verbs, but also the cultural context in which they are so frequently utilised, i. e. the explanation of behavior. In this sense, contrary to Van Cleave and Gauker and as explained in the previous chapter, I deny that the pragmatic role of mental state verbs does not rely on the production of explanations and narratives regarding other people's behavior. As I pointed out, I do not think their classification really draws any concrete boundary between folk-psychology practices and other uses of mental state verbs. While these are definitely used in more than one relevant way, as after all was pointed out in chapter 3, this does not cancel out the fact that some of their uses are indeed relevant for folk-psychology; but most importantly for my purpose, it does not invalidate the fact that many of our folk-psychology explanations are *given* in these terms, which is what is essential for LALAS. The idea is that the child gets acquainted with the practices of folk-psychology and with three essential facts:

- When dealing with other agents' actions, we often explain them in terms of agent-descriptions of states of affairs.
- These associations between agents and descriptions are rendered in a recurring syntactic form, which is that of Subj + verb + embedded sentence.
- This recurring structure, as a consequence, has in many communicative situations a specific role, i. e. that of providing the explanation required in folk-psychology narratives.

The assumption here is that children are not only sensitive to the grammatical form used in speech (see next paragraph for a justification of this claim,) but will also learn how to engage in specific social games like giving reasons for people's behavior, and will learn the "rules" underlying them. I am using scare quotes for a very specific reason; I am not claiming that children learn a propositional rule that indicates explicitly the way they are supposed to provide explanations of people's behavior, much less that such a rule is meant to possess a specific syntactic structure. What is claimed is rather that the child is sensitive to the recurring formulas and format that are employed in folk-psychology narra-

tives, and that in particular two things will be salient: the structural pattern, on the one hand, and the pragmatic function, on the other. Note that this means adopting, at least partially, the perspective that is proposed by Montgomery (2005), that is, to take seriously the idea that language learning relies on learning roles and patterns in communicative situations, more than on locking in reference relations between words and objects. More on this perspective, and how it is related to, for instance, Tomasello's account of word learning, in section 6.1.5.2.

By "pragmatic" function is intended in this case the fact that this pattern is often used in the context of describing people's behavior: in this sense, storytelling and narration are a legitimate *testing ground* for folk-psychology explanation because they provide the context in which the systematic association between agents and descriptions is used (Berio, 2020a). At the same time, in the context of the cultural practice, what structures the input received by the child is partially syntactic form: it is thanks to the development of syntactic skills that the child can individuate the structural pattern underlying the recurring explanation format that is given in the input, and learn that this is at least one way, if not the most popular way, in which this kind of information is given. The variety of the input received will also be essential: in many situations, stories might not include a perceivable referent, i. e. might not be making use of visual aids or illustrations, while many others might be told with visual props. In both cases, very different situations involving different agents and contexts will be described using the same descriptive means. In this large quantity and variety of data, the child will be able to individuate common patterns: here is where the structural alignment mechanism described above will kick in. Using the same language to describe a large array of different situations allows the child to compare them and individuate the underlying pattern; in other words, from the ability to individuate patterns in speech, the learner gets to the comparison of different situations involving different agents; from comparison of these situations, the child can individuate the common pattern of relating agents and descriptions and consequently use it to solve FBTs (Berio, 2020a). This is fundamental to making the schema available. Importantly, however, the schema is not just the result of the acquisition of the syntactic skills, but rather the result of the syntactic stimuli being embedded in cultural practices.

6.1.6.2 The use of schemata

One of the mechanisms that supports LALAS's story is that of *selective attention*[46]. The idea has been exploited by Barsalou and others (Barsalou, 2005; Tillas, 2015a) to argue in the context of conceptualization processes. The idea is that attention focuses on selected features in perceptual stimuli; in the case of Barsalou's and Tillas' model, this is crucial in determining which perceptual features will be stored as having fundamental importance in a modal representation. Cognitive psychology distinguishes between top-down and bottom-up attention processes, i.e. attention that is guided by the goals of the individual and higher-level processes and attention driven by external stimuli, like a loud noise (James, 1890; Eysenck and Keane, 2012). Top-down elements that can guide selective attention include episodic memory traces (Uncapher et al., 2011) and, according to Tillas, conceptual features; the idea is that already stored information can determine how the stimuli are processed. For example, Tillas reading is:

> Later on, selective attention is also driven by top-down effects from stored representations. [...] stored representations of branches, which also carry information about the position of the represented branches in the visual field, drive attention, for example, to the top part of the visual field (e.g. via influencing oculomotor movements) when perceiving a tree and so forth. (Tillas, 2015a, p.225)

I argue that something similar happens with schemata present in folk-psychology: on the one hand, the fact that these explanations are given is a pattern that gets detected; on the other hand, the child will be especially drawn to the recurring structure composed by the association of an agent and a description, since such associations are salient in folk-psychology explanation. In other words, the fact that explanations are given in mentalistic terms will guide the child in understanding that *that* kind of information is relevant when we are explaining people's behavior, and consequently to look for these association patterns in the given stimuli and to use them when answering relevant questions (Berio, 2020a). Apperly (2011) argues persuasively that, when faced with a false belief task, a child has to take into consideration an extremely large amount of information, with a complex and perceptually rich scene that has to be computed efficiently and rapidly. This is surely true, and Apperly uses this argument to claim that a lot of what we do in our daily interactions is to use schemata and scripts of social situations in order not to have to deal online with this amount of information, instead using pre-existent slots and frames. In this way, going to the bakery is not an impossible task, but one that can just be solved by appealing to socially

46 For a shorter version of this argument, see Berio (2020a).

and linguistically constructed frames that help navigation. There is a sense in which I claim something very similar, in assuming that the schemata acquired through language provide precisely this kind of "short-cut"; the child is already able to do a lot of things before 4 years of age, including using well-developed skills to track attention, goal-directed action, and so on. The average infant also seems to be able to register associations between locations and objects, at least according to some interpretations of the implicit false belief task literature (Onishi and Baillargeon, 2005; Scott and Baillargeon, 2009; Song and Baillargeon, 2008). However, having an abstract schema is useful in so far as it guides attention during interaction, as the child is implicitly instructed to pay attention to specific "mentalistic" explanations; moreover, there is a quick, handy way to store information regarding the association between the agent and the description, and provides a means for solving false belief reasoning tasks that are more difficult, like the classic FBT.

If it is true, then, when a *model* of the situation is created by the child to deal with the incoming information (Apperly, 2011), it is plausible that the model itself makes use of said schemata, and that it comes with pre-acquired expectations not only at a coarse-grained level, i.e. that bakers usually behave so and so, but also at a more fine-grained level, i.e. that agents act according to some points of view, and that this is easily captured by associating and agent with a description of a situation. Looking at the false belief situation, the 3-year-old child comes to the scene with the ability to keep track of associations between locations and agents, between agents and preferred movements, between agents and determined perceptual perspectives, and so on. For the child to have these abilities, no abstraction is probably necessary. Things change, however, in the case of the elicited response, explicit false belief task. Take the example of the classic Sally and Anne situation; here the child has to make an explicit prediction of how Sally will behave. The child that has an agent-situation schema to hand is facilitated in at least two ways: for a start, the association has been stored in a sufficiently abstract format, which might be a more effective way of storing it in the first place, i.e. more easily kept in mind. She also has a readily available schema concerning information relevant for predicting the behavior; she knows that when dealing with other agents' behavior, the association between the agent and the description is what is usually requested, since this is something that she has learned through exposure to narratives and folk-psychology practice. The schema provides an easy way to recall this information, since it is readily stored in a quite general format. Having the association represented in a more abstract form can make it available for inference processes that are not plausibly involved in Onishi and Baillargeon's (2005) gaze following task, for example, where measures are taken that do not necessarily imply explicit elabora-

tion. Moreover, a working hypothesis is that these schemata make it easier to compare different associations; since it is clear, thanks to the abstraction aided by language, that very different situations can be encoded in the same dual-format, these can also be compared more easily, as values of the same parameter. This can make keeping in mind and operating on different, potentially conflicting, associations easier and more straightforward, thus facilitating inferential reasoning with them. Finally, the fact that the schema has formed as a result of the recurring occurrence of an agent associated with a state of affairs *presented in the form of a linguistic description* can be what facilitates the child's expression of the prediction in the same format, i. e. verbally.

At this point, one might wonder how this account actually fits within the data presented before, and how it is related to a more general view of mentalizing. The next part of this chapter will be dedicated to making this suggestion more concrete, explaining how the idea of schemata learned through language can fit at least two current models of false belief reasoning. In this way, the proposal will be embedded in a more general view of mentalizing processes. Before going further, let me sum up what the main advantages of having a schema are:

- It helps with quicker identification: once the abstraction is present,
  - Attention is drawn to the recurring relation between agents and descriptions: the child knows "where to look" for relevant information.
  - It makes it easier to store new associations in pre-assigned slots, as they fit a pre-recognized pattern.
  - The association between Sally and the position of the toy is *more easily* kept in mind because it can be stored in this format, which is not linguistic, but has formed via abstraction from (also) linguistic input.
  - This allows the association to be used in interaction and communication as well.
- It might make different or conflicting associations easier to compare: agent *x* and description *p* can be more easily compared to agent *y* and description *q*.

A claim LALAS makes is that the different *labels* used to describe attitudes (that is, the different verbs) will be understood as specific relations with the state described by the sentence, in a way that will help the formation of the full-blown mental state *concepts* that might be involved in higher-level mentalizing, in more sophisticated folk-psychology, and ultimately in elaborated concepts of mental states. In other words, labeling the different attitudes in different ways (think, believe, et.c) will help to mark conceptual differences. It is essential to remember, in this case, that children start differentiating different mental state verbs when they are 4 years old, as explained in chapter 3. This is, then, compatible with the idea proposed in this case, that differentiation between different mental

state verbs comes with understanding how the agent and the description can be variably related to each other. There is no developmental revolution with mental terms acquisition: there is, however, a boost, granted by the fact a more general association between agent and description is available for storage and use. Moreover, it is the labeling of different relations between agents and descriptions of states of affairs that potentially leads to the full-blown formation of concepts of belief, thinking, and so on.

6.2 Applying LALAS to mentalizing: two false belief stories

It has been argued so far that language provides recurring structures that form a schema that is re-applied in subsequent tasks, where associations between agents and states of affairs in the form of descriptions can be used. The schema language provides, it has been argued, produces a more general frame of reference; not just an association between *that* agent and the location of *that* object, for instance, but the relation between *agent x* and *description z*. What will now be explored is how this idea fits in a more general picture of mentalizing.

As will be seen, it is not argued here that the *only* reason why children fail the false belief task is that they have not acquired the relevant linguistic competence; it is very likely, and compatible with LALAS, that the reasons why children fail FBTs are varied, including very different factors like executive control and working memory. However, it is argued that language provides a fundamental aid in this sense. To see how, I will go through two different theories of mindreading. These have not been presented in the previous chapters because they do not address the problem of language and mentalizing, however, I will argue that not only do they leave some room for a role for linguistic development, but LALAS completes these views in providing a mechanism that fits them in a productive way. As will be seen, the two views are not completely compatible; deciding between the two is a matter of empirical questions, on the one hand, and theoretical stances on the other. In both cases, however, LALAS makes a concrete contribution in providing a specific mechanism for implementing the role of mental state terms and syntactic acquisition, thus having the advantage of being compatible with two promising, albeit opposing views.

6.2.1 Double systems mindreading

Before describing the two views, it is worth saying something about why I decided to focus on so-called double mechanisms for mentalizing.

In chapter 4, I explained the so-called *developmental paradox* concerning mindreading: while there is solid evidence that explicit FBTs cannot be passed until one's fourth birthday, even once the demands of the task are modified, there is also increasing evidence that other kinds of mentalizing activities are found in much younger children. The validity of some of these studies is controversial, but the data still seem to point in the direction of a quite complex developmental picture and developmental trajectory as far as mentalizing skills are concerned. There is good reason to doubt that TOM, or whatever the underlying mentalizing processes are, is an all-or-nothing matter, and while each proponent will try to argue that their theory accounts for all the data, it has been convincingly argued by many (Bermúdez, 2009; De Bruin and Newen, 2012, 2014; Apperly and Butterfill, 2009; Apperly, 2011) that a single mechanism for mentalizing cannot account for the variety of data.

When considering the role of language in mentalizing, this is even more evident; since many of the abilities that include intention reading and goal-directed action understanding seem to be at the root of more complex mentalizing behavior appear to emerge significantly earlier than the onset of linguistic abilities (see section 4.2.2 in chapter 4 for the review of this data), it is clear that a view that accommodates the role of language has two choices: either ignore the fact that a continuity is likely to be present, and give an account in exclusively linguistic terms, or postulate that language acquisition can account for parts of the more complex and varied mentalizing skills. While the first route has been the one pursued by many theories presented in the previous chapter, the latter is the path chosen here. As a consequence, I focus on two pluralistic accounts of mindreading, for two reasons: firstly, they are the most detailed available and combine different instances, including elements from ST and TT; secondly and most importantly, as will be seen, they leave room for a role for language and can be efficiently integrated with the present proposal. While I have already argued in Berio (2020a) that they are the most compatible with my approach, I will here elaborate on how exactly can my account fit in these frameworks.

6.2.2 Apperly and Butterfill

Apperly and Butterfill (2009) and Apperly (2011) argue (the latter at length) in favor of their two-system view, making a parallel to number cognition, where there is significant evidence that, while some aspects of basic number cognition are present in very young infants, others develop with age.

Mostly, the data they base their argument on regards three basic core themes: the fact that infants and non-human animals are capable of some

forms of belief reasoning, but not others, as explained in the previous chapters; the fact that sophisticated belief reasoning is cognitively demanding for adults, relying heavily on language abilities and working memory (Apperly et al., 2008); and finally the fact that, despite this, folk-psychology is part of common practice. This kind of evidence gives good reason to think, the argument goes, that more than one system is involved in mentalizing. More specifically, the idea is that there are two systems: a flexible albeit demanding one that requires sophisticated cognitive abilities, and one which is automatic and mandatory, being already present at birth and had potentially in common with most other animals. The flexible, demanding belief reasoning treated by Apperly and Butterfill (2009) is described as that related to reason-giving explanations, which normally require (p.960):

1. Complex causal constructions, where distal and complex causes are involved in the explanation.
2. Abductive reasoning, because they employ inferences to the best available explanation.
3. A normative dimension, because they are supposed to describe the agent's *rational* Behavior.
4. Ascriptions of states with propositional contents.

These features of belief ascription are thought to be the cognitively demanding part of it. On the basis of these features, Apperly and Butterfil reject some accounts that try to explain away the logical problem. Theories of innate belief reasoning competences, like Leslie (1994a); Onishi and Baillargeon (2005); Fodor (1992), while explaining why adults' mindreading skills are superior to infants' by appealing to conceptual abilities changing over time because of limited general processing (in the case of Fodor) or executive function (Leslie), do leave other explanatory gaps[47]. According to Apperly and Butterfill, these theories are lacking in that they do not give an account of how, in adults, both flexible and demanding mindreading *and* mandatory and automatic mindreading, depending on the context and instance, seem to be possible[48]. Explanations given in terms of reasoning about factual states rather than mental states (Perner, 1991; Csibra et al., 1999; Csibra and Gergely, 2007) are also considered ineffi-

47 As mentioned in the previous chapter, I have a less generous interpretation of Leslie's view, partially due to the fact that executive function, as pointed out in 4, does not seem to be able to completely explain the gap.

48 Most of the theories I will mention in these paragraphs are not concerned with the role of language, which is why they were not mentioned in the previous chapter. See section 5.1 for the explanation.

cient, for two fundamental reasons: on the one hand, they do not cover the full range of mindreading abilities, since they do not consider perspective taking, differences in perception or belief. On the other hand, they do not take into account the fact that reasoning about facts can also involve complex causal constructions and have a normative dimension. They also point out problems with two different sets of accounts; on the one hand, *automatization* accounts rely on the idea that, once a hard-to-get ability like false belief reasoning is achieved, enough practice and experience can make the task effortless in adults (Suddendorf and Whiten, 2003). While this is plausible, it does not explain how infants are indeed able to solve some mindreading tasks, and not others, since automatization only comes after learning. Association accounts are also briefly considered. While prior to De Bruin and Newen (2012), the paper engages with claims made elsewhere (Baldwin and Baird, 2001; Povinelli et al., 2000; Perner and Ruffman, 2005; Ruffman and Perner, 2005), according to which statistical regularities like head orienting-approaching to an object associations can be exploited by infants and non-human animals to formulate predictions of behavior. One of the criticisms of these approaches offered in Apperly and Butterfill (2009) is that data about statistical regularities in behavior are potentially lacking; moreover, according to Apperly and Butterfill (2009), they fail to account for cases in which perspective interference, like those in Kovacs et al. (2010), where self-perspective seems to interfere with mindreading abilities.

On the basis of these arguments, they propose a different story, comprising the flexible system sketched above and one operating with *belief-like states*, or *registrations*. Registrations are defined as proxies for beliefs and as relations to objects and properties, but not to propositions, and more specifically:

> One stands in the registering relation to an object and location if one encountered it at that location and if one has not since encountered it somewhere else. Registrations resemble beliefs in having correctness conditions that may not be obtained: A registration fails to be correct when the object registered is not where it is registered as being. (Apperly and Butterfill, 2009, p.962)

This registration relation can be used as motivation for action, in the sense that having a registration relation with an object can be the basis for later returning to the location where the object was registered, for instance; at the same time, it can be used to predict actions or interactions with an object, i.e. to predict the movement towards a location to interact with an object registered there. Registration of location can be followed by registration of other properties, for instance stance. Registrations must however be limited in two ways; they have causal power on action only independently, i.e. they cannot be combined, unless

in a codifiable way[49]; also, they cannot involve quantifiers or complex combinations of properties, or distinguish between what is represented and how it is represented. Therefore, they can be used as proxies in a limited number of tasks, only limited to Level 1 perspective taking, for a start, and not Level 2.

According to the authors, the two systems are most likely independent; since the flexibility of the late-blooming mechanism is supposed to be opposite to the rigidity and efficiency of the registration-based one. Limits on the format of the belief-like representations produced at the registration level would mean that input from it could not be useful to the second system, and vice versa. As will be seen, this is the major point of departure between this account and the associationist one proposed by De Bruin and Newen (2014) and De Bruin and Newen (2012).

The view evolves in the later version presented in Apperly (2011), although the fundamental idea remains that of two independent systems that operate in different situations and bear different relations to development. In this sense, the fundamental take is that the automatic system does operate on associations, whereas the high-level mindreading system is the one that operates on beliefs and attributions. Apperly leaves open the question of whether or not System 2 operates with concepts of mental states, arguing that the debate about concepts is in this sense too unresolved for a decisive answer, and this might be the reason why talk of *belief-like* states and *registrations* is actually abandoned in the later version of the theory; what is relevant on the other hand is that a normative dimension is recognized as part of System 2, which operates with higher-level cognitive mechanisms that track the behavior of *rational* agents. System 1, on the other hand, relies on easier association mechanisms. While level 1 perspective taking is something that is a competence of System 1, level 2 perspective taking, i. e. taking into consideration *how* another agent sees an object or a situation, is to be attributed to System 2. The developmental paradox is explained in terms of the two systems operating on different levels and being acquired at different times[50].

49 While this is not completely clear in the text, I am under the impression that this means that registrations can only be considered in a simple sum (one after the other) but not combined in more complex ways.

50 In this sense, the authors choose to join a growing body of literature that explains developmental phenomena by posing two different systems, as already mentioned in chapter 2 when discussing Carruthers' view. The idea is that System 1 processes are automatic, experiential, implicit associative, intuitive, rapid, and often domain-specific, while System 2 processes are abstract, rule based, domain-general, evolutionarily recent, heritable, analytic, linked to general intelligence, and so on (Evans, 2008). Exactly how these properties are distributed depends

At this point, it is relevant to mention the role of language in Apperly's theory. While discarding the idea that syntactic acquisition and semantic development in particular play a role in mentalizing abilities or in the development of System 2, Apperly argues that a function of language in its social dimension is instead that of providing situations and scenarios in which children are trained to understand other agents. In this framework, as already anticipated, a significant role is attributed to social schemata and scripts that are used in particular situations. The concept of a script derives in this case from cognitive psychology, and Apperly relies on social psychology (Cantor et al., 1982; Fiske and Taylor, 1984; Gilbert, 1998). The idea is that social scripts contain information about how particular events and situations usually unfold, including the roles of people involved, the objects present in the situation, and the normative dimension: what we are supposed to do, how it is correct to behave, and so on. These schemata are at play when we deal relatively mindlessly with everyday situations, like going to the bank: we behave in certain ways because of complex social scripts that we use without too much effort, since they are based on generalizations that make these situations ultimately doable.

The view, as Apperly recognizes, is similar to that proposed by Nelson (2005) and Hutto (2008b), but:

> [...] I believe this endowment has a further role, not envisaged by these authors, in enabling children and adults to identify what information is likely to be relevant in a given situation. The social origins of much of this information ensure that there will normally be widespread interpersonal agreement on what is relevant in a given situation. (Apperly, 2011, p.160)

Apperly does not go into details about how language makes some information more relevant, but he does stress the fact that social scripts of the kind mentioned, which are partially derived from language, do make mindreading a *tractable* problem, because they help to deal with the *abductive* reasoning problem. "Inference to the best explanation", Apperly argues (pp. 118 – 119), is of central importance in mindreading tasks. When interpreting Sally's behavior, the child has to identify which kind of beliefs are relevant for the current situation, and the task is not trivial given the great quantity of information that is present

on the account, which for (Evans, 2008) is a good reason to abandon the labels System 1 and System 2 in a generic sense to cluster all these theories together. It is worth mentioning that fairly recently dual system theories have been proposed in reasoning (Evans and Stanovich, 2013) and decision making Kahneman and Frederick (2005) as well, even if, as Evans (2008) argues, dual system views have been present in the cognition literature for longer, under different labels.

on the scene. Scripts and social models make this kind of information tractable, and help to determine which kind of information is relevant.

As should be clear by now, this view is highly compatible with LALAS; in a certain sense, it could even be said that LALAS expands Apperly's idea by assigning it a specific role for language. Apperly declares more than once in the book that acquisition of determined syntactic structures is not what makes language relevant, and thus poses a distinction between the social role that he wants to attribute to linguistic communication and the more specific role of mastery of language and determined linguistic information. In this sense, LALAS finds a place for the latter in light of the former, proposing to make the role of linguistic acquisition a specific piece of the story in social development, and giving linguistic structures a specific role to play in the more general social development account. In LALAS's case, the prediction is that syntactic information and language in general contribute to the formation of narration schemata that are part of the way children understand narratives and engage in social dialogue. As has been seen when treating Bermúdez (2009) view, Apperly is not the only one appealing to frames and abstractions to explain efficient mentalizing, as Bermúdez appeals to Minsky (1975) and frames as situational schemata of a similar kind. The difference between these proposals and LALAS, then, is I suggest that language information structured syntactically can help the formation of similar schemata as well[51].

When presenting the two systems idea, Apperly gives a rough division of labor of the processes involved in mentalizing, assuming that both System 1 and System 2 have to deal with them, making the working hypothesis that these can be divided into *inference*, *storage*, and *use*. The idea is that these processes come one after the other and are reflected at both levels, being processes that have to happen for both mentalizing skills to be present.

Apperly spells out in these terms three cases where the FBT is not passed by 3-year-olds. The first case is the standard prediction task in which the child has to predict the agent's behavior, thus going from inference, to storage, to use. In the second case, the child only has to report on Sally's belief, without making

51 Apperly argues that System 2 has to argue with the "complicated business of representing relations between agents and propositions." (Apperly, 2011, p.147): this is not something that gets further elaboration. I steer clear, here, of the problems related to the notion of *proposition* because what I am suggesting is that linguistic information structured according to certain patterns allow for the relating of agents to what are ultimately descriptions of situations in representations, but whether or not this happens through a "layer" in between sentences and linguistic input, and the representation, i.e. through propositions, is not really a concern for the account. Special thanks are due to Prof. Gottfried Vosgerau for this observation.

predictions, thus going from inference to storage. In the third case the child is told the content of Sally's belief and what is really the case, and she has to report it back, thus making no real inference, but having to store the information. The fact that all of these tasks are failed by 3-year-olds, it is argued by Apperly, might be because of the storage of the associations involved.

Crucially, the "higher" flexible system and the "lower" efficient system are not connected: so for example, when the child is performing the implicit false belief task, and her gaze predicts the fact that the agent will look in the wrong box, this happens at the level of the lower system. However, since there is no connection with the higher system, which is not functional for the 3-year-old yet, an output cannot be produced at that level: therefore, the child cannot make an explicit prediction.

Note, then, that this fits perfectly with LALAS, in so far as the idea is that language helps create an association schema that facilitates, among other things, effective storage of new associations. In this sense, assuming there is really a problem of storage in the mentalizing tasks described, this might be due to the fact that children have not yet acquired the schemata that allow for effective storage (Berio, 2020a). In this case, then, the information coming from linguistic input does not only provide input for the creation of the complex situational scripts used in folk-psychology, but also provides a cognitive tool for the use of information acquired otherwise.

Finally, it is worth mentioning the stance that Apperly takes on *modularization*, according to which a high-level process can get modularized with experience, i.e. become automatic and easier with sufficient practice. The idea then is that limiting the input and output range can make very complex processes easier to compute, and thus make it quicker to deal with the amount of information. In this case, an effect given by practice with given social scripts and situations can make the process automatic, and the output can even be used as a basis for further, more sophisticated processes. This intuition is important, because there is a sense in which LALAS predicts exactly that; language provides us with "shortcuts" that make dealing with information easier, but contrary to Apperly's theory this does not only come in handy when forming social scripts, but also when it comes to organising, storing, and using information that can be dealt with more abstractly because of language. This has the obvious advantage of accommodating the data about sentential complements and linguistic development. Before making a final assessment on Apperly's theory and its possibilities, I will now turn to a rival double-mechanism account.

6.2.3 De Bruin and Newen (2012)

As seen above, Apperly and Butterfill argue that the two systems have to be different to explain behavior in false belief understanding, because this solves the "developmental paradox". De Bruin and Newen (2012), on the other hand, argue against this choice for fundamentally two reasons:

1. They consider the choice implausible: it would need to be explained why two different mechanisms, that have nothing in common but are used for the same social-cognition purposes, have evolved;
2. They argue that data suggests otherwise, since there seems to be a correlation between performance in implicit false belief tasks and explicit false belief tasks, with performance in the former being predictive of performance in the latter (Aschersleben et al., 2008; Kristen et al., 2011). They therefore argue that what is needed is not only to account for the dissociation between performance in implicit and explicit tasks, but also the continuity of the two.

1 is, as should already be clear, in line with the general framework I argue for, since I think there are concrete reasons for assuming that language interacts with pre-existing skills to enhance cognitive abilities, and that postulating two non-communicative systems can in this case be misleading. However, note that Apperly (2011) is more open to the possibility of interaction between the two systems and assumes that information initially processed by a flexible mechanism can be treated automatically once downward modularization has occurred. In any case, 2 raises an interesting issue, that is to be kept in mind: a fundamental choice has to be made whether or not to postulate a continuity between the two mechanisms[52].

In De Bruin and Newen (2012, 2014), there are two systems:

1. Association module;
2. Inhibition, selection, representation system (operative system).

It is the interaction between the two modules that allows for the increasing ability to form complex associations used in FBTs.

By "association" is meant a structured representation of goal-directed behavior (De Bruin and Newen, 2012, p. 244). The assumption is the infants have the ability to:

1. Recognize human subjects as different from inanimate objects;

52 This independently from the fact that some data might be caught in the latest replication crisis.

2. Understand subject's action as directed towards an object.

The claim is that the association system starts out by providing associations between an agent and an object based on the basic relation of *movement towards*. This is based on studies that show that infants develop, around 5 months of age, attention to the *goals* of the observed agents (Woodward, 2003), for instance. Subsequently, around 12 months of age, this behavior gets connected to the *looking* behavior of the subject as well, allowing for more distal associations: in other words, it allows for registering not only associations between agents' movements and objects but also associations between agents' glances and objects. The association module in question is what guides, in the account, not only the infant's understanding of herself as an intentional agent, but also the understanding of the intentional behavior of others. The module gives a way for the infant to develop a "bi-directional capacity of action-perception"; in other words, the infant is able to use its own cognitive perceptual and motor resources to understand the goal-directed movement of an agent, and to register the association between a moving agent, the goal of the action, and the object. The capacity of action-perception, then, is *bi-directional* because it involves both action perception and action production. The characterization is then meant, firstly, to ground theory of mind abilities on a cognitively basic level, assuming they are built up from relatively simple association abilities; and secondly, to accommodate data from the simulationist and mirror-neuron perspective in respect to so-called "shared-representations" (Kovacs et al., 2010) and, in general, to evidence found through mirror-neuron research (Rizzolatti and Sinigaglia, 2006; Rizzolatti et al., 2001; Rizzolatti and Gentilucci, 1988; Rizzolatti et al., 1987).

The *operating system* is described as a system that works on the associations formed by the association module, and that at first operates only on incongruent motor associations, then "learns" to operate on incongruent perceptual associations as well, and finally is able to deal with symbolic associations. Fundamentally, both systems can function by taking as input all three kinds of associations (i.e. motor, distal, and symbolic).

The presence of the operating system allows one to:

1. Inhibit associations that are not relevant or that would lead to unwanted movement: this is supposed to prevent the system from generating the replication of a movement when all is needed is a representation of it; in other words, it allows one to discriminate between perception and action.
2. Select the right motor information, on the basis of the relevant perceptual information, in order to correctly move to the following step.
3. Represent the selected information.

The representation produced by step 3 can be the basis for movement; in case the toddler has learned to associate the movement towards box A with getting a toy, but the toy is moved to another box B, she can now perform the same movement toward box B, thanks to the inhibition of movement towards box A, the selection of the motor information (the movement to perform) in relation to the perceptual information (the toy is in box B) and the representation of the selected information. But the same representation can be used in understanding another person's behavior, as in step 2 the motor information can be selected on the basis of the perception of the movement of the agent, generating in step 3 a representation of the movement expected by the agent. This, according to the account, explains why in experiments like Woodward (2003), infants look longer when an agent mistakenly tries to reach a toy in the box where it was, instead of in the one where it is; the expectations are not met. In this case, there is a motor association to be used. Matters become even more complicated when an incongruent association is involved, i.e. when the child does not only have to inhibit the response according to step 1 and represent a relation according to step 3, but also needs to select the representation of the agent, which is different from her own, as in step 2. This involves the representation of visual information that "specifies the agent's previous perception of the object", for instance. The case of incongruent association is then particularly hard to deal with for the two interacting systems, as is the case in the spontaneous response false belief tasks. In this case, the child has to decouple the information regarding perceptual input.

Exactly the same can be said with regards to what De Bruin and Newen (2012) call "symbolic associations", like for instance those entertained between a linguistic symbol, and an agent. Even in this case, the operating system will function on three subsequent steps: it will inhibit a response, select an appropriate association, and produce a representation. The difference between performance in an elicited response false belief task and a spontaneous response false belief task is, then, seen in the literature for reasons that are similar to the case of perceptual association above. While the spontaneous response task does not require a representation of the agent's false belief, but can be solved by the child inhibiting her own representation and forming and selecting the association between the agent and the symbol, in an explicit false belief task, to be able to give the correct answer, the child has to operate on the symbolic level as well, and to represent the incongruent relation between the agent and the symbol verbally. Not only does the child need to represent the association; in this case, she has to form a meta-representation, representing not only what the other agent represents but also how she does it, i.e. the propositional attitude.

Let us see how the classic Sally-Anne task is explained (De Bruin and Newen, 2012, p.252):

> Children read a cartoon scenario in which Sally puts her ball in the basket. The association module registers a symbol-based association between Sally and the basket one that is congruent with the infants own symbol-based association. After Sally leaves, Anne places the ball in the box. The association module registers a symbol-based association that is incongruent with the one registered previously. When Sally returns, the operating system enables the infants to predict where she will look by: (i) inhibiting the infants own symbol-based association, (ii) selecting the symbol-based association that specifies Sallys belief about the location of the doll. In the final step, the operating system (iii) represents this information as a verbal prediction about Sallys behavior to search in the basket. On our association view, this last step is not required in the verbal spontaneous-response FBTs by Scott et al. (Scott et al., 2011)

Note that in De Bruin and Newen (2014), the authors abandon the use of "symbolic" as a term for denoting associations but substitute it with the clearer distinction between *perceptual* and *cognitive* perspectives, which basically reflects the distinction between level 1 perspective taking and level 2 perspective taking: on the one hand, what is visible from another agent and on the other hand how it is visible for another agent, as explained in chapter 4. Children are able to represent "in order to" relations to objects, in the sense that they can understand goal-directed action; however, it is argued, "the emergence of linguistic competence, long-term memory, and executive functioning more in general allows children to (re)configure the information encoded in their associations in much more abstract in-order-to relations." (De Bruin and Newen, 2014, p.309). In any case, what is thought to be lacking for children that still do not pass the false belief task is the ability to represent the incongruent cognitive perspective, or at least to do it in a way that allows them to formulate a verbal prediction about it.

While not immediate, there is a connection between this approach and that of LALAS, since there is the common idea that what is lacking is an adequate way to represent associations. LALAS' prediction is basically, then, that language helps the learning child to be able to do exactly what is required for the operative system to operate more efficiently on those associations, and to store them in the right "format", i.e. a more abstract one that allows representation and verbal prediction. Not only does LALAS predict that acquiring the right schema will facilitate performance in the false belief task while forming an association that is in the right format for the operating system to represent a cognitive perspective; additionally, the idea is that this capacity builds on the existing ability to exploit associations between agents and locations, which is a perspective fundamentally in line with the view advocated by Newen and de Bruin. In LALAS schemata, the agent is not related to just a location, but to a complex situation; as argued, hav-

ing an abstract representation for this relation might be what facilitates easier comparison between different sets of agent–description pairs, on the one hand, and what allows for faster operation with these associations, on the other hand. In this sense, linguistic input would function as a training ground for the operative system, which would then learn to recognize cognitive perspective – i.e., in LALAS terms, to spot the right kind of association – and to use these schemata to solve tasks in which different associations of this kind have to be kept in mind. The obvious advantage of this account compared to Apperly and Buttefill's explanation is that a direct relation between implicit and explicit false belief tasks is part of this theory.

6.2.4 Relating LALAS to other views

LALAS can be related to many of the views that have been examined in the previous chapter. In Berio (2020a), I have compared my schemata-based approach with some of these frameworks; here, I will elaborate on the subject more in depth, showing how LALAS is positioned in the literature. One obvious similarity is with Garfield et al. (2001) (henceforth LS, Language and Social cognition); however, note that while LS is presented as strongly necessitating some kind of modular stance, LALAS does not require it. There is also a difference in scope, on the other hand, due to the fact that, while advocating a role for language in developing mentalizing skills, LS does not specify which kind of features of language do which kind of mentalizing work easier, and does not distinguish very efficiently between different kinds of mentalizing at all. Finally, according to LS, there are two separate sets of skills, i.e. language and social cognition that, while separate and modular, are independently necessary for the development of ToM. LALAS makes a different claim; that is, the acquisition of determined language skills and vocabulary enhances performance in determined mentalizing tasks because it provides a way to form schemata that allows for complex articulation to be made faster and more precisely. This difference also entails a different take on language, as seen not in the Chomskyian perspective of something only secondarily involved in communication, but rather as a fundamentally social tool, hence part of social cognition skills instead of something only interacting with it in terms of ToM. However, it is worth emphasizing that my proposal does not need to be read as opposing LS; there is a sense in which the theoretical assumptions about weak modularity and weak innatism can be accommodated in LALAS as well. In this sense, LALAS's specific bootstrapping mechanism can be included in a picture like that in Garfield et al. (2001), and it can be seen as complementing it.

LALAS is furthermore compatible with Montgomery's (2005) intuition that learning mental state terms and use is less a matter of learning by ostension, and more a matter of learning roles. Similarly to Van Cleave and Gauker (2013), however, Montgomery focuses specifically on the pragmatics involved in mental state language; as argued, LALAS's scope extends a bit beyond that.

Recall also the idea in Baldwin and Saylor (2012): the proposal is that abstraction, as the necessary process for forming concepts of mental states, is achieved by analogy and structural alignment of different situations aided by language. This is an interesting proposal because it gives language the kind of role that LALAS also advocates; however, this mechanism is thought to be the one at play for the *concept* of belief, and does not have so much to do with learning to apply different perspectives. It is a move towards forming a representation of belief, which is necessary at later stages; the focus of LALAS is on the previous stage, following the proposal in Montgomery (2005) that a full-blown representation of a mental state as such is not essential at first, where the application of a schema of intentional actions and perspective can be sufficient. The idea in Baldwin and Saylor (2012) is that, when making sense of reference to something that is physically absent, structural mapping intervenes making room for the idea of an intended referent – in this case an internal focus of attention. Unable to retrieve the pattern *word "dog" – perception of DOG*, the child would then associate the word with a *dog-focus* as the referent of the conversation. As has been explained, this is a different take, as in LALAS associative structures are thought to be what is primarily gained by language input.

Finally, there are two theories that are fundamentally related to LALAS, and to which the present account owes many theoretical insights. In a sense, the theory puts together Hutto's claim, that folk-psychology narratives are of fundamental importance for the development of skills related to mentalizing, with de Villiers' idea of a syntactic bootstrapping mechanism. I argue that children make use of the syntactic clues provided by language as they form abstract schemata that can be used in subsequent situations, re-employing structures that are not necessarily linguistic in nature; these structures derive their role from their use in folk-psychology narratives. Note that this also avoids the cross-linguistic issues that emerge with the original de Villiers' proposal: sentential constructions with desire verbs in German, for example, are not an issue for LALAS, because it is not just the ability to deal with sentential complement structures that allows more abstract representations, but also how these structures are used and how they are embedded in social practices. The child applies to folk psychology what she has learned from it: that explanations of behavior are given in light of associations between agents and descriptions of things, and that these can be kept in mind and manipulated in an abstract format. None of this learning is explicit,

and while syntactic abilities play a role in it, they are not the only element. In the next section, I will highlight how the mechanism for LALAS is compatible with the schemata being learned, at least in part, by non-syntactic means.

6.3 Not only language: development of schemata through other kinds of input

One might be tempted to take LALAS as a hypothesis about a *unique* role of language in boosting mentalizing abilities, and there is a sense in which this proposal is definitely focused on how language provides an ideal scaffolding mechanism to produce useful abstractions and generalizations that can facilitate mentalizing development. However, it is worthwhile specifying that I am not claiming that the only way to form schema-boosting mentalizing abilities is through exposure to linguistic structures; on the contrary, the fact that the schema is itself non-linguistic seems to leave room for the possibility that it gets acquired some other way (see also Berio (2020a)).

As a matter of fact, there seems to be evidence that something similar can be achieved. In Wellman and Peterson (2013), a set of deaf children with hearing parents was trained with thought bubble scenarios similar to those used in Hale and Tager-Flusberg (2003) and described in chapter 4, where speech bubbles represented the thoughts of the characters in narratives illustrated with cartoon storyboards, and things could be placed in the speech bubble to signify the characters' thought.

Interestingly, this kind of training was particularly effective in increasing the performance of deaf children with respect to false belief reasoning of different kinds, thus suggesting that this kind of training can foster mentalizing abilities. Similar results were achieved by training autistic subjects (Paynter and Peterson, 2013). Interestingly, there is evidence that thought bubbles can be successfully used for mentalizing tasks that are passed before the fourth birthday as well, and this holds for both typically developing children and autistic children, as in a study by Kerr and Durkin (2004): here, children familiar with thought bubbles were able to identify their content as thoughts of the relevant character and were able to understand that those could be potentially false.

In these thought bubble experiments, the content of somebody's thought is directly represented in a form that is very easily grasped, making something that is supposedly "invisible" directly traceable. In the context of LALAS, this is also interesting because the association between an agent and a "point of view", i.e. a representation of a state of affairs, is represented non-linguistically (so, not through a *description*) of a state of affairs, but perceptually. Assuming that chil-

dren make use of abstract schemata that allow them to more easily represent associations between agents and state of affairs represented (in one way or the other, but with descriptions in the case of language), the thought bubbles seem to be training exactly this representational ability. In thought bubble studies, there is a substantial aid to false belief reasoning because one can actively operate with these associations (moving the images in the thought bubbles, seeing the connection between the agent, the representation, the object, etc.). This is likely to be excellent training for false belief reasoning, because familiarity over these associations is gained through folk-psychology narratives. This kind of engagement is likely to be training the same abilities that are trained by language, in developing an efficient way to work with associations between agents and states of affairs[53].

6.4 Belief-reasoning strategy ruled out in 4-year-olds?

It has been repeated more than once in this chapter that operating over schemata learned through language is not necessarily the same as forming concepts of belief and mental states, and that on the contrary language makes it easier to deal with pairs of agents and states of affairs. One reason to think this is indeed the case has to do with Fabricius et al. (2010) studies, where 4- to 5-year-olds are tested in more complicated versions of the false belief location and container task. What is interesting when considering Fabricius' paradigm is that what makes the tasks more difficult is the fact that the FBT is turned into a true belief task, where the location of the object, for example, is changed twice without the character being present. (So, in the equivalent of Maxi's story, that would mean that the chocolate that Maxi thinks is in cupboard A is removed from the cupboard, placed in cupboard B, but then put again in cupboard A before Maxi comes back in.) What makes Fabricious' results interesting[54] is the pattern of the results. When asked where Maxi will look for the chocolate, predictably, 3-year-olds pass the task, assuming Maxi will look where the chocolate actually is (this is in line with the fact that they fail the false belief task); equally predictable

53 Thanks to Guy Dove and the audience at my talk at SSPP 2019 for pointing to this congruence between my idea and the experiments with thought bubbles, which I also incorporated in Berio (2020a).

54 In a previous experiment (Fabricius and Khalil, 2003), the same results were achieved, but they were subject to a lot of criticism by Perner and Horn (2003) because of the many questions used, which potentially caused confusion in the children. This is why I focus on the more recent version, in which the questioning was simplified.

is the fact that 6-year-olds have no problem replying to the question correctly. However, strikingly, 4- to 5-year-olds fail the task, and they tend to predict that Maxi will look in the cupboard where the chocolate was initially hidden.

This, it has been argued in Fabricius et al. (2003); Hedger and Fabricius (2011); Fiebich (2013), is a good reason to assume that 4-year-olds indeed might not fully understand belief, because they do not seem to be able to attribute a true belief in this case. Rather, they rely on a easier strategy based on perceptual access reasoning, based on two rules: (1) seeing → knowing and (2) knowing → getting it right. This strategy, which is less complex than relying on belief reasoning, allows the 3-year-olds to pass the false belief task, because they correctly assume that Maxi will be wrong, but it does not help them correctly predict behavior in the case where Maxi was not present but will nevertheless get it right.

This kind of result is in line with LALAS in predicting that operating with concepts like *belief* is not what guides the 4-year-olds passing the false belief task. If Fabricius is right about PAR, moreover, and the idea that heuristics like assumptions about ignorance and presence guide the predictions made by 4-year-olds, this also supports the prediction that LALAS makes, that schemata are applied in the context of specific embedded narratives. In folk-psychology narratives, it is very plausible that the child learns exactly this kind of heuristic: when something changes, and somebody does not know about it, they are usually wrong and they get it wrong. It is no wonder that in such a context the child would predict an erroneous behavior, as this fits with the idea that explanations are given as embedded in the general practice of explaining behavior according to the patterns and rules normally used in a social context. One might wonder why, if the child has indeed developed schemata for dealing with associations between agents and states of affairs, it would be the case that these associations cannot override the heuristics of PAR.

However, the design of the task might make the actual relation between Maxi and the state of affairs non-salient: something happens, which does not really change anything in the scenario. Nevertheless, the child is asked a specific question about the prediction of behavior. There is the concrete possibility that the child gets confused because the explanation does not seem to be bound to a relation between Maxi specifically and the state of affairs, but instead appeals to some other principle: this might actually be the reason why the child "decides to" rely on the heuristic, and does not answer correctly. In other words, pragmatic reasons might lead the child to answer the question in that way. This idea is confirmed by the studies in Oktay-Gür and Rakoczy (2017), which were specifically designed to test whether the true belief task of the kind used by Fabricius (2010) is undermined by how pragmatically odd the situation is for the child.

In the Fabricius task, they argue, the child is asked a trivial question about a belief of a person that has perfect access to the situation; the child, in this situation, might be led to believe themselves just wrong and to have missed something. Moreover, beliefs are not salient enough in the situation: while the FBT requires the child to actually compare their own belief with that of the other actor, in this case something is asked about a non-contradictory situation, which might confuse the child, who is used to different inquiries. Following these hypotheses, Oktay-Gür and Rakoczy (2017) tested if the same difficulties with true beliefs were retained in a less awkward pragmatic situation: in this case, there were two protagonists involved, one witnessing the event of the location being changed, and the other not witnessing it. In this case, children aged 4 had very little difficulties attributing the true belief to the agent that did witness the change of location, and their performance correlated with their ability to attribute the false belief. Naturally, this speaks for the fact that the task as designed in Fabricius might have been too hard for the children and, pragmatically speaking, difficult to handle, and thus supports the hypothesis that a lack of plausibility and familiarity might be the reason why children did not use the information available according to LALAS. However, one might think, as in this case no perceptual access was absent for the true belief condition, the children in this task might have still used the PAR strategy for making a difference between true and false belief characters. Even so, in the other experiments in the study, Oktay-Gür and Rakoczy (2017) show that, if the true belief task is not simplified in terms of relevance and salience of mental states, but the "true belief character" lacks no perceptual access to any event (e.g. they leave the room, but nothing happens while they are away), 4-year-olds still fail the task in the same way they do in Fabricius (2010): this, it is argued, speaks harshly against the idea that a perceptual access explanation is indeed resolutive enough. While PAR might indeed be relevant for some folk-psychology narratives, then, and in this sense support LALAS's predictions, it does not rule out the main core of the claim made here, i.e. that 4-year-olds' performance improves partially thanks to the generalizations aided by language use and acquisition.

6.5 LALAS and the bigger picture

At this point, an important issue that should be tackled is how such a model can fit a more general model of interaction between language and thought. After all, LALAS's prediction is that language actively enhances cognitive performance in determined mentalizing tasks, and it does it by providing additional clues to pos-

sibly abstract patterns and schemata. In what follows, I will position LALAS's claim within a more general picture of the interaction between language cognition by recalling the analysis that was made in chapter 2.

Let us start by looking at the table presented at the end of chapter 2. Recall that the different views analyzed at the beginning of the book allow for different roles for language and different interactions with other cognitive abilities. As formulated, LALAS claims that schemata derive from linguistic input, and that engagement in narrative practice is what allows the child to familiarize themselves with how forms of explanations are to be used and how to structure them.

Table 6.1: Theories' claims and predictions.

	Language directs attention	Language aids memory	Language "directs" thought	Language as representational format	Language as propositional thought
LASSO	√	√	√	x	√
LFH	√	√	√	x	x
"Inner speech"	?	√	?	√	√
Quinean bootstrapping	x	x	x	√	√
Theory–theory	√	x	√	x	√
Rewiring hypothesis	√	√	√	√	√
LALAS	√	√	?	?	?

In this sense, language is thought to be:

1. Directing attention: as argued, using the same linguistic structures guides the retrieval of meaningful patterns and invites comparison between different situations through structural alignment;
2. Aiding memory: schemata, albeit non-linguistic, provide a way to better store new representations;
3. Providing a new representational format; however, the new representational format is not necessarily linguistic, but just a higher level of abstraction.

However, LALAS makes no prediction about whether or not language is the means through which we gain control of our thoughts or is per se the representational format of propositional thought.

Considering point 2, notice that this is also compatible with the idea that linguistic knowledge helps to store meaningful information about the folk-psychology narrative at hand; in other words, it is perfectly possible that having language mastery can also help the infant to keep in mind several facts and perspectives in a complex folk-psychology narrative in linguistic form, for example remembering specific strings like "She thinks that it is the case that *x*, he thinks that *y* " and so on. However, this is not the specific prediction of LALAS, and while one could expect this to be the case in long complex narratives, the role of the schemata in this case would be to keep track of associations in a less detailed and non-linguistic format. On the issue of whether language is what constitutes propositional thought, LALAS is silent, since the idea is that language interacts with pre-existing resources in order to enhance existing mechanisms. The same can be said about language as "controlling" thought: the mechanism proposed here is theoretically neutral on the question whether language is what determines "endogenous" control over thought, as argued for in Tillas (2015a).

There are at least two possible interpretations of LALAS that can be explored. One is very specific, and puts LALAS to work in the LASSO framework. The second is more general in the sense that it makes a weaker claim, inserting LALAS in the theoretical framework proposed by Clark and others, i. e. that language enhances cognitive abilities in several ways. I will explore the two options in this order in what follows.

6.5.1 LASSO and LALAS

Recall what has been explained about LASSO's theory in chapter 2: the idea is that propositional structures in thought come from language, and are achieved because thought can piggyback on language. Conceptual units are associated in a network with labels and modal representations, where the weights are determined by frequency of associations and a net is built over Hebbian mechanisms. This allows for conceptual units to be structured and reorganized in propositional form, thus granting the possibility of propositional thought in the realm of an associationist framework.

The key detail of an association-based analysis like this is the notion that linguistic labels can carry information about the structures they appear in – in other words, syntactic information has a place in the network. This is what forms the foundation of the claim that language provides enough scaffolding for thought to be propositionally structured: sentences, in which words appear,

are syntactically structured entities, and are associated with conceptual components in a way that mirrors their unity and structure.

Now, let us try to apply this to the specific case of syntatic structures that involve mental state verbs and their use in FBTs. Exposure to the same constructions implies, in such a picture, information about the type of argument that a certain linguistic entry, in this case a verb, takes. In this case we will have something that should go, roughly, along the following lines. Consider potential linguistic data the child is exposed to:

– *Nora thinks that there are smarties in the box.*

– *Sameer knows that you took his toy.*

– *Ian knows that Kara has already eaten candy today.*

In these instances, the mental verb takes a sentential complement as an argument. This strengthens the connection between this sort of linguistic entry (*think*, *know*) and another kind of linguistic entry – mostly sentences, which are what is embedded in a sentential complement. As Tillas (2015a) puts it (p.227), "[...] a subject learning a new word also learns (about) its relations to other words"; this is because, rather than appearing in isolation, words appear in a linguistic context. Sentences that follow one of these verbs will typically express a state of affairs, or an event; in other words, they will be a description, and in this way they will be understood by the child, who is used to such stimuli in everyday interaction. After all, lots of interactions with children, from a very young age, provide them with descriptions of the world surrounding them. These verbs, then, will be associated with the upcoming description of a situation. So far, this fits with LALAS assumptions. Thanks to exposure to language and learning, in LASSO this implies forming connections between different concepts too. Note that one might say a "class" in this sense emerges: verbs that take complement structures as arguments. (It is of fundamental importance to remember that an associationist account like the one presented does not exclude the representation of grammatical relations.) At a conceptual level, what we will have is the representation of a mental state emerging as the result of the encounter with so many different situations in which the same basic elements can be individuated: a description of a situation, and an agent. This representation will be connected to the one at the linguistic level, which is highly structured given that the information regarding the structure is, as said, stored. This is potentially what provides the child with the structured representation described by LALAS: in this case, this forms directly because of language, thanks to associations that restructure already existing representations. The structure learned at the linguistic level is connected with a conceptual representation of a mental state towards a spe-

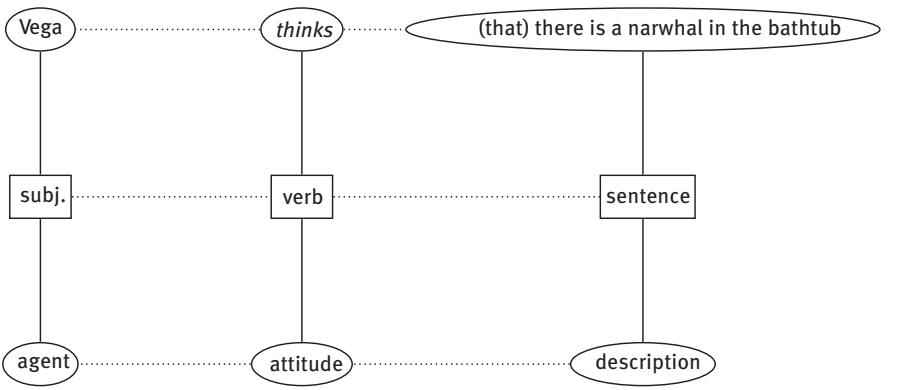

Figure 6.1: From the linguistic input to the recombination of conceptual structures, with the underlying syntactic input

cific situation; this, in turn, is structured in a way that mirrors the linguistic stimuli, but it does not have to be rehearsed in a linguistic form to be used in subsequent tasks: the proposed story gives a way in which the representation can be formed and it accounts for the role of the syntactic role in the formation, without assuming a level of representation different from the conceptual level or the NL one.

The idea in applying LALAS to LASSO, then, is to exploit the associations between labels and conceptual structures, in order to assume that abstract conceptual structures are formed thanks to the linguistic input. In this case, LALAS can be considered to be the thing which provides the child with *propositional thoughts* related to folk-psychology. Notice that, while it might seem that this is the same prediction that LALAS alone makes, this is not the case: applying LASSO in this case means implying a series of things also implied by Tillas, which include the fact that propositional thinking (in and outside folk-psychology) derives from language, and that it structures a level that is *conceptual* in nature. Also, it means assuming that direct connections are formed between the linguistic stimuli and the conceptual structures; this is not necessarily the case in every interpretation of LALAS, as (as has been explained) language might play the role of providing the model for an abstract schema, without playing the role of rearranging concepts in a propositional form, as indicated by LASSO.

6.5.2 LALAS and language as a tool

Inscribing LALAS in LASSO is a theoretical possibility, but not the only possible out-come. In general, LALAS is compatible with views that were described in chapter 2 as the *rewiring hypothesis*, and in general with the idea that language expands our cognitive tool kit.

Recall the cognitive scaffolding view originated by Vygotsky (1962), according to which higher mental functions (e. g. interaction with caregivers and the acquisition of language abilities) are made possible by language. Clark's version of the rewiring hypothesis goes a bit further in assuming that a function (among others) of language is to allow for the formation of representations that are more easily recallable and reusable: in other words, language provides mechanisms for further abstraction, on the one hand, and better memory on the other. In a nutshell, LALAS schemata are thought to be exactly that kind of tool, as they are used for the faster and more efficient processing of associations.

A further possible role for language is recognized by Clark (2006) as "labelling as short-cut": the idea is that using labels for different perceptual representations makes other representations available faster. More generally, this goes in the direction of a role for language in acting as "glue" for cognitive representations of different kinds, allowing one to cluster perceptually different input; this is, then, exactly what I believe the role to be for mental state terms and the syntactic environment in which they occur, as they allow one to find analogies across situations, and to integrate the resulting generalizations in a schema that can be used when engaging in socio-cultural practices like folk-psychology narratives.

This idea goes hand in hand with that of *hybrid thoughts*, which implies that language can interact with language-independent resources in the formation of new skills. In the case of numerical skills, a linguistic system might be interacting with a non-linguistic system (see 2.3.1 in chapter 2): notice, however, how, in the case of LALAS, one might say that linguistic information and abilities also work to enhance and improve a system already in place, i. e. that of comprehending associations between agents and states of affairs and understanding goal-directed actions. This is possibly more comprehensible if one thinks of the hypothesis made by Clark (2013) and Lupyan (2012) about the role of language as a constraint in guiding *top-down* inferences on the cognitive processing of perceptual information. In other words, words and clusters of words – sentences – can be fundamental in influencing low-level processes regarding how the stimulus is processed. The linguistic information that comes with the use of mental state terms, then, might be what is used to form predictions about the kind of further input received. Let us imagine, for instance, that hearing the mental state verb in

"Lea thinks" generates expectations, on the one hand, of hearing a full sentence as the rest of the sentence and, on the other hand, of having to represent a state of affairs. The combination between hearing mental state verbs in a determined syntactic context and hearing them as the result of the description of a state of affairs or situation might indeed be what brings one to form generalizations that are at the core of the creation of the schemata that are proposed in LALAS, as an important ability for the child to pass false belief reasoning tasks. In this way, language shapes the expectations about the kind of representation that needs to be recruited, according to Clark (2013): this might be exactly the kind of mechanism, then, that brings one to generalize over a large number of different situations and abstract from the relevant pattern that provides the child with the schema used in mentalizing.

This kind of idea is not only supported by Clark, but fits with the general take on cognition that sees language as enhancing computational abilities, one way or the other. A recent example is that of Dove (2017), who argues that language does not only provide scaffolding for further thinking, but also has a fundamental role in transforming how we conceptualize objects and external stimuli. This resonates well with the idea that in the case of LALAS there is a direct influence of language on how we relate to social situations, not only because of communicative constraints, but also because of how language shapes the stimuli and reveals (and to some extent creates) patterns in how certain situations are dealt with in the social environment. It is also very in line with the idea proposed by Camp (2009) that syntactic information provides a window into the potential combinatorial possibilities of thought; language, on this approach, provides the possibility of stimulus-independent and recombinable thought, granting connections between thoughts that would otherwise not be combined (in this sense, in a way that is similar to Tillas (2015a)). This is also possibly compatible with LALAS, since the main idea is that one thing language allows one to do is to form systematic relations between different perceptual situations, and to use these relations in productive new ways that were not available before.

Finally, note that Clark also thinks that linguistically formulated thoughts can be the object of further thinking and reasoning, and that in this case language can allow increasing abstract thought. I am generally of the idea that second-order folk-psychology depends on a similar mechanism: when the folk-psychology narratives become increasingly complex and involve increasing levels of recursion, language is probably the best candidate for it. However, recall that LALAS is a hypothesis about what happens at a lower level of processing, and it is concerned with how one goes from failing to passing a simple explicit FBT. However, it is a hypothesis compatible with the idea that more sophisticat-

ed forms of mentalizing require language as the unique format of representation. While I do not intend here to subscribe to the entirety of Clark's research plan, including predictive coding and his general view of cognition, I think the approach to language he proposes has significant resonance with LALAS.

6.6 Summary and conclusions

This chapter has been dedicated to elaborating a proposal about how language can indeed be a meaningful resource in solving mentalizing tasks, building on the previous chapters. I am not of the idea that this is the only way in which language can help the learning child to pass the standard explicit FBT; however, I do believe that the mechanism I name LALAS can be a fundamental part of the story.

I described LALAS as a mechanism which helps the child to form abstract schemata that assist in memorizing new associations between agents and (descriptions of) states of affairs, and in using them in subsequent tasks. The mechanism is thought to be at play when the child reaches the necessary mastery of sentential complement structures and syntax in general. Being able to master sentential complements gives the child the means to individuate meaningful patterns in folk-psychology explanations and to form an abstract associations between agents and descriptions. Once familiar with how these associations are used in the context of folk-psychology, the child slowly masters this resource by applying it in false belief task comprehension, which allows her to form quicker associations, to better spot the relevant associations, and to use them for prediction. While the schema is itself non-linguistic, it forms with the aid of linguistic input, with syntax playing an important role in providing a way to parcel the stimuli and reveal the underlying pattern. Structural alignment is part of this story, as language invites comparison between different situations that have a description of them in common; this, as underlined, fits with several approaches to syntax acquisition and syntactic parsing and it does not force me to commit to an associationist account of syntax. While syntax helps in individuating the pattens as described, it is not necessary to adopt an I–language nor a logical form approach, making LALAS theoretically less demanding than than might be thought; in general, schemata like those postulated here can be integrated within more general situational models that make computing complex perceptual and social situations easier and more efficient.

LALAS can be adapted to at least two different "double mechanism" accounts of mentalizing, with slightly different outputs. In the case of Apperly (2011), LALAS can be the mechanism underlying better storage abilities when

the higher-level mechanism is at play; in the case of the model in De Bruin and Newen (2014), LALAS can give a complimentary explanation of how the operative system gets trained in operating on cognitive perspectives. As stressed, my proposal borrows a lot from many of the approaches described in the previous chapters, and (as argued) provides a way of solving some of their shortcomings. At the same time, I concede that schemata like those described in LALAS might form through other mechanisms, or that at least they could be facilitated by specific training, like thought bubble training, showing that LALAS is flexible enough to account for how language and other cognitive skills might result in improvement in the same domain.

LALAS is particularly compatible with those accounts that, in chapter 2, were shown to be particularly convincing in light of evidence that goes well beyond mentalizing. The fact that very different data support the idea of language as a powerful cognitive tool, and that LALAS fits in this general model is certainly an advantage of my proposal. Potentially, LALAS could be integrated in LASSO. Since LASSO makes demanding assumptions about the nature of cognition, I underlined how one might prefer a "language as cognitive tool" framework and how, even then, LALAS could be part of the story. In the long run, a complete picture of cognition would need to take a side, but I am convinced that the fact that my account is compatible with more than one option is at the moment a strength.

The advantages of LALAS are of various kinds. Firstly, it provides a way to combine the theoretical frameworks that stress the role of folk-psychology narratives (like Hutto's or Nelson's frameworks) with accounts that underline the role of specific language features, like complement structures. The idea is that this combination can, on the one hand, overcome the problems faced by accounts like the syntactic bootstrapping hypothesis, and on the other hand, give a solid concrete complementation to accounts that focus on the role of social interaction. The syntactic bootstrapping hypothesis faces problems deriving from not having a straightforward account of how linguistic information interacts with non-linguistic information, which can bring one to the assumption that mentalizing is carried out in linguistic format: as argued, such a position entails some difficulties, both theoretical and empirical. At the same time, cross-linguistic data seem to pose a challenge to these theories, since they rely on specific constructions that are not universally acquired at the same time. LALAS solves these problems by assuming that if specific linguistic information, including syntactic information, is playing an important role in boosting cognitive power, at the same time the schemata used in folk-psychology are to be considered to be embedded in social narratives. It is the fact that they are recurrent and frequent means of explanations of behavior that makes these narratives so use-

ful in the context of mentalizing, and in solving the false belief task. Finally, this approach accommodates the findings that stress how social interaction, narrative abilities, and mentalistic language used by caregivers all contribute to the formation of more effective skills at play in mentalizing.

Part IV: **Conclusions**

Chapter 7:
Conclusions and outlook for future research

7.1 The scope of this work

This book started with a general picture of the debate regarding the relation between language and thought. While less specific than the rest of this book, the first chapter served the fundamental function of framing the debate in the context of the influence that learning a specific language has on the development of other cognitive skills.

The choice to focus on specific processes, language items and abilities (namely, abilities in social cognition and the acquisition of related vocabulary) stems precisely from the analysis conducted in the first chapter. If, on the one hand, it makes sense to investigate what the relation between language and thought is, in light of the more general picture of cognition, I also argued in chapter 2, on the other hand, that a sensible approach includes considering, specifically, which kind of thinking could be affected by which kind of linguistic abilities. This is in line with the pluralistic perspective advocated by Beaulac (2014) and especially with the empirical evidence described in chapter 2. The choice of dedicating the central part of this book to mental state verbs and the impact of their acquisition on social skills, then, follows from this particular stance on what the general aim of a picture of the relation between language and thought should be.

It is in the same spirit that I chose to give special attention to a specific category of nouns and properties, even in the most general of the chapters, as I decided to focus on color terms. This also had the function of proving that the inquiry is worthwhile in the first place because, despite the communicative views about language that are still present in the debate, I believe there is evidence that the acquisition of language does have an impact on non-linguistic cognition. In this sense, the empirical reviews presented in chapter 2 have served the purpose of clarifying, on the one hand, how fine-grained the analysis of the interaction between linguistic and non-linguistic processes needs to be and, on the other, how already present evidence supports the idea of language as a powerful tool in cognition.

The analysis in chapter 2 also excluded models of language and cognition like Carruthers' (2012) on the basis of theoretical and empirical considerations, and highlighted some of the advantages in other models, especially the associationist models offered by Lupyan and Tillas (Lupyan, 2012; Tillas, 2015a), but also partially the model proposed by Clark (Clark, 2006). There is a strong

https://doi.org/10.1515/9783110748475-008

sense in which LALAS as proposed in chapter 6 does indeed go in the direction that is prescribed by these accounts. Linguistic input acts as a cognitive enhancement for mentalizing skills because it allows one to form abstract schemata which are used in some versions of the false belief task, and thus explains at least partially the increasing abilities of children who already master mental state language. While, as specified, LALAS does not predict that language is the only way in which these cognitive structures can form, it does predict that language is a valuable resource for mentalizing. This is fundamentally in line with the data presented in the rest of the book.

However, it is important to specify that LALAS is somewhat neutral on the role of language in terms of the representational format for thinking. As a hypothesis on how acquisition of linguistic structures boosts abilities in the growing child, LALAS predicts that the combination of social interactions, folk-psychology narratives and structured language input allows the child to form more efficient abstract generalizations that help to navigate specific tasks that rely on them. It does not, however, directly answer the question of whether we need to consider language as the way *endogenous control* is gained over thinking, as is predicted by Tillas, or as the *lingua franca* of cognition, as predicted by Carruthers, or as an instrument of cognitive self-manipulation as described by Lupyan and Clark. At the same time, LALAS is not an explanation of how children form a *theory* of mentalizing; while I do not exclude that theory-like models are to be considered for sophisticated forms of mentalizing, this is not implied by LALAS. As stressed more than once, LALAS is concerned not with the acquisition of mental state concepts, but rather with the idea that language provides useful tools for keeping track of relevant information, in a way that possibly precedes the formation of folk-psychology conceptual structures. It is very plausible that the fairly sophisticated full-blown concepts of belief, thought, and knowledge are language based in more than one relevant sense, and in this case it is possible that Carey and/or Gopnik's take (Carey, 2009; Gopnik, 2001) on the matter are indeed right.

In chapter 3, I touched on several issues connected with language acquisition, and highlighted the main characteristics of mental state verbs. Having a clear idea of when they are acquired, what the main semantic and pragmatic properties in English are, and how this relates to their syntactic acquisition and to data coming from cross-linguistic and cross-cultural studies is essential for having a clear picture of their impact on cognitive development. The contrast between the informational change and the conceptual change approaches (Gleitman, 2009; Gopnik, 2001) has been presented because it helps highlight the challenges that underlie the acquisition of, for example, mental state verbs, and because it is indeed connected with the issue underlying mentalizing as

well; positions like Gopnik's combine the assumption that semantic development precedes verbal expression with the idea that a theory involving concepts of mental states lies at the core of some mentalizing skills (Gopnik and Astington, 1988; Gopnik and Meltzoff, 1997; Gopnik, 2001). None of these assumptions is endorsed in this work, and I have argued on the one hand that there are good reasons to assume that syntactic and use-based theories of acquisition are probably right in assuming that learning the meaning of verbs is a more complex matter, and on the other hand that it is not necessary to assume a theory-like approach in young children when they pass the false belief task at age 4. Independently, on the specific positions, an explicit aim of this book was to connect the literature on mentalizing and language acquisition to some of the linguistical issues underlying the acquisition of verbs, cross-linguistic data, and specific linguistic structures. While some theories like the syntactic bootstrapping mechanism proposed by de Villiers (de Villiers, 2005) already connect the two issues, these are not usually put in relation with larger bodies of literature in the philosophy of mind, or inscribed in a more general developmental framework of mentalizing. Moreover, this is to my knowledge the first time that these theories have been systematically placed in relation with broader pictures of the interaction between language and thought in cognition. In these senses, I think this work makes a new contribution to the debate. In chapter 3, it emerged that, on the one hand, folk-psychology explanations are culture-specific, and that the use of mental state verbs varies with them; and on the other hand that, if there is semantic and syntactic variability across languages, it is still possible to identify what makes these verbs generally hard to learn, on the one hand, and what ties them to specific communication practices on the other. Moreover, the chapter served as a demonstration of how linguistic debates like that concerning the role of syntactic information and semantic pairing are relevant when one is considering how language development impacts psychological processes, even in a philosophical perspective, as many philosophical themes emerge in the literature dealing with a spectrum of issues, from the relation between conceptual development and language expression to the relation between the semantics of verbs and their pragmatic and communicative functions. Especially relevant is the fact that, from the review in the chapter, the time-line for the mastery of mental state verbs in English emerges clearly, with fairly solid and reliable results.

In chapter 4, I firstly delineated the issue of mindreading itself, highlighting what the main issues in the literature are and what the data on mentalizing in general look like, before digging into the literature explicitly addressing the issue of the interface between language and mentalizing. Of the several, intertwined issues that are present in the literature, some are related to domain specificity and how "special" mental representations of mental states are, whereas

others are issues related to the correct developmental framework. The fact that a distinction between implicit and explicit false belief tasks emerges so clearly from the literature is of fundamental importance and, while I raise some doubts about the ecological validity of some of the studies, a lesson to be learned from the first part of the chapter is that an account of the impact of language on mentalizing abilities needs to take into consideration how it can fit into a more general account of the development of mentalizing-related skills. The picture painted by the available evidence is that of a complex set of abilities that interact to allow (Western) children to go from failing the false belief task to passing it with flying colors in less than a year. An account of the role played by language has to fit this picture, taking into consideration how different skills and abilities interact with each other. The data emerging from the empirical analysis in chapter 4 strongly support two things: that exposure to and the use of mental state terms and structures in which mental state terms are used is connected to false belief reasoning abilities, in possibly a causal way; and that narratives, narrative skills, and pragmatic abilities also play a relevant role in the development of mentalizing skills. These conclusions have been reached through an in-depth analysis of a large body of literature, which makes clear the fact that language interacts with other skills (including perspective taking, pragmatic abilities, attention and goal-directed action) in a complex way; however, the fact that many intertwined factors are related in the case of mentalizing skills does not mean that the role of language can be dismissed, as is evident once studies analyzing the role of syntactic abilities, semantic development, and pragmatic skills have been considered. This is also supported by the clinical data, since a relation between language abilities and mentalizing abilities is found in autistic children, schizophrenic patients and perhaps in SLI, even if the results seem to be less clear in that case. The fact that specific language impairment seems mostly to imply a delay in mentalizing skills, and that aphasia seems to leave some mentalizing skills intact, is also of great importance because it supports the hypothesis that, while linguistic input might indeed be enhancing and accelerating the development of mentalizing skills, there are strong reasons to doubt that false belief reasoning has to be conducted in a linguistic format. A rather strong case for the importance of linguistic input, on the other hand, comes from the literature on deafness, since considerable delays are found in false belief reasoning for deaf children with hearing parents, along with strong correlations between language abilities, caregivers' vocabulary and false belief performance.

Chapter 5 was dedicated to an analysis of the existent accounts of language and mentalizing, which were evaluated in light of the empirical data presented in the previous chapters. While data about language acquisition presented in chapter 3 serve as a back-ground for understanding the arguments presented in chap-

ter 5, the data presented in chapter 4 are used to highlight the strength of the accounts that are presented in the literature. The debate around the issue is, if not wide-spread, quite intense, and the chapter reconstructed some of the main arguments around it, for example the tension between the syntactic bootstrapping hypothesis and the objections raised on the basis of cross-linguistic data by Perner et al. (2005), and backed up by analysis by Van Cleave and Gauker (2013). In general, approaches that focus on specific syntactic components are related to more general Vygotskyan views, like that of Garfield et al. (2001), which is underspecified in many ways but still brings attention to a fundamental fact, i.e. the interaction between social development and linguistic input. The stress on the social component is also something that emerges in the analysis by Montgomery (2005) and Nelson (2005),which also have in common with Van Cleave and Gauker (2013) the fact that they stress that recognizing a role for language does not necessarily imply that *concepts* are involved in the instances of mentalizing that become possible around the fourth birthday. This is relevant because, as was argued in chapter 6, there are good empirical and theoretical reasons to think that this is indeed not necessary. Montgomery (2005) also has a strong point, which is similar to the analysis in Hutto (2009), in calling attention to how the use of mental state terms and syntactic structures is ultimately related to specific social situations. Whether or not one decides to appeal to Wittgensteinian rules and language games, data presented in chapter 5 support the idea that linguistic input has to be considered in its social context in order to have a complete picture of how acquiring language and expanding one's vocabulary can actively contribute to enhancing false belief reasoning. Hutto and Bermúdez (2009) stress two factors that acquire a certain importance in my own proposal; the role of folk-psychology narratives, on the one hand, and the role of schemata and social scripts, on the other. Theory-like mindreading, for both Bermúdez and Hutto, is a rare case. The comparison between the two views is fruitful because it shows how, by departing from the same consideration about how rare the use of explicit rules and representations can be in mentalizing, one can argue for two very different views: a narrative based one in the case of Hutto, and a linguistic-because-propositional one in the case of Bermúdez. In this sense, Bermúdez's view is also comparable with Gordon's ascent routines, despite the simulationist approach that characterizes Gordon's framework. Analysis of this view made quite clear that it is of fundamental importance, not only to take into consideration data that positively underline a relation between mentalizing and language, but also data, like that related to aphasia, that exclude a constitutive relation. Aside from the emphasis on socio-cultural interactions and practice, other important lessons can be learned from other positions analyzed in chapter 5, among which the specific mechanism of structural alignment

proposed by Baldwin and Saylor (2012) is an interesting one. As further elaborated in chapter 6, structural alignment does provide a concrete mechanism for how language can contribute to making sense of complex linguistic input. Baldwin and Saylor suggest that it is a way for the child to form mental state concepts; as articulated in chapter 6, I argue that structural alignment has the potential to be used to explain false belief understanding in a milder sense, without assuming it is conducive to the formation of concepts, at least not when the explicit FBT is passed. The analysis carried out in this chapter had the purpose of underlying the strengths and weaknesses of the theories available, and to make clear how what is needed is an account that, on the one hand, takes the data into consideration, and on the other, fits a more general developmental story.

Chapter 6 was dedicated to my own proposal for the role of language in supporting false belief reasoning, partially developed in Berio (2020a). The account presented as LALAS does not cover any possible influence of language on any sort of social cognition ability, but rather focuses on what role mental state terms used in syntactic contexts and in socially and culturally determined narratives can have in enhancing performance in explicit false belief tasks. A fundamental claim in LALAS is that the syntactic component and the communicative impact of language are not to be considered as two alternatives when it comes to realizing what makes the difference in the development of mentalizing skills; on the contrary, if I am right, they represent clues of a different nature that support the development of the same skills, which allow the child to navigate the social world long before they acquire the conceptual representations that adults tend to associate with folk-psychology. The first part of the chapter focused on the claims made by LALAS, which rely on structural alignment as proposed by Gentner (Gentner, 1978; Gentner et al., 2011) as a mechanism driven by language that allows for comparison of very different input. The fundamental idea is that being exposed to linguistic input provides the child with sufficient material to form generalized structures, associating an agent with a description of a state of affairs, which I call schemata. These schemata are embedded in specific cultural practices that include folk-psychology narratives. Another central claim made by LALAS is that how these structures are used and their explanatory power in terms of relating with other people behavior is also a fundamental part of the story of how false belief reasoning improves so dramatically in children around the fourth birthday. In explaining my account, I also made clear how it is compatible with different theories of parsing and how it is positioned in relation to theories of syntactic acquisition. In relating my view to other previously presented accounts, I stressed how I do not assume that 4-year-olds possess full-blown concepts of mental states, nor that they rely exclusively on syn-

tactic information to solve FBTs. On the contrary, LALAS leaves room for many interrelated factors to collaborate in the formation of false belief abilities. This is very clear in the second part of the chapter, where I related LALAS to two different accounts of mentalizing that rely on two different systems, De Bruin and Newen (2014) and Apperly (2011). I argued that LALAS can fill the gap in both accounts when it comes to indicating exactly what makes language a reliable resource in the development of false belief reasoning. I also made the specific point that LALAS predicts that schemata can form through and thanks to language, but that the same abstract format for storing associations might emerge thanks to other kinds of training focused on associating agents with state of affairs, like in the case of speech-bubble studies. I concluded the chapter by making a point about the bigger picture, i.e. by relating LALAS to views about language and cognition in general. I gave two concrete examples of how LALAS can be inserted in a theory of cognition: I related it on the one hand to LASSO Tillas (2015a) and on the other to Clark's idea (Lupyan and Clark, 2015). I argued that LALAS is compatible with both an association-based account and with the idea that language partially *rewires* cognition, thus situating my account within a broader picture.

In what follows, I will discuss what I think are promising directions for this work. In particular, I will give some indications of how I think direct empirical evidence for LALAS can be found. While the section is rather programmatic in indicating a possible path, where a lot has to be established and further elaborated, I think there are particularly relevant considerations for the present work in indicating a promising direction for research.

7.2 LALAS and empirical evidence: what we have and what we lack

In this section, I will briefly describe the kind of empirical evidence that could support the account I proposed in the last chapter[55].

There is a sense in which LALAS is built on existing empirical evidence, since it is designed to accommodate the empirical results on the influence of language on mentalizing abilities in light of the difficulties that other accounts face. A fundamental concern in this book has been that of providing an account that not only explains and further specifies the importance of linguistic input for

55 In Berio (2020a) I present some indications for future research; I further elaborate here and go more in depth on what kind of evidence would be needed to support my hypothesis.

false belief reasoning, but could also be situated within a larger view of cognition, which is something that, as I argued in previous chapters, other accounts do not always do. The evidence brought by de Villiers and colleagues and described in chapter 4[56], which supports a role for sentential complement structures in the development of false belief reasoning, is perfectly in line with LALAS: one of the main predictions of my account, as a matter of fact, is that structured input in language does provide the child with valuable clues for the construction of useful schemata employed in false belief reasoning. The same holds as a consequence for the data which suggest a role for one's mental state lexicon, both in production and comprehension; if it is the case that semantic and syntactic skills help the child to individuate the patterns in linguistic input, and that training to recognize these structures in the context of explanation in folk-psychology narratives and stories is of fundamental importance not only for the mechanism, but also for using them in the right context and for facilitating false belief "reasoning" when it is required, then LALAS has empirical support. This seems to be the case given the evidence presented in chapter 4[57].

Another prediction LALAS makes is that lesions or conditions that interfere with the syntactic and general linguistic abilities in adults will not necessarily result in problems in false belief reasoning; the reason is that false belief reasoning does not necessarily include manipulation of linguistic symbols, and false belief reasoning does not necessarily have to be linguistic. However, LALAS does not exclude that some forms of explicit belief reasoning have to be carried out in natural language, as Bermúdez (2009) suggests. On the contrary, LALAS is a hypothesis about how language helps in the relatively early stages of mentalizing, when false belief reasoning starts emerging. In this sense, LALAS is empirically supported by a variety of data: on the one hand, the data suggesting that aphasic patients maintain their false belief reasoning abilities (Varley and Siegal, 2000; Siegal et al., 2001; Apperly et al., 2006) and, on the other hand, the fact that conditions that delay language acquisition also show a delay in false belief reasoning (deafness with non-hearing patients, for example, Schick et al. (2007); de Villiers and de Villiers (2011)). In this sense, the take away of LALAS is that language constitutes a very valuable clue, but not that mental state terms and structures cause a cognitive revolution. This is also in line with the fact that, despite the positive findings from de Villiers and colleagues (de Villiers and de Villiers, 2011; de Villiers and Pyers, 2002; Hale and Tager-Flus-

56 For instance de Villiers and de Villiers (2011); de Villiers and Pyers (2002); Hale and Tager-Flusberg (2003).

57 For instance, Diessel and Tomasello (2001); Milligan et al. (2007); Wellman et al. (2001); Ruffman et al. (2003).

berg, 2003), many analyses suggest that language is a predictor of false belief performance in many different ways, including vocabulary acquisition and also interaction with caregivers and siblings (Perner et al., 1994; Cassidy et al., 2005; Happ, 1995; Tager-Flusberg and Sullivan, 1994a). Furthermore, LALAS predicts a role for narrative play and narrations skills in so far as it predicts that experiencing narratives and engaging in folk-psychology activity is of fundamental importance for the child to familiarize themselves, not only with linguistic input, but also with its pragmatic and social function. This is empirically supported by the data that stress the role of narratives (Ruffman et al., 2003, Ruffman, Slade and Crowe, 2003), storytelling (Symons, 2004), and pretense play (Nielsen and Dissanayake, 2010), but also with the correlation between pragmatic skills and false belief reasoning (Bosco and Gabbatore, 2017; Angeleri and Airenti, 2014). Finally, LALAS makes a more specific claim about how language helps generalizations and the formation of abstract schemata; in this sense, data in support of this prediction come from literature stressing the role of language in structural alignment (Gentner, 1978; Gentner et al., 2011; Gentner and Medina, 1998) and more in general with data showing the role of language in fostering generalization and aiding category formation (Lupyan, 2009; Lupyan et al., 2007; Lupyan and Mirman, 2013).

In summary, many of LALAS's predictions are supported by available data. However, it is clearly not a set case, and more specific evidence can also be collected. On the one hand, further data on the potential contribution of language in understanding abstract relations are definitely needed: while the data collected so far are more than promising, an exploration of this function of language would definitely be of use not only for LALAS, but also in delivering empirical support for many of the theories cited in chapter 2, among others Lupyan's account and Clark's view (Lupyan, 2012; Clark, 1996, 1998). If it is indeed the case that language allows us to cut-through our representational space to form new associations and to enhance our associations abilities, and also to allow for the formation of abstract conceptual and non-conceptual structures, this represents a milestone in marking the success of non-communicative views of language over others. As argued in chapter 2, I believe it is already the case that supra-communicative views of language have a clear theoretical advantage, but the investigation of exactly which linguistic components are relevant and play a pivotal role during acquisition and what kind of concrete impact they have on which specific abilities is of fundamental importance for a theory of cognition in general: as argued in chapter 2, accounting for specific mechanisms is the ideal direction for theories of language in cognition. While LALAS is a step in this direction, in delineating a specific mechanism at play which is fed by both linguistic structures and narrative and pragmatic practice, the ideal picture

would be one that relates this mechanism to analogous ones in other realms of cognition and development. In this sense, an exploration of the role of structural alignment and of linguistic structures in promoting abstract representations is indeed fundamental. This would ideally result in the integration of the conclusions in this work in a theory with a larger scope, in a way that relates this material to a broader picture of how language acquisition shapes cognition and cognitive abilities. As a consequence, the themes of mentalizing and mindreading would be better integrated in a theory of language and cognition.

Despite the great amount of literature on mentalizing, there is work to be done. For example, the cultures and languages briefly discussed in chapter 3 and chapter 4, which make considerably reduced use of mental state terms and engage less in folk-psychology than the average Western culture, are crucial. In this sense, I do believe that exploration of false belief reasoning abilities in these populations is of fundamental importance, not only to eventually reconsidering the potential importance and universality of false belief reasoning as an important paradigm of the human social skill set, but also because, in the case of LALAS, it can turn out to be illuminating with respect to which cognitive components interact in the development of mentalizing abilities. In this sense, it would ideally be possible not only to explore, on the one hand, how different folk-psychology practices relate to false belief reasoning, which is an exploration that has partially begun; the investigation could proceed hand in hand with verifying, for example, how non-linguistic training can help promote a performance in FBTs more similar to the performance of English-speaking children. Studies which make use of *speech bubbles*, then, could be crucial in assessing whether such a training can produce any improvement in false belief reasoning in populations that engage less in these folk-psychology narratives than do the most commonly studied languages and cultures. LALAS's prediction is that false belief reasoning in 4-year-olds is aided by language in two senses: as a source for structural comparison, and as a source of social interaction that leaves the child accustomed to and familiar with the relevant social practices. In this sense, investigating the extent to which one factor is compensating the other in cases in which they are not both strong as sources of evidence is fundamental for LALAS.

In conclusion, further evidence for LALAS could come from two different directions; on the one hand, more evidence for a role of language in aiding the formation of abstract structures could be collected, strengthening the account on the more "general" side and relating it to an idea of language in cognition. In the opposite direction, more specific data can be collected on how linguistic structures that are present cross-linguistically and practices that are present cross-culturally can interact to promote false belief reasoning.

7.3 Conclusions

This book set out to answer the following question: "To what extent does language acquisition influence our cognitive abilities, and more specifically how does this happen in the context of false belief reasoning?" In order to reply to this question, I drew on three different bodies of literature, providing an overview: of what the issues underneath the relation between language and thought are; of the kind of evidence there is for a role for language in mentalizing; and of how available accounts deal with the empirical evidence, on the one hand, and with the more general picture of language and cognition, on the other. The answer I provided focuses on the specific role of mental state terms in aiding false belief reasoning; as I underlined, I do not claim that this is the only possible way in which language can contribute to social cognition, let alone to cognition in general. My purpose was more modest, in providing a model for a specific mechanism that lies at the interface between our experience as communicators, our experience as mentalizers, and our experience as members of communities that engage in folk-belief practices. As underlined in this last chapter, many questions still need to be answered, both of an empirical and a theoretical kind. However, a step has been made in the right direction, in considering language as a communicative tool and a cognitive tool at the same time, and in delineating exactly what its role is in contributing to the development of mentalizing abilities that constitute part of our social behavior. I set out with the aim of being as considerate as possible to the cross-linguistic and cross-cultural data, but a lot of work has yet to be done in that direction as well; while progress is slow, there are reasons to be hopeful, since more and more data are produced in this sense. In conclusion, while many questions remain unanswered, I believe LALAS contributes to a future complex and pluralistic view of how the language we acquire as children helps us to shape our cultural world as much as our cognitive possibilities.

Bibliography

Abbeduto, Leonard and Sheldon Rosenberg (1985). Children's Knowledge of the Presuppositions of Know and Other Cognitive Verbs. *Journal of Child Language* 12(3), 621–641.

Agrillo, Christian, Laura Piffer, Angelo Bisazza, and Brian Butterworth (2012). Evidence for Two Numerical Systems That Are Similar in Humans and Guppies. *PLoS ONE* 7(2), e31923.

Aichhorn, Markus, Josef Perner, Benjamin Weiss, Martin Kronbichler, Wolfgang Staffen, and Gunther Ladurner (2009). Temporo-Parietal Junction Activity in Theory-of-Mind tasks: Falseness, Beliefs, or Attention. *Journal of Cognitive Neuroscience* 21 (6), 1179–1192.

Aikhenwald, Alexandra Y. (2004). *Evidentiality*. New York: Oxford University Press.

Aksu-Ko, Ayhan (1988). *The Acquisition of Aspect and Modality: the Case of Past Reference in Turkish*. Cambridge University Press.

Aksu-Ko, Ayhan (2000). Understanding Sources of Belief and Marking of Uncertainty. In *Proceedings of the 30th Annual Child Language Research Forum*, Stanford, CA, 123–130.

Aksu-Ko, Ayhan and Dan I. Slobin (1986). A Psychological Account of the Development and Use of Evidentials in Turkish. In Wallace L. Chafe and Joanna Nichils (Eds.), *Evidentiality: the Linguistic Coding of Epistemology*, 159–167. NJ: Norwood. Ablex Publishing Corporation.

Anderson, Lloyd B. (1986). Evidentials, Paths of Change and Mental Maps: Typologically Regular Asymmetries. In *Evidentiality: the Linguistic Coding of Epistemology,* 168–188 NJ: Norwood. Ablex Publishing Corporation.

Angeleri, Romina and Gabriella Airenti (2014). The Development of Joke and Irony Understanding: A study with 3-to-6-year-old children. *Canadian Journal of Experimental Philosophy,* 68(2), 133–146.

Apperly, Ian A. (2011). *Mindreaders: The Cognitive Basis of "Theory of Mind"*. New York, NY: Psychology Press.

Apperly, Ian A., Elisa Back, Dana Samson, and Lisa France (2008). The Cost of Thinking About False Beliefs: Evidence from Adults' Performance on a Non-Inferential Theory of Mind Task. *Cognition* 106(3), 1093–1108.

Apperly, Ian A. and Stephen A. Butterfill (2009). Do Humans Have Two Systems to Track Beliefs and Belief-Like States? *Psychological Review* 116(4), 953–970.

Apperly, Ian A. and E.J. Robinson (2001). Children's Difficulties Handling Dual Identity. *Journal of Experimental Child Psychology* 78(4), 374–397.

Apperly, Ian A. and E.J. Robinson (2003). When Can Children Handle Referential opacity? Evidence for Systematic Variation in 5- and 6-year-old Children's Reasoning About Beliefs and Belief Reports. *Journal of Experimental Child Psychology* 85(4), 297–311.

Apperly, Ian A., Dana Samson, Naomi Carroll, Glyn Humphreys, and Shazia Hussain (2006). Intact First- and Second-order False Belief Reasoning in Patient with Severely Impaired Grammar. *Social Neuroscience* 1(3–4), 334–348.

Aschersleben, Gisa, Tanja Hofer, and Bianca Jovanovic (2008). The Link Between Infant Attention to Goal-Directed Action and Later Theory of Mind Abilities. *Developmental Science* 11(6), 862–868.

ASHA (2012). Recommended Revisions to the DSM-5. Technical report, *American Speech-Language-Hearing Association* (ASHA), Washington, DC.

https://doi.org/10.1515/9783110748475-009

Astington, J. W. and J. M. Jenkins (1999). A Longitudinal Study of the Relation Between Language and Theory-of-Mind Development. *Developmental Psychology* 35(5), 1311–1320.

Bach, Kent (2002). Do Belief Reports Report Beliefs? *Pacific Philosophical Quarterly* 78(3), 215–241.

Baillargeon, Rene, David Buttelmann, and Victoria Southgate (2018). Invited Commentary: Interpreting Failed Replications of Early False-Belief Findings: Methodological and theoretical considerations. *Cognitive Developmen*t 46, 112–124.

Baldo, Juliana V., Nina F. Dronkers, David Wilkins, Carl Ludy, Patricia Raskin, and Jiye Kim (2005). Is Problem Solving Dependent on Language? *Brain and Language* 92(3), 240–250.

Baldwin, Dare A. and Jodie A. Baird (2001). Discerning Intentions in Dynamic Human Action. *Trends in Cognitive Science* 5(4), 171–178.

Baldwin, Dare A. and Megan M. Saylor (2012). Language Promotes Structural Alignment in the Acquisition of Mentalistic Concepts. In Janet Wilde Astngton and Jodie A. Baird (Eds.), *Why Language Matters for Theory of Mind*. Oxford, UK: Oxford University Press.

Baron-Cohen, Simon (1991). Do People with Autism Understand What Causes Emotion? *Child Development* 62(2), 385–395.

Baron-Cohen, Simon, Alan M. Leslie, and Uta Frith (1985). Does the Autistic Child Have a "Theory of Mind"? *Cognition* 21(1), 37–46.

Baron-Cohen, Simon, Howard Ring, John Moriarty, and Bettina Schmitz (1994). Recognition of Mental State Terms: Clinical Findings in Children with Autism and a Functional Neuroimaging Study of Normal Adults. *British Journal of Psychiatry* 165, 640–649.

Barsalou, Lawrence W. (2005). Situated Conceptualization. In Henri Cohen and Claire Lefebvre (Eds.), *Handbook of categorization in Cognitive Science*, 11–73. Amsterdam: Elsevier.

Bartsch, Karen and Henry M. Wellman (1995). *Children Talk About the Mind*. Oxford, New York: Oxford University Press.

Beaulac, Guillame (2014). *Language, Mind and Cognitive Science: remarks on theories of the Language-Cognition Relationships in Human Minds*. PhD Thesis, University of Western Ontario.

Berio, Leda (2020a). Culturally Embedded Schemata for False Belief Reasoning. *Synthese Special Issue: The cultural evolution of human social cognition,* DOI 10.1007/s11229–020–02655–7

Berio, Leda (2020b). Language Relativity and Flexibility of Mental Representations: Color Terms in a Frame Based Analysis. In *Cognitive Structures: Linguistic, Philosophical, and Psychological Perspectives* 2016, Düsseldorf, in press. Springer.

Berlin, Brent and Paul Kay (1969). *Basic Color Terms: Their Universality and Evolution*. Berkeley, Los Angeles, Oxford: University of California Press.

Bermúdez, José Luis (2005*). Philosophy of Psychology, A Contemporary Introduction*. Milton, UK: Routledge.

Bermúdez, José Luis (2009). Mindreading in the Animal Kingdom? In Robert W. Lurz (Ed.), *The Philosophy of Animal Minds*, 145–164. Cambridge, MA: Cambridge University Press.

Bermdez, José Luis (2003). *Thinking without Words*. New York: Oxford University Press.

Bialystok, Ellen, Gigi Luk, Kathleen F. Peets, and Sujin Yang (2010). Receptive Vocabulary Differences in Monolingual and Bilingual Children. *Bilingualism: Language and Cognition* 13(4), 525–531.

Bloom, Lois (1993). *Language Development from Two to Three*. Cambridge, NY: Cambridge University Press.

Bloom, Lois, Patsy Lightbrown, Lois Hood, Melissa Bowerman, Michael Marastos, and Michael P. Marastos (1975). Structure and Variation in Child Language. *Monographs of the Society for Research in Child Development* 40(2), 1–97.

Bloom, Paul and Tim P. German (2000). Two Reasons to Abandon the False Belief Task as a Test of Theory of Mind. *Cognition* 77(1), B25–B31.

Bonawitz, Elizabeth, Stephanie Denison, Thomas L. Griffiths, and Alison Gopnik (2014). Probabilistic Models, Learning Algorithms, Response Variability: Sampling in Cognitive Development. *Trends in Cognitive Science* 18(10), 497–500.

Booth, James R., William S. Hall, Gregory C. Robison, and Sue Yeong Kim (1997). Acquisition of the Mental State Verb Know by 2- to 5-Year-Old children. *Journal of Pyscholinguistic Research* 26(6), 581–603.

Borghi, Anna M. and Ferdinand Binkofski (2014). *Words as Social Tools: An Embodied View of on Abstract Concepts*. New York; Heidelberg; Dordrecht; London: Springer.

Boroditsky, Lera, Wendy Ham, and Michael Ramscar (2002). What is Universal in Event Perception? Comparing English and Indonesian. In *Proceedings of the annual meeting of the Cognitive Science Society*, Volume 24.

Boroditsky, Lera, Lauren A. Schmidt, and Philipps Webb (2003). Sex, Syntax, and Semantics. In Dedre Gentner and Susan Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought*. Cambridge, MA: MIT Press.

Bosco, Francesca and Ilaria Gabbatore (2017). Sincere, Deceitful, and Ironic Communicative Acts and the Role of the Theory of Mind in Childhood. *Frontiers in Psychology*, 8–21.

Bradley-Johnson, Sharon (1991). Book Review: Bankson Language rfTest -l2 R (BLT-2). *Journal of Psychoeducational Assessment* 9(4), 379–381.

Brennan, Susan E. and Herbert H. Clark (1996). Conceptual Pacts and Lexical Choice in Conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22(6), 1482–1493

Brosseau-Liard, Patricia E. and Susan A.J. Birch (2011). Epistemic States and Traits: Preschoolers Appreciate the Differential Informativeness of Situation Specific and Person Specific Cues to Knowledge. *Child Development* 82(6), 1788–1796.

Brne, Martin (2005). "Theory of Mind" in Schizophrenia: A Review of the Literature. *Schizophrenia bulletin* 31(1), 21–42.

Buttelmann, David, Malinda Carpenter, and Michael Tomasello (2009). Eighteen-month-old Infants Show False Belief Understanding in an Active Helping Paradigm. *Cognition* 112(2), 337–342.

Buttelmann, David, Harriet Over, Malinda Carpenter, and Michael Tomasello (2014). Eighteen-month-olds Understand False Beliefs in an Unexpected-Contents Task. *Journal of Experimental Child Psychology* 119, 120–126.

Byom, Lindsey J. and Bilge Mutlu Mutlu (2013). Theory of Mind: Mechanisms, Methods, and New Directions. *Frontiers in Human Neuroscience*.

Call, Joseph and Michael Tomasello (2008). Does the Chimpanzee Have a Theory of Mind: 30 Years Later. *In Trends in Cognitive Science*, 83–96.

Camp, Elisabeth (2009). Putting Thoughts to Work: Concepts, Systematicity, and Stimulus-Independence. *Philosophy and Phenomenological Research* 78(2), 275–311.

Cantor, Nancy, Walter Mischel, and Judith C. Schwartz (1982, January). A Prototype Analysis of Psychological Situations. *Cognitive Psychology* 14(1), 45–77.

Carey, Susan (2009). The Origins of Concepts. Oxford, NY: Oxford University Press.

Carey, Susan (2011a). Concept Innateness, Concept Continuity, and Bootstrapping. *Behavioral and Brain Sciences* 34(3), 152–162.

Carey, Susan (2011b). Precis on "The Origin of Concepts". *Behavioral and Brain Sciences* 34, 113–167.

Carpenter, Malinda (2009). Just How Joint is Joint Action in Infancy? *Topics in Cognitive Science* 1(2), 380–392.

Carruthers, Peter (1998). *Language, Thought and Consciousness*. Cambridge University Press, MA.

Carruthers, Peter (2002). The cognitive functions of Language. *Behavioral Brain Science* 25(6), 657–74.

Carruthers, Peter (2006). *The Architecture of the Mind: Massive Modularity and the Flexibility of Thought*. Oxford, NY: Oxford University Press.

Carruthers, Peter (2012). Language in Cognition. In Eric Margolis, Richard Samuels, and Stephen P. Stich (Eds.), *The Oxford Handbok of Philosophy of Cognitive Science*, 382–401. New York: Oxford University Press.

Carruthers, Peter (2013a). Animal Minds Are Real, (Distinctively) Human Minds Are Not. *American Philosophical Quarterly* 50(3), 233–248.

Carruthers, Peter (2013b). The Fragmentation of Reasoning. In Paulo Quintanilla (Ed.), *La coevolucin de mente y lenguaje: Ontognesis y filognesis*. Lima: Fondo Editorial de la Pontificia Universidad Catlica del Per.

Carruthers, Peter (2016). Two Systems for Mindreading? *Review of Philosophy and Psychology* 7, 141–162.

Carruthers, P. (2017). The illusion of Conscious Thought. *Journal of consciousness studies* 24(9–10), 228–252.

Casasanto, Daniel, Lera Boroditsky, Webb Phillips, Jesse Greene, Shima Goswami, Simon Bocanedra-Thiel, Ilia S. Diaz, Olga Fotokopoulu, Ria Pita, and David Gil (2004). How Deep are Effects of Language on Thought? Time Estimation in Speakers of English, Indonesian, Greek and Spanish. In *Proceedings of the 26th Annual Conference Cognitive Science Society*, Austin, Texas. Cognitive Science Society.

Cassidy, Kimberly Wright, Deborah Shaw Fineberg, Kimberly Brown, and Alexis Perkins (2005). Theory of Mind May Be Contagious, but You Don't Catch it From Your Twin. *Child Development* 76(1), 97–106.

Cattell, Ray (2007). *Children's Language. Consensus and Controversy* (2nd (Revised) ed.). London, UK; New York, NY: Continuum.

Chafe, Wallace L., Johanna Nichols, and Roy O. Freedle (Eds.) (1986). *Evidentiality: the Linguistic Coding of Epistemology*. NJ: Ablex Publishing Corporation.

Cheng, Ken (1986). A Purely Geometric Module in the Rat's Spatial Representation. *Cognition* 23(2), 149–78.

Choi, Soonja (1995). The Development of Epistemic Sentence-Ending Modal Forms and Functions in Korean Children. In Joan L. Bybee and Suzanne Fleischman (Eds.), *Modality in Grammar and Discourse*, 165–204. Amsterdam: Benjamins.

Choi, Soonja and Alison Gopnik (1995). Early Acquisition of Verbs in Korean: A Cross-Linguistci Study. *Journal of Child Language* 22(3), 497–529.

Chomsky, Noam (1980a). *Language and Learning: the Debate Between Jean Piaget and Noam Chomsky*. Cambridge, MA: Cambridge, MA: Harvard University Press.
Chomsky, Noam (1980b). Rules and Representations. *Behavioral and Brain Sciences* 3(1), 1–15.
Chomsky, Noam (1981). *Lectures on the Theory of Government and Binding*. Berlin: De Gruyter.
Chomsky, Noam (1995). The Minimalist Program. Cambridge, MA; London, UK: MIT Press.
Church, Alonzo (1950). On Carnap's Analysis of Statements of Assertion and Belief. *Analysis* 10(5), 97–99.
Churchland, Paul M. (1981). Eliminative Materialism and The Propositional Attitudes. *Journal of Philosophy* 78, 67–90.
Churchland, Paul M. (1988). *Matter and Consciousness*. Cambridge, MA: MIT Press.
Clark, A. (1996). Linguistic Anchors in the Sea of Thought? *Pragmatics and Cognition* 4(1), 93–103.
Clark, Andy (1998). Magic Words: How Language Augments Human Computation. In Peter Carruthers and Jill Boucher (Eds.), *Language and Thought: Interdisciplinary Themes*. Cambridge, MA: Cambridge University Press.
Clark, Andy (2006). Language, Embodiment, and the Cognitive Niche. *Trends in Cognitive Science* 10(8), 370–374.
Clark, Andy (2013). Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science. *Behavioral and Brain Sciences* 36(3).
Clements, Wendy A. and Josef Perner (1994). Implicit understanding of belief. *Cognitive Development* 9(4), 377–397.
Cohen, Adam, Jona Y. Sasaki, and Tamsin C. German (2015). Specialized Mechanisms for Theory of Mind: Are Mental Representations Special Because They Are Mental or Because They Are Representations? *Cognition* 136, 49–63.
Collingwood, Robin G. (1946). *The Idea of History*. Oxford, NY: Clarendon Press.
Coppola, Marie and Ann Senghas (2001). Children Creating Language: How Nicaraguan Sign Language Acquired a Spatial Grammar. *Psychological Science* 12(4), 323–328.
Courtin, Cyril (2000). The Impact of Sign Language on the Cognitive Development of Deaf Children: The Case of Theories of Mind. *The journal of deaf studies and deaf education* 5(3), 266–276.
Courtin, Cyril and Anne-Marie Melot (1998). Development of Theories of Mind in Deaf Children. *Psychological Perspectives on Deafness*, 79–102.
Covington, Michael A., He Congzhou, Cati Brown, Lorina Nai, Jonathan McClain, Bess Sirmon Fjordbak, James Semple, and John Brown (2005). Schizophrenia and the Structure of Language: The Linguist's View. *Schizophrenia Research* 1(77), 85–98.
Crimmins, Mark and John Perry (1989). The Prince and the Phone Boot: Reporting Puzzling Beliefs. *The Journal of Philosophy* 86(2), 685–711.
Csibra, Gergely and Gyrgy Gergely (2007). 'Obsessed with Goals': Functions and Mechanisms of Teleological Interpretation of Actions in Humans. *Acta Psychologica* 124 (1), 60–78.
Csibra, Gergely, Gyrgy Gergely, Szilvia Br, Orsolya Kos, and Margaret Brockband (1999). Goal Attribution Without agency Cues: the Perception of 'Pure Reason' in Infancy. *Cognition* 72(3), 237–267.
Custer, Wendy L. (1996). A Comparison of Young Children's Understanding of Contradictory Representations in Pretense, Memory, and Belief. In *Child Development* 67(2):678–88.

Dahlgren, Sven Olof and Anegen Trillingsgaard (1996). Theory of Mind in Non-Retarded Children with Autism and Asperger's Syndrome. A Research Note. *Journal of Child Psychology and Psychiatry* 37(6), 759 – 763.

Davidson, Donald (1968). On Saying That. *Synthese* 19(1/2), 130 – 146.

De Bruin, L.C. and A. Newen (2014). The Developmental Paradox of False Belief Understanding: a Dual System Solution. *Synthese* 191, 297 – 320.

De Bruin, Leon C. and Albert Newen (2012). An Association Account of False Belief Understanding. *Cognition* 123(2), 240 – 259.

de Villiers, Jill (2007). The Interface of Language and Theory of Mind. Lingua. *International review of general linguistics. Revue internationale de linguistique generale* 117(11), 1858 – 1878.

de Villiers, Jill (2014). What Kind of Concepts Need Language? *Language Science* 46 (Part B), 100 – 114.

de Villiers, Jill (2018). Perspectives on Truth: The Case of Language and False Belief Reasoning. In Kristen Syrett and Sudha Arunachalam (Eds.), *Semantics in Language Acquisition*, 222 – 245. Amsterdam: John Benjamins Publishing Company.

de Villiers, Jill and Peter de Villiers (2014). The Role of Language in Theory of Mind development. *Topics in Language Disorders* 34(4), 313 – 328.

de Villiers, Jill and Peter A. de Villiers (2009). Complements Enable Representation of the Contents of False Beliefs: the Evolution of a Theory of Theory of Mind. *In Language Acquisition*, 169 – 195. London: Palgrave Macmillan.

de Villiers, Jill and Peter A. de Villiers (2012). Linguistic Determinism and the Understanding of False Beliefs. In Peter Mitchell and Kevin J. Riggs (Eds.), *Children's Reasoning and the Mind*, 191 – 228. New York, NY: Hove U.K. Psychology Press.

de Villiers, Jill and S.A. Fitneva. *Referential Transparency for Opaque Containers*. Unpublished manuscript

de Villiers, Jill, Kathryn Hobbs, and Bart Hollebrandse (2014). Recursive Complements and Propositional Attitudes. In Tom Roeper and Margaret Speas (Eds.), *Recursion: Complexity in Cognition*, 221 – 242. Cham: Springer International Publishing.

de Villiers, Jill, E. Hosler, Carol A. Miller, M. Whalen, and J. Wong (1997). *Language, Theory of Mind, and Reading Other Peoples Emotions: A study of Oral Deaf Children*. Paper Presented at Society for Research in Child Development, Washington, DC, April.

de Villiers, Jill and Jennie Pyers (1997). Complementing Cognition: The Relationship between Language and Theory of Mind. In *Proceedings of the 21st Annual Boston University Conference on Language Development*.

de Villiers, Jill and Jennie E. Pyers (2002). Complements to Cognition: A Longitudinal Study of the Relationship Between Complex Syntax and False-Belief-Understanding. *Cognitive Development* 17(1), 1037 – 1060.

de Villiers, Jill G. (2001). Extention, Intension and Other Minds. In *Research in Child Language Acquisition: Proceedings of the 8th Conference of the International Association* for the Study of Child Language.

de Villiers, Jill G. (2004). Getting Complements on Your Mental State (Verbs). In S.Baauw J.Kampen, van (Ed.), *Proceedings of the 2003 GALA Conference*, Volume 1, Utrecht, 13 – 26.

de Villiers, Jill G. (2005). Can Language Acquisition give Children a Point of View? In *Why Language Matters for Theory of Mind*, 186 – 219. New York: Oxford University Press.

de Villiers, Jill G., Margaret Speas, Jay Garfield, Harper Gernet-Girard, and Tom Roeper (2009). Evidentials in Tibetan: Acquisition, Semantics and Cognitive Development. *In Evidentiality: a Window Into Language and Cognitive Development*.

de Villiers, Peter A., Frances Burns, and Barbara Pearson (2003). The Role of Language in the Theory of Mind Development of Language-Impaired Children: Complementing Theories. In Shanley Allen, Asli Ozyurek, Sotaro Kita, and Amanda Brown (Eds.), *Proceedings of the 27th Annual Boston University Conference on Language Development*, 232 – 242. Somerville, MA: Cascadilla Press.

de Villiers, Peter A. and Jill G. de Villiers (2011). Deception Dissociates from False Belief Reasoning in Deaf Children: Implications for the Implicit versus Explicit Theory of Mind Distinction. *British Journal of Developmental Psychology* 30(1), 188 – 209. Publisher: Wiley/Blackwell (10.1111).

Dehaene, Stanislas, Elizabeth Spelke, Philippe Pinel, Ruxandra Stanescu, and Susanna Tsivskin (1999). Sources of Mathematical Thinking: Behavioral and Brain-Imaging Evidence. *Science* 284(5416), 970 – 974.

Deleau, Michel (1996). L'attribution d'etats Mentaux chez Les Enfants Sourds et Entendants: Une Approche du Role de l'Experience Langagiere sur Une Theorie deLl'esprit. *Bulletin de Psychologie* 50(427), 48 – 56.

Dennett, Daniel C. (1991). *Consciousness Explained*. Boston: Little, Brown and Co.

Desjardins, N. and Dare A. Baldwin (1992). *Infants' understanding that object labels license inferences about underlying properties*. Unpublished manuscript, University of Oregon.

Devine, Rory T. and Claire Hughes (2014). Relations Between False Belief Understanding and Executive Function in Early Childhood: A MetaAnalysis. *Child Development* 85(5), 1777 – 1794.

Devitt, Michael (2006). *Ignorance of Language*. Oxford, NY: Oxford University Press.

Diaz, Vanessa and M. Jeffrey Farrar (2018). Do Bilingua and Monolingual Preschoolers Acquire False Belief Understanding Similarly? The Role of Executive Functioning and Language. *First Language* 38(4), 382 – 398.

Diessel, Holger and Michael Tomasello (2001). The Acquisition of Finite Complement Clauses in English: A Corpus-Based Analysis. *Cognitive Linguistics* 12(2), 97 – 141.

Ding, Nai, Lucia Melloni, Hang Zhang, Xing Tian, and David Poeppel (2016). Cortical Tracking of Hierarchical Linguistic Structures in Connected Speech. *Nature Neuroscience* 19, 158 – 164. Nature Publishing Group

Donald, Merlin (1991). *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Cambridge, MA: MIT Press.

Dorrenberg, Sebastian, Hannes Rakoczy, and Ulf Liszkowski (2018). How (Not) to Measure Infant Theory of Mind: Testing the Replicability and Validity of Four Non-Verbal Measures. *Cognitive Development* 46, 12 – 30.

Dove, Guy (2017). Language as a Disruptive Technology: Abstract Concepts, Embodiment and the Flexible Mind. *Philosophical Transactions of The Royal Society B* 373 (1752).

Dudley, Rachel, Naho Orita, Valentine Hacquard, and Jeffrey Lidz (2015). Three-Year-Olds' Understanding of Know and Think. In *Experimetal Perspectives on Presuppositions. Studies in Theoretical Psycholinguistics*, Volume 45, pp. 241 – 262. Cham: Springer.

Egan, Frances (1991). Propositional Attitudes and the Language of Thought. *Canadian Journal of Philosophy* 21(3), 370 – 388.

Eilan, Naomi, Christoph Hoerl, Teresa McCormack, and Johannes Roessler (Eds.) (2005). *Joint Attention: Communication and Other Minds*. Oxford, New York: Oxford University Press.
Erbach, Kurt, Peter Sutton, Hana Filip, and Kathrin Byrdeck (2017). Object Mass Nouns in Japanese. *Proceedings of the 21st Amsterdam Colloquium*.
Evans, Jonathan St. B. T. (2008). Dual-Processing Accounts of Reasoning, Judgement, and Social Cognition. *Annual Review of Psychology* 59, 225 – 278.
Evans, Jonathan St. B. T. and Keith E. Stanovich (2013). Dual-process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science* 8(3).
Everett, Caleb (2017). *Numbers and the Making of Us: Counting and the Course of Human Cultures*. Cambridge, MA; London, England: Harvard University Press.
Everett, Daniel L. (2005a). Comment to Everett's Cultural Constraints on Grammar and Cognition. *Current Anthropology* 46(4), 621 – 646.
Everett, Daniel L. (2005b). Cultural Constraints on Grammar and Cognition in Pi-raha. Another Look at the Design Features of Human Language. *Current Anthropology* 46(4).
Eysenck, Michael W. and Mark T. Keane (2012). *Psicologia Cognitiva* (Fourth Italian edition ed.). Neapel: Idelson-Gnocchi.
Fabricius, William V., Ty W. Boyer, Amy A. Weimer, and Kathleen C. Carroll (2010). True or False: Do 5-year-olds understand belief? *Developmental Psychology* 46(6), 1402 – 1416.
Fabricius, William V. and Suzanne L. Khalil (2003). False Beliefs or False Positives? Limits on Children's Understanding of Mental Representation. *Journal of Cognition and Development* 4(3), 239 – 262.
Fajans, Jane (1997). *They Make Themselves: Work and Play Among the Baining of Papua New Guinea*. Chicago, IL: University of Chicago Press.
Faller, Martina (2007). The Cuzco Quechua Conjectural: an Illocutionary Analysis of an Epistemic Evidential. *Proceedings of SULA 4*.
Farmer, M. (2000). Language and Social Cognition in Children with Specific Language Impairment. *Journal of Child Psychology and Psychiatry* 41(5), 627 – 636.
Farrant, Brad M., Janet Fletcher, and Murray T. Maybery (2006). Specific Language Impairment, Theory of Mind, and Visual Perspective Taking: Evidence for Simulation Theory and the Developmental Role of Language. *Child Development* 77(6), 1842 – 1853.
Farrant, Brad M., Murray T. Maybert, and Janet Fletcher (2012). Language, Cognitive Flexibility, and Explicit False Belief Understanding: Longitudinal Analysis in Typical Development and Specific Language Impairment. *Child Development* 83(1), 223 – 35.
Fedorenko, Evelina and Rosemary Varley (2017). Language and Thought Are Not The Same Thing: Evidence from Neuroimaging and Neurological Patients. *Annals of the New York Academy of Sciences* 1369(1), 132 – 153.
Fernandez, Camilia (2011). Mindful Storytellers: Emerging Pragmatics and Theory of Mind Development. *First Language* 33(1), 20 – 46.
Fiebich, Annika (2013). Mindreading with Ease? Fluency and Belief Reasoning in 4- to 5-year-olds. *Synthese* 191(5).
Figuras-Costa, Berta and Paul Harris (2001). Theory of Mind Development in Deaf Children: a Nonverbal Test of False-Belief Understanding. *Journal of Deaf Studies and Deaf Education* 6(2), 92 – 102.
Fisher, Cynthia (1996). Structural Limits on Verb Mapping: The Role of Analogy in Children's Interpretations of Sentences. *Cognitive Psychology* 31(1), 41 – 81.

Fisher, Cynthia, D. Geoffrey Hall, Susan Rakowitz, and Lila Gleitman (1994). When it is Better to Receive than to Give; Syntactic and Conceptual Constraints on Vocabulary Growth. *Lingua* 92, 331–392.

Fiske, Susan T. and Shelley E. Taylor (1984). *Social Cognition*. New York: Random House.

Fitneva, Stanka A. and Tomoko Matsui (2009). Special Issue: Evidentiality: a Window into Language and Cognitive Development. *New Directions for Child and Adolescent Development* 2009(125), 1–103.

Flavell, John H. (1993). The Development of Children's Understanding of False Belief and the Appearance-Reality Distinction. *International Journal of Psychology* 28(5), 595–604.

Fodor, J.A. (1985). Fodor's Guide to Mental Representation: The Intelligent Auntie's Vade-Mecum. *Mind* 94(373), 76–100.

Fodor, Jerry A. (1975). *The Language of Thought*. Cambridge, MA: Harvard University Press.

Fodor, Jerry A. (1978). Propositional Attitudes. In *RePresentations: Philosophical Essays on the Foundations of Cognitive Science*, MA. MIT Press.

Fodor, Jerry A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Fodor, Jerry A. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: Cambridge University Press.

Fodor, Jerry A. (1992). A Theory of the Child's Theory of Mind. *Cognition* 44 (3), 283–296.

Fodor, Jerry A. (2008). *LOT2: The Language of Thought Revisited*. Oxford, New York: Oxford University Press.

Fonteneau, Elisabeth and Jules Davidoff (2007). Neural Correlates of Color Categories. *Neuroreport* 27(18), 1323–13271323–1327.

Frege, Gottlob (1892). Über Sinn und Bedeutung. *Zeitschrift fr Philosophie und philosophische Kritik*. 100 (1):25–50

Frith, Christopher D. (2004). Schizophrenia and Theory of Mind. *Psychological Medicine* 34(3), 385–389.

Gamannossi, Beatrice Accorti and Giuliana Pinto (2014). Theory of Mind and Language of Mind in Narratives: Developmental Trends from Kindergarten to Primary School. *First Language* 34(3), 262–272.

Garfield, Jay L., Candida C. Peterson, and Tricia Perry (2001). Social Cognition, Language Acquisition and The Development of Theory of Mind. *Mind & Language* 16 (5), 494–541.

Gelman, Susan A. and John D. Coley (1991). Language and Categorization: The Acquisition of Natural Kind Terms. In *Perspectives on Language and Thought: Inter relations in Development*, pp. 146–196. Cambridge, NY: Cambridge University Press.

Gentner, Dedre (1978). On Relational Meaning: The Acquisition of Verb Meaning. *Child Development* 49(4), 988–998.

Gentner, Dedre, Florencia K. Anggoro, and Raquel S. Klibanoff (2011). Structure Mapping and Relational Language Support Children's Learning of Relational Categories. *Child Development* 82(4), 1173–1188.

Gentner, Dedre and Virginia Gunn (2001). Structural Alignment Facilitates the Noticing of Differences. *Memory and Cognition* 29(4), 565–577.

Gentner, Dedre and José Medina (1998). Similarity and the Development of Rules. *Cognition* 65(2–3), 263–297.

Gerber, Elizabeth R. (1985). Rage and Obligation: Samoan Emotion in Conflict. In Geoffrey M. White and John Kirkpatrick (Eds.), *Person, Self, and Experience: Exploring Pacific Ethnopsychologies*, 121–167. Berkeley: University of California Press.

Gibson, James J. (1986). *The Ecological Approach to Visual Perception*. Hillsdale, New Jersey: Hillsdale.

Gick, Mary L. and Keith J. Holyoak (1983). Schema Induction and Analogical Transfer. *Cognitive Psychology* 15(1), 1–38.

Gilbert, Aubrey L., Terry Regier, Paul Kay, and Richard B. Ivry (2007). Support for Lateralization of the Whorf Effect Beyond the Realm of Color Discrimination. *Brain and Language* 105(2), 91–98.

Gilbert, Daniel T. (1998). Ordinary Personology. In Susan T. Fiske, Gardner Lindzey, and Daniel T. Gilbert (Eds.), *The Handbook of Social Psychology*, 89–150. New York, NY: McGraw Hill.

Gillette, Jane, Henry Gletiman, Lila Gleitman, and Anne Lederer (1999). Human Simulations of Vocabulary Learning. *Cognition* 73(2), 135–176.

Gleitman, Lila (2009). The Structural Sources of Verb Meanings. *Language Acquisition* 1(1), 3–55. Publisher: Routledge.

Gleitman, Lila R., Kimberly Cassidy, Rebecca Nappa, Anna Papafragou, and John C. Trueswell (2009). Hard Words. *Language learning and development* 1(1), 23–64.

Gleitman, Lila R. and Barbara Landau (1985). *Language and Experience: Evidence from the Blind Child*. Cambridge, MA; London, England: Harvard University Press.

Goetz, Peggy J. (2003). The Effects of Bilingualism on Theory of Mind Development. *Bilingualism: Language and Cognition* 6(1), 1–15.

Goldman, Alvin I. (1986). *Epistemology and Cognition*. Cambridge, MA: Harvard University Press.

Goldman, Alvin I. (2006). *Simulating Minds: the Philosophy, Psychology and Neuroscience of Mindreading*. Oxford, NY: Oxford University Press.

Goldstone, Robert L. (1994). Influences of Categorization on Perceptual Discrimination. *Journal Experimental Psychology.* 123(2), 178–200.

Goldstone, Robert L., Richard M. Shiffrin, and Yvonne Lippa (2001). Altering Object Representations Through Category Learning. *Cognition* 78(1), 27–43.

Gopnik, Alison (2001). Theories, Language, and Culture: Whorf without Wincing. In Melissa Bowerman and Stephen C. Levinson (Eds.), *Language Acquisition and Conceptual Development*. Cambridge, NY: Cambridge University Press.

Gopnik, A. (2011). Commentary on Susan Carey "The Origins of Concepts", in Precis on "The Origins of Concepts". *Behavioral and Brain Sciences,* 44–47, Cambridge University Press

Gopnik, Alison and Janet W. Astington (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.

Gopnik, Alison and Elizabeth Bonawitz (2015). Bayesian Models of Child Development. *Wiley Interdisciplinary Reviews (WIRES) Cognitive Science* 6(2), 75–86.

Gopnik, Alison and Alvin I. Goldman (1993). How We know Our Minds: the Illusion of First-person Knowledge of Intentionality. In *Readings in Philosophy and Cognitive Science*, 315–346. Cambridge, MA; London, England: MIT Press.

Gopnik, Alison, Thomas L. Griffiths, and Christopher G. Lucas (2015). When Younger Learners Can Be Better (or at Least More Open Minded) than Older Ones. *Current Directions in Pyschological Science* 24(2), 87–92.

Gopnik, Alison and Andrew N. Meltzoff (1997). *Words, Thoughts, and Theories*. Cambridge, MA: Cambridge University Press.
Gopnik, Alison, Andrew N. Meltzoff, and Patricia Kuhl (1999). *The Scientist in The Crib*. New York, NY: Harper Collins.
Gopnik, Alison and Elizabeth Seiver (2009). Reading Minds: How infants Come to Understand Others. *Zero to Three* 30(2), 28 – 32.
Gordon, Robert M. (1986). Folk Psychology and Simulation. *Mind & Language* 1(2), 158 – 171.
Gordon, Robert M. (1992). The Simulation Theory: Objections and Misconceptsions. *Mind & Language* 7(1 – 2), 1 – 10.
Gordon, Robert M. (2005). Simulation and Systematic Erros in Prediction. *Trends in Cognitive Science* 9(8), 361 – 362.
Gordon, Robert M. (2007). Ascent Routines for Propositional Attitudes. *Synthese* 159 (2), 151 – 165.
Grice, Herbert Paul (1957). Meaning. *Philosophical Review* 66(3), 377 – 388.
Grice, Herbert Paul (1968). Utterer's Meaning, Sentence-Meaning, and Word-Meaning. *In Philosophy, Language, and Artificial Intelligence. Studies in Cognitive Systems*, Volume 2, 225 – 242. Dodrecht: Springer.
Grice, Herbert Paul (1982). *Meaning revisited. In Mutual knowledge.* New York: Academic Press.
Grice, Herbert Paul (1989). *Studies in the Ways of Words*. Cambridge, MA: Harvard University Press.
Guajardo, Nicole R. and Anne C. Watson (2017). Narrative Discourse and Theory of Mind Development. *The Journal of Genetic Psychology* 163(3), 305 – 325.
Hale, Courtney Melinda and Helen Tager-Flusberg (2003). The Influence of Language on Theory of Mind: A Training Study. *Developmental Science* 6(3), 346 – 359..
Hamilton, Antonia F., Rachel Brindley, and Uta Frith (2009). Visual Perspective Taking Impairment in Children with Autistic Spectrum Disorder. *Cognition* 113(1), 37 – 44.
Happ, Francesca G. E. (1995). The Role of Age and Verbal Ability in the Theory of Mind Task Performance of Subjects with Autism. *Child Development* 66(3), 843 – 855.
Harrigan, Kaitlyn, Valentine Hacquard, and Jeffrey Lidz (2016). Syntactic Bootstrapping in the Acquisition of Attitude Verbs: Think, Want and Hope. In Kyeong-min Kim et al (Ed.), *Proceedings of the 33rd West Coast Conference on Formal Linguistics*, Somerville, MA, 196 – 206. Cascadilla Proceedings Project.
Harrington, Leigh, Richard Siegert, and John McClure (2005). Theory of Mind in Schizophrenia: A Critical Review. *Cognitive Neuropsychiatry* 10(4), 249 – 286.
Harris, Paul L. (2005). Conversation, Pretense, and Theory of Mind. In *Why Language Matters for Theory of Mind*, 70 – 83. Oxford, UK: Oxford University Press.
Harris, Paul L., Carl N. Johnson, Deborah Hutton, Giles Andrews, and Tim Cooke (1989). Young Children's Theory of Mind and Emotion. *Cognition and Emotion* 3, 379 – 400.
Harris, Zellig S. (1957). Co-Occurrence and Transformation. *Linguistic Structure*. 33(3, 1), 283 – 340.
Heal, Jane (1986). Replication and Functionalism. In *Language, Mind and Logic*. Cambridge, MA: Cambridge University Press.
Hebb, Donald Olding (1949). *The Organization of Behavior.* New York, NY: Wiley and Sons.

Hedger, Joseph A. and William V. Fabricius (2011). True Belief Belies False Belief: Recent Findings of Competence in Infants and Limitations in 5-Year-Olds, and Implications for Theory of Mind Development. *Review of Philosophy and Psychology* 2 (429).

Heinrich, Joseph, Steven J. Heine, and Ara Norenzayan (2010). The Weirdest People in the World? *Behavioral and Brain Sciences* 33(2–3), 61–83.

Heyes, Cecilia (2018). *Cognitive Gadgets: the Cultural Evolution of Thinking*. Cambridge, Massachusetts: Harvard University Press.

Hinzen, Wolfram (2006). *Mind Design and Minimal Syntax*. Oxford, New York: Oxford University Press.

Hirai, Shin and Masako Jitsumori (2009). Counting Absolute Numbers of Items, from 1 to 8, in Pigeons. *Learning & Behavior* 37(4), 365–379.

Horton, William S and Richard J. Gerrig (2005). Conversational Common Ground and Memory Processes in Language Production. *Discourse Processes* 40(1).

Horton, William S. and Richard J. Gerrig (2016). Revisiting the Memory-Based Processing Approach to Common Ground. *Topics in Cognitive Science* 8(4), 780–795.

Horton, William S. and Boaz Keysar (1996). When Do Speakers Take into Account Common Ground? *Cognition* 59(1), 91–117. Publisher: Elsevier.

Hubbard, Edward M., Ilka Diester, Jessica F. Cantlon, Daniel Ansari, Filip van Opstal, and Vanessa Troiani (2008). The Evolution of Numerical Cognition: From Number Neurons to Linguistic Quantifiers. *Journal of Neuroscience* 28(46), 11219–11824.

Huemer, Michael, Josef Perner, and Brian Leahy (2018). Mental Files Theory of Mind: When do Children Consider Agents Acquainted with Different Object Identities? *Cognition* 171, 122–129.

Huffman, Mary Lyndia Crotteau, Elizabeth F. Loftus, and Stephen J. Ceci (1994). Repeatedly Thinking about a Non-Event: Source Misattributions among Preschoolers. *Consciousness and Cognition* 3(3–4), 388–407.

Hughes, Claire and Judy Dunn (1998). Understanding Mind and Emotion: Longitudinal Associations with Mental-State Talk between Young Friends. *Develomental Psychology* 34(5), 1026–1037.

Hume, David (1748). *A Treatise of Human Nature* (1975 ed.), Volume 2nd. Oxford, NY: Clarendon Press.

Huttenlocher, Janellen, Patricia Smiley, and Rosalind Charney (1983). The Emergence of Action Categories in the Child: Evidence from Verb Meaning. *Psychological Review* 90(1), 72–93.

Hutto, Daniel (2009). Folk Psychology as Narrative Practice. *Journal of Consciousness Studies* 16(6–8), 9–39(31).

Hutto, Daniel and Michaela Kirchhoff (2015). Looking Beyond the Brain: social Neuroscience Meets Narrative Practice. *Cognitive System Research* 34(35), 5–17.

Hutto, Daniel D. (2008a). First Communions: Mimetic Sharing without Theory of Mind. In Jordan Zlatev, Timothy P. Racine, Chris Sinha, and Esa Itkonen (Eds.), *The Shared Mind: Perspectives on Intersubjectivity*, 245–276. Amsterdam/ Philadelphia: John Benjamins Publishing Company.

Hutto, Daniel D. (2008b). *Folk Psychological Narratives: The Sociocultural Basis of Understanding Reasons*. Cambridge, MA

Hutto, Daniel D. (2008c). The Narrative Practice Hypothesis: Clarifications and Implications. *Philosophical Explorations* 11(3), 175–192.

Hutto, Daniel D. (2011). Understanding Fictional Minds without Theory of Mind! *Social Minds in Fiction and* Criticism, 45(2), 276 – 282

Hutto, Daniel D. (2017). Basic Social Cognition without Mindreading: Minding Minds without Attributing Contents. *Synthese* 194, 827 – 846.

Jackendoff, Ray S. (1990). *Semantic Structures*, Volume 18. Cambridge, MA: MIT Press.

James, William (1890*). Principles of Psychology*, Volume 1. New York: Holt.

Javor, Rebeka (2016). Bilingualism, Theory of Mind and Perspective-Taking: The Effect of Early Bilingual Exposure. *Psychological and Behavioral Sciences* 5(6), 143 – 148.

Jenkins, Jennifer M. and Jane Wilde Astington (1996). Cognitive Factors and Family Structure Associated with Theory of Mind Development in Young Children. *Develomental Psychology* 32(1), 70 – 78.

Johnson, Carl Nils and Michael P. Maratsos (1977). Early Comprehension of Mental Verbs: Think and Know. *Child Development* 48(4), 1743 – 1747.

Johnson, Carl Nils and H. M. Wellman (1980). Children's Developing Understanding of Mental Verbs: Remember, Know and Guess. *Child Development* 51(4), 1095 – 1102.

Kadosh, Roi Cohen and Ann Dowker (Eds.) (2015). *The Oxford Handbok of Numerical Cognition*. New York, NY: Oxford University Press.

Kahneman, Daniel and Shane Frederick (2005). A Model of Heuristic Judgment. In Robert G. Morrison Keith J. Holyoak (Ed.), *The Cambridge Handbook of Thinking and Reasoning*, 267 – 294. Cambridge, UK: Cambridge University Press.

Kamawar, Deepthi and David R. Olson (1999). Children's Representational Theory of Language: the Problem of Opacque Contexts. *Cognitive Development* 14(4), 531 – 548.

Karmiloff-Smith, Annette (1994). Beyond Modularity: a Developmental Perspective on Cognitive Science. *European Journal of Disorders of Communication* 29(1), 95 – 105.

Karmiloff-Smith, Annette, Edward Klima, Ursula Bellugi, Julia Grant, and Simon Baron-Cohen (1995). Is There a Social Module? Language, Face Processing, and Theory of Mind in Individuals with Williams Syndrome. *Journal of Cognitive Neuroscience* 7(2), 196 – 208.

Kerr, Sharyn and Kevin Durkin (2004). Understanding of Thought Bubbles as Mental Representations in Children with Autism: Implications for Theory of Mind. *Journal of Autism and Developmental Disorders* 34(6).

Kiang, Michael, Gregory Light, Jocelyn Prugh, Seana Coulson, David L. Braff, and Marta Kutas (2007). Cognitive, Neurophysiological, and Functional Correlates of Proverb Interpretation Abnormalities in Schizophrenia. *Journal Int. Neuropsychology Society* 13, 653 – 663.

Kidd, Evan (2003). *An Investigation of Children's Sentence Processing: a Developmental Perspective*. PhD Thesis, La Trobe University, Bundoore, Australia.

Kiparsky, Paul and Carol Kiparsky (1970). Fact. In Manfred Bierwisch and Karl E. Heidolph (Eds.), *Progress in Linguistics*. The Hague: Mouton.

Kircher, Tilo T. J., Tomasina M. Oh, Michael J. Brammer, and Philip K. McGuire (2005). Neural Correlates of Syntax Production in Schizophrenia. *British Journal of Psychiatry* 186, 209 – 214.

Kovacs, gnes Melinda, Erno Tgls, and Ansgar Denis Endress (2010). The Social Sense: Susceptibility to others' Beliefs in Human Infants and Adults. *Science* 330 (6012), 1830 – 1834.

Kripke, Saul A. (1979). A Puzzle about Belief. In *Meaning and Use*, 239 – 288. Israel: Reidel Dordrecht.

Kristen, Susanne, Beate Sodian, Claudia Thoermer, and Hannah Perst (2011). Infants' Joint Attention Skills Predict Toddlers' Emerging Mental State Language. *Develomental Psychology* 47(5), 1297 – 1219.

Kulke, Louisa and Hannes Rakoczy (2018). Implicit Theory of Mind – an Overview of Current Replications and Non-Replications. *Data in Brief* 16, 101 – 104.

Kyuchukov, Hristo and Jill de Villiers (2009). Theory of Mind and Evidentiality in Romani-Bulgarian Children. *Psychology of Language and Communication* 3(2), 21 – 34.

Lalumera, Elisabetta (2014). Whorfian Effects in Color Perception: Deep or Shallow? *Baltic International Yearbook of Cognition, Logic and Communication* 9(1), 1 – 13.

Lee, Jungmee (2010). The Korean evidential -te. A Modal Analysis. In *Proceedings of Colloque de Syntaxe et Semantique Paris 2009: Empirical issues in Syntax and Semantics*, Volume 8, 287 – 311.

Lee, Thomas Huntak and Ann Law (2000). Evidential Final Particles in Child Cantonese. In Eve V. Clark (Ed.), *The Proceedings of the Thirtieth Annual Child Language Research Forum,* 131 – 138. Stanford: Center for the Study of Language and Information.

Leekam, Susan and Josef Perner (1991). Does the Autistic Child Have a Metarepresentational Deficit? *Cognition* 40(3), 203 – 218.

Leekam, Susan, Claire Sewell, Josef Perner, and Laura Healey (2008). False Signs and the Non-Specificity of Theory of Mind: Evidence that Preschoolers have General Difficulties in Understanding Representations. *British Journal of Developmental Psychology* 26 (4), 485 – 497.

Leslie, Alan M. (1987). Pretense and Representation: the Origins of "Theory of Mind". *Psychological Review* 94(4), 412 – 426.

Leslie, Alan M. (1994a). Pretending and Believing: Issues in the Theory of ToMM. *Cognition* 50(1), 211 – 238.

Leslie, Alan M. (1994b). ToMM, ToBy, and Agency: Core Architecture and Domain Specificity. In Lawrence Hirschfeld and Susan A. Gelman (Eds.), *Mapping the Mind: Domain Specificity in Cognition and Culture*. Cambridge, NY: Cambridge University Press.

Leslie, Alan M. (2000). Theory of Mind as a Mechanism of Selective Attention. In Michael Gazzaniga (Ed.), *The New Cognitive Neurosciences* (2nd ed.). 1235 – 1247. Cambridge, MA: MIT Press.

Leslie, Alan M., Ori Friedman, and Tim P. German (2004). Core Mechanisms in Theory of Mind. *Trends in Cognitive Science* 8(12), 528 – 533.

Leslie, Adam M. and Laila Thaiss (1992). Domain Specificity in Conceptual Development: Neuropsychological Evidence from Autism. *Cognition* 43(3), 225 – 251.

Leung, Eleanor H. and Harriet L. Rheingold (1981). Development of Pointing as a Social Gesture. *Develomental Psychology* 17(2), 215 – 220.

Levinson, Stephen C. (1996). Frames of Reference and Molyneux's Question: Cross-Linguistic Evidence. In Paul Bloom, Mary A. Peterson, Lynn Nadel, and Merrill F. Gerrett (Eds.), *Language and Space*, 109 – 169. Cambridge, MA; London, England: MIT Press.

Levinson, Stephen C. (1997). From Outer to Inner Space: Linguistic Categories and Non-linguistic Thinking. In Jan Nuyts and Eric Pederson (Eds.), *Language and Conceptualization*, 13 – 45. Cambridge, MA: Cambridge University Press.

Levinson, Stephen C. (2003). *Space in Language and Cognition: Explorations in Cognitive Diversity*. Cambridge, UK: Cambridge University Press.

Levinson, Stephen C., Sotaro Kita, Daniel B. M. Haun, and Bjrn H. Rasch (2002). Returning the Tables: Language affects Spatial Reasoning. *Cognition* 84(2), 155 – 188.

Lewis, David K. (1986). *On the Plurality of Worlds*. Blackwell Publishers Ltd.

Li, Peggy and Lila Gleitman (2002). Turning the Tables: Language and Spatial Reasoning. *Cognition* 83(3), 265 – 294.

Limber, John (1973). The Genesis of Complex Sentences. In Timothy E. Moore (Ed.), *Cognitive Development and the Acquisition of Language*. New York, NY: Academic Press.

Lind, Sophie E. and Dermot M. Bowler (2009). Language and Theory of Mind in Autism Spectrum Disorder: the Relationship between Complement Syntax and False Belief Task Performance. *Journal of Autism and Developmental Disorders* 39(6), 929 – 937.

Liu, Qiang, Hong Li, Jennifer L. Campos, Qi Wang, Ye Zhang, Jiang Qiu, Qinglin Zhang, and Hong-jin Sun (2009). The N2pc Component in ERP and the Lateralization Effect of Language on Color Perception. *Neuroscience Letters* 454(1), 58 – 61.

Lohmann, Heidemarie and Michael Tomasello (2003). The Role of Language in the Development of False Belief Understanding: A Training Study. *Child Development* 74 (4), 1130 – 1144. Publisher: Wiley/Blackwell.

Low, Jason and Josef Perner (2012). Implicit and Explicit Theory of Mind: State of the art. *British Journal of Developmental Psychology* 30(1), 1 – 13.

Lupyan, Gary (2009). Extracommunicative Functions of Language: Verbal Interference Causes Selective Categorization Impairments. *Psychonomic Bulletin & Review* 16(4), 711 – 718.

Lupyan, Gary (2012). Linguistically Modulated Perception and Cognition: The Label-Feedback Hypothesis. *Frontiers in Psychology* 3(54).

Lupyan, Gary and Andy Clark (2015). Words and the World: Predictive Coding and the Language-Perception-Cognition Interface. *Current Directions in Pyschological Science* 24(4), 279 – 284.

Lupyan, Gary and Daniel Mirman (2013). Linking Language and Categorization: Evidence from Aphasia. *Cortex* 49(5), 1187 – 1194.

Lupyan, Gary, David H. Rakison, and James L. McClelland (2007). Language is not just for Talking: Redundant Labels Facilitate Learning of Novel Categories. *Psychological Science* 18(12), 1077 – 83.

Lutz, Catherine (1982). The Domain of Emotion Words on Ifaluk. *American Ethnologist* 9(1), 113 – 128.

Machery, Edouard (2008). Massive Modularity and the Flexibility of Human Cognition. *Mind& Language* 23(3), 363 – 272.

Macnamara, John, Erica Baker, and Chester L. Olson (1976). Four-Year-Olds' Understanding of "Pretend", "Forget" and "Know": Evidence for Propositional Operations. *Child Development* 47(1), 62 – 70.

MacWhinney, Brian (2000). *The CHILDES project: Tools for Analyzing Talk*. Hillsdale, NJ: Erlbaum.

Mahon, Merle and Alison Crutchley (2006). Performance of Typically-Developing School-Age Children with English as an Additional Language on the British Picture Vocabulary Scales II. *Child Language Teaching and Therapy* 22(3), 333 – 351.

Markman, Ellen M. (1992). Constraints on Word Learning: Speculations About their Nature, Origins, and Word Specificity. *In Modularity and Constraints in Language and Cognition*, 59 – 102. NJ: Lawrence Erlbaum Associates.

Masataka, Nobuo (2003). From Index-Finger Extension to Index-Finger Pointing: Ontogenesis of Pointing in Preverbal Infants. In Sotaro Kita (Ed.), *Pointing: Where Language, Culture and Cognition Meet*. Mahwah, New Jersey, London: Taylor and Francis Group.

Matthews, Robert J. (2007). *The Measure of Mind*. Oxford, NY: Oxford University Press.

Mayer, Andreas and Birgit E. Truble (2012). Synchrony in the Onset of Mental State Understanding Across Cultures? A Study Among Children in Samoa. *International Journal of Behavioral Development* 37(1), 21 – 28.

Mazza, Monica, Vittorio Di Michele, Rocco Pollice, Massimo Casacchia, and Rita Roncone (2008). Pragmatic Language and Theory of Mind Deficits in People with Schizophrenia and their Relatives. *Psychopathology* 41(4), 254 – 263.

McCready, Eric and Norry Ogata (2007). Evidentiality, Modality, and Probability. *Linguistics and Philosophy* 30, 147 – 206.

Meristo, Marek, Gary Morgan, Alessandra Geraci, Laura Iozzi, Erland Hjelmquist, Luca Surian, and Michael Siegal (2012). Belief Attribution in Deaf and Hearing Infants. *Developmental Science* 15(5), 633 – 640.

Meristo, Meristo, Karin Strind, and Erland Hjelmquist (2016). Early Conversational Environment Enables Spontaneous Belief Attribution in Deaf Children. *Cognition* 157, 139 – 145.

Mickan, Anne, Maren Schiefke, and Anatol Stefanowitsch (2014). Key is a llave is A Schlssel: A Failure to Replicate an Experiment from Botodistky et al. 2003. In *Yearbook of the German Cognitive Linguistics Association, Volume* 2, 39 – 50. Berlin: De Gruyter Mouton.

Miller, Carol A. (2010). False Belief and Sentence Complement Performance in Children with Specific Language Impairment. *International Journal of Language and Communication Disorders* 39(2), 191 – 213.

Milligan, Karen, Janet Wilde Astington, and Lisa Ain Dack (2007). Language and Theory of Mind: Meta-Analysis of the Relation between Language Ability and False-Belief Understanding. *Child Development* 78(2), 622 – 46.

Minsky, Marvin (1975). A Framework for Representing Knowledge. In Patrick Winston (Ed.), *The Psychology of Computer Vision*. New York, NY: McGraw-Hill.

Mitchell, Peter, Elizabeth J. Robinson, Jenny E. Isaacs, and Rebecca M. Nye (1996). Contamination in Reasoning about False Belief: An Instance of Realist Bias in Adults but not Children. *Cognition* 59(1), 1 – 21.

Mithen, Steven J. (1990). *Thoughtful Foragers: A study of Prehistoric Decision Making*. Cambridge, NY: Cambridge University Press.

Mo, Lei, Guiping Xu, Paul Kay, and Li-Hai Tan (2011). Electrophysiological Evidence for the Left-Lateralized Effect of Language on Preattentive Categorical Perception of Color. *PNAS* 108(34), 14026 – 14030.

Mo, Shuliang, Su Yanjie, Mark A. Sabbagh, and Jiaming Xiu (2014). Sentential Complements and False Belief Understanding in Chinese Mandarin-speaking Preschoolers: a Training Study. *Cognitive Development* 29, 50 – 61.

Mody, Maria and John W. Bellievau (2013). Speech and Language Impairments in Autism: Insights from Behavior and Neuroimaging. *North American Journal of Medicine and Science* 5(3), 157 – 161.

Moeller, Mary Pat and Brenda Schick (2006). Relations Between Maternal Input and Theory of Mind Understanding in Deaf Children. *Child Development* 77 (3), 751 – 766.

Montgomery, Derek E. (2005). The Developmental Origins of Meaning for Mental Terms. In Janet W. Astington and Jodie A. Baird (Eds.), *Why Language Matters for Theory of Mind*. Oxford, NY: Oxford University Press.

Moore, Chris and Jane Davidge (1989). The Development of Mental Terms: Pragmatics and Semantics. *Journal of Child Language* 16(3), 633 – 641.

Morissette, Paul, Marcelle Ricard, and Thrse G. Decarie (1995). Joint Visual Attention and Pointing in Infancy: a Longitudinal Study of Comprehension. *British Journal of Developmental Psychology* 13(2), 163 – 175.

Mossaheb, Nilufar, Harald N. Aschauer, Susanne Stoettner, Michaela Schmoeger, Nicole Pils, Monika Raab, and Ulrike Willinger (2014). Comprehension of Metaphors in Patients with Schizophrenia-spectrum Disorders. *Compr. Psychiatry* 55(4), 928 – 937.

Mossler, Daniel G., Robert S. Marvin, and Mark T. Greenberg (1976). Conceptual Perspective Taking in 2- to 6-year-old children. *Developmental Psychology*. 12(1), 85 – 86.

Mukanata, Yuko (2000). Challenges to the Violation-of-Expectation Paradigm: Throwing the Conceptual Baby out with the Perceptual Processing Water? *Infancy* 1(4), 471 – 477.

Naito, Mika and Kayo Koyama (2006). The Development of False-Belief Understanding in Japanese Children: Delay and difference? *International Journal of Behavioral Development* 30(4), 290 – 304.

Nappa, Rebecca, David January, Lila Gleitman, and John Trueswell (2004). Paying Attention to Attention: Perceptual Priming Effects on Word Order. In *Proceedings of the 17th annual CUNY Sentence Processing Conference*, Volume 26.

Nelson, Katherine (2005). Language Pathways into the Community of Minds. In Janet W. Astington and J.A. Baird, Jodie A. (Eds.), *Why Language Matters for Theory of Mind*, 26 – 49. Oxford, UK: Oxford University Press.

Newen, Albert (2011). Phenomenal Concepts and Mental files: Phenomenal concepts are theory-based. *Philosophia Naturalis* 47 – 48(1 – 2), 155 – 183.

Newen, Albert and Petra Vetter (2017). Why Cognitive Penetration of our Perceptual Experience is still the Most Plausible Account. *Consciousness and Cognition* 47, 26 – 37.

Nguyen, Thien-Kim and Janet Wilde Astington (2014). Reassessing the Bilingual Advantage in Theory of Mind and its Cognitive Underpinnings. *Bilingualism: Language and Cognition* 17(2), 396 – 409.

Nichols, Shaun and Stephen Stich (2000, February). A Cognitive Theory of Pretense. *Cognition* 74(2), 115 – 147.

Nichols, Shaun and Stephen P. Stich (2003). *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*. Oxford University Press.

Nieder, Andreas and Earl K. Miller (2004). Analog Numerical Representations in Rhesus Monkeys: Evidence for Parallel Processing. *Journal of Cognitive Neuroscience* 16(5), 889 – 901.

Nielsen, Mark and Cheryl Dissanayake (2010). An Invenstigation 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.

Norman, Donald Arthur and Tim Shallice (1986). *Attention to Action: Willed and Automatic Control of Behaviour*. New York: Plenum Press.

Oktay-Gür, Nese and Hannes Rakoczy (2017). Children's Difficulty with True Belief Tasks: Competence Deficit or Performance Problem? *Cognition* 166, 28 – 41.

Onishi, Kristine H. and Rene Baillargeon (2005). Do 15-month-old Infants Understand False Belief? *Science* 308(8), 255 – 258.

Ozturk, Ozge and Anna Papafragou (2016). The Acquisition of Evidentiality and Source Monitoring. *Language Learning and Development* 12(2), 199 – 230.

Pack, Adam A. and Louis M. Herman (2006). Dolphin Social Cognition and Joint Attention: Our Current Understanding. *Aquatic Mammals* 32(4), 443 – 460.

Papafragou, Anna, Kimberly Cassidy, and Lila Gleitman (2007). When We Think About Thinking: The Acquisition of Belief Verbs. *Cognition* 105(1), 125 – 165.

Papafragou, Anna, Justin Hulbert, and John Trueswell (2008). Does Language Guide Event perception? Evidence from Eye Movements. *Cognition* 108(1), 155 – 184.

Papafragou, Anna and Peggy Li (2001). Evidential Morphology and Theory of Mind. In Borbora Skarabela, Sarah Fish, and Anna H. J. Do Do (Eds.), *Proceedings from the 26th Annual Boston University Conference on Language Development,* Somerville MA. Cascadilla Press.

Papafragou, Anna, Peggy Li, Youngon Choi, and Han Chung-hye (2007). Evidentiality in Language and Cognition. *Cognition* 103(2), 253 – 299.

Paynter, Jessica and Candida C. Peterson (2013). Further Evidence of Benefits of Thought-bubble Training for Theory of Mind Development in Children with Autistm Spectrum Disorder. *Reseach in Autism Spectrum Disorders* 7(2), 344 – 348.

Pearson, Amy, Danielle Ropar, and Antonia F. de C. Hamilton (2013). A Review of Visual Perspective Taking in Autism Spectrum Disorder. *Frontiers in Human Neuroscience* 7(652).

Pereplyotchik, David (2017). *Psychosyntax. The Nature of Grammar and its Place in the Mind.* Number 129 in Philosophical Studies Series. Cham, Switzerland: Springer International Publishing.

Perner, Josef (1991). *Understanding the Representational Mind.* Cambridge, MA: MIT Press.

Perner, Josef, Michael Haemer, and Brian Leahy (2015). Mental Files and Belief: A Cognitive Theory of How Children Represent Belief and its Intentionality. *Cognition.* 145:77 – 88

Perner, Josef and Robert Horn (2003). Ignorance or False Negatives: Do Children of 4 to 5 Years Simulate Belief With Not Knowing = Getting it Wrong? *Journal of Cognition and Development* 4(3), 263 – 273.

Perner, Josef, Michael Huemer, and Brian Leahy (2015). Mental Files in Development: A Cognitive Analysis of How Children Represent Belief and its Intensionality. *Cognition* 145, 77 – 88.

Perner, Josef and Brian Leahy (2016). Mental files in Development: Dual Naming, False Belief, Identity and Intensionality. *Review of Philosophy and Psychology* 7, 491 – 508.

Perner, Josef and Susan Leekam (2008). The Curious Incident of the Photo that was Accused of Being False: Issues of Domain Specificty in Development, Autism, and Brain Imaging. *The Quarterly Journal of Experimental Psychology* 61(1), 76 – 89.

Perner, Josef, Susan R. Leekam, Shalini Davis, Nicola Odgers, and Deborah Myers (1998). Misrepresentation and Referential Confusion: Children's Difficulty with False Beliefs and Outdated Photographs. *Unpublished manuscript*

Perner, Josef, Susan R. Leekam, and Heinz Wimmer (1987, June). 3-year-olds' Difficulty with False Belief: the Case for a Conceptual Deficit. *British Journal of Developmental Psychology* 5(2), 125 – 137.

Perner, Josef, M. Christine Mauer, and Monika Hildenbrand (2011). Identity: Key to Children's Understanding of Belief. *Science* 333(6041), 474 – 477.
Perner, Josef and Ted Ruffman (2005). Infant's Insight into the Mind: How Deep? *Science* 308(5719), 214 – 216.
Perner, Josef, Ted Ruffman, and Susan R. Leekam (1994). Theory of Mind is Contagious: You Catch it From Your Sibs. *Child Development* 65(4), 1228 – 1238.
Perner, Josef, Manuel Sprung, Petra Zauner, and Hubert Haider (2003). Want That Is Understood Well before Say That, Think That, and False Belief: A Test of de Villiers's Linguistic Determinism on German-Speaking Children. *Child Development* 74(1), 179 – 188
Perner, Josef, Sandra Stummer, Manuel Sprung, and Martin Doherty (2002). Theory of Mind Finds it Piagetian Perspective: Why Alternative Naming Comes with Understanding Belief. *Cognitive Development* 17(3 – 4), 1451 – 1472.
Perner, Josef, Petra Zauner, and Manuel Sprung (2005). What Does "That" Have to Do with Point of View? In Janet W. Astington and Jodie A. Baird (Eds.), *Why Language Matters for Theory of Mind*. Oxford, NY: Oxford University Press.
Perry, John (2001). *Knowledge, Possibility, and Consciousness*. Cambridge, MA: MIT Press.
Peterson, Candida C. (2000). Kindred spirits: Influences of Siblings' Perspectives on Theory of Mind. *Cognitive Development* 15(4), 435 – 455.
Peterson, Candide C. and Michael Siegal (1998). Changing Focus on the Representational Mind: Deaf, Autistic and Normal Children's Concepts of False Photos, False Drawings and False Beliefs. *British Journal of Developmental Psychology* 16(3), 301 – 320.
Peterson, Candida C. and Michael Siegal (1999). Representing Inner Worlds: Theory of Mind in Deaf, Autistic and Normal Hearing Children. *Psychological Science* 10(2), 126 – 129.
Peterson, Candida C. and Michael Siegal (2002). Insights into Theory of Mind from Deafness and Autism. *Mind and Language* 15(1), 123 – 145.
Piaget, Jean (1936). *The Origins of Intelligence in the Child*. Milton, UK: Routledge & Kegan Paul.
Piaget, Jean (1952). *The Origins of Intelligence in Children*. New York: International University Press.
Pica, Pierre, Cathy Lemer, Vronique Izard, and Stanislas Dehaene (2004). Exact and Approximate Arithmetic in an Amazonian Indigene Group. *Science* 306 (5695), 499 – 503.
Pickering, Martin J. and Simon Garrod (2013). An Integrated Theory of Language Production and Comprehension. An Integrated Theory of Language Production and Comprehension. *Behavioral and Brain Science* 36(4), 329 – 347.
Pillow, Bradford H. (1989). Early Understanding of Perception as a Source of Knowledge. *Journal of Experimental Child Psychology* 47(1), 116 – 129.
Pinker, Steven (1989). *Learnability and Cognition*. MIT Press.
Polich, Laura (2005). *The Emergence of the Deaf Community in Nicaragua: With Sign Language You Can Learn so Much*. Washington, DC: Gallaudet University Press.
Poulin-Dubois, Diane, Hannes Rakoczy, Kimberly Burnside, Cristina Crivello, Sebastian Drrenberg, Katheryn Edwards, Horst Krist, Louisa Kulke, Ulf Liszkowski, Jason Low, Josef Perner, Lindsey Powell, Beate Priewasser, Eva Rafetseder, and Ted Ruffman (2018). Do Infants Understand False Beliefs? We Don't Know yet A Commentary on Baillargeon, Buttelmann and Southgate's commentary. *Cognitive Development* 48, 302 – 315.
Poulin-Dubois, Diane and Jessica Yott (2017). Probing the Depth of Infants' Theory of Mind: Disunity in Performance Across Paradigms. *Developmental Science* 21(4), e12600.

Povinelli, Daniel J., Jesse M. Bering, and Steve Giambrone (2000). Toward a Science of Other Minds: Escaping the Argument by Analogy. *Cognitive Science* 24 (3), 509 – 541.

Pratt, Chris and Peter Bryant (1990). Young Children Understand That Looking Leads to Knowing (So Long as They Are Looking into a Single Barrel). *Child Development* 61(4), 973 – 982.

Priewasser, Beate, Eva Rafetseder, Carina Gargitter, and Josef Perner (2018). Helping as an Early Indicator of a Theory of Mind: Mentalism or Teleology? *Cognitive Development* 46, 69 – 78.

Proft, Marina, Britta Schnemann, and Hannes Rakoczy (2019). Children's Understanding of the Aspectuality of Intentions. *Journal of Experimental Child Psychology* 181, 17 – 33.

Pyers, Jennie E and Ann Senghas (2009). Language Promotes False-Belief Understanding: Evidence From Learners of a New Sign Language. *Psychological Science* 20(7), 805 – 812.

Quine, Willard V.O. (1960). Word and Object. Cambridge, MA: MIT Press.

Rakison, David H. and Yevdokiya Yermolayeva (2010). Infant Categorization. *WIREs Cognitive Science* 1(6), 894 – 905.

Rakoczy, Hannes (2003). *The Role of Experience and Discourse in the Developing Understanding of Pretense: a Training Study.* In Paper presented at SRDC.

Rakoczy, Hannes, H. (2017). Comparative Metaphysics: Thinking about Objects in Space and Time. In J. Call, Josep (Ed.), Handbook of Comparative Psychology, Volume 2: Perception, learning and cognition. Washington, D.C.: American Psychological Association.

Range, Friederike, Ulrike Aust, Michael Steurer, and Ludwig Huber (2008). Visual Categorization of Natural Stimuli by Domestic Dogs. *Animal Cognition* 11(2), 339 – 47.

Rapp, Alexander M., Karin Langhor, Dorothee E. Mutschler, and Barbara Wild (2014). Irony and Proverb Comprehension in Schizophrenia: Do Female Patients "Dislike" Ironic Remarks? *Schizophrenia Res. Treatment* 2014(

Recanati, Franois (2012). *Mental Files.* Oxford, NY: Oxford University Press.

Reilly, Sheena, Bruce Tomblin, James Law, Cristina McKean, Fiona Mensah, Angela Morgan, Sharon Goldfeld, Jan M. Nicholson, and Melissa Wake (2014). Specific Language Impairment: a Convenient Label for Whom? *International Journal of Language and Communication Disorders* 49(4), 416 – 451.

Richard, Mark (1989). How I Say What You Think. *Midwest Studies in Philosophy*, 317 – 337.

Rizzolatti, Giacomo, Leonardo Fogassi, and Vittoro Gallese (2001). Neurophysiological Mechanisms Underlying the Understanding and Imitation of Action. *Nature Neuroscience Reviews* 2, 661 – 670.

Rizzolatti, Giacomo and Maurizio Gentilucci (1988). Motor and Visual-Motor Functions of the Premotor Cortex. In John Wiley (Ed.), *Neurobiology of Neocortex.* New York, NY: Wiley and Sons.

Rizzolatti, Giacomo, M. Gentilucci, Maurizio, Luca Fogassi, Giuseppe Luppino, Massimo Matelli, and Luca Ponzoni-Maggi (1987). Neurons Related to Goal-Directed Motor Acts in Inferior Area 6 of the Macacque Monkey. *Experimental Brain Research* 67 (1), 220 – 224.

Rizzolatti, Giacomo and Corrado Sinigaglia (2006). *So Quel Che Fai. Il Cervello che Agisce e i Neuroni Specchio.* Milano: Raffaello Cortina Editore.

Robbins, Joel and Alan Rumsey (2008). Introduction: Cultural and Linguistic Anthropology and the Opacity of Other Minds. *Anthropoligical Quarterly* 81(2), 407 – 420.

Roberson, Debi and Jules Davidoff (2000). The Categorical Perception of Colors and Facial Expressions: the Effect of Verbal Interference. *Memory and Cognition* 28(6), 977 – 986.

Roberson, Debi, Jules Davidoff, and Nick Braisby (1999). Similarity and Categorisation: Neuropsychological Evidence for a Dissociation in Explicit Categorisation Tasks. *Cognition* 71(1), 1–42.

Roberson, Debi, Jules Davidoff, Ian R.L. Davis, and Laura R. Shapiro (2004). Color Categories: Evidence for the Cultural Relativity Hypothesis. *Cognitive Psychology* 50(4), 378–411.

Roberson, Debi, Hyensou Pak, and Richard Hanley (2008). Categorical Perception of color in the left and right visual field is verbally mediated: evidence from Korean. *Cognition* 107(2), 752–762.

Robinson, Elizabeth J. and Philipp Whitcombe (2003). Children's Suggestibility in Relation to Their Understanding About Souces of Knowledge. *Child Development* 74(1), 48–62.

Rodriguez-Ferrera, Silvia, Ryan A. McCarthy, and Peter J. McKenna (2001). Language in Schizophrenia and its Relationship to Formal Thought Disorder. *Psychological Medicine* 31(2), 197–205.

Ruffman, Ted and Josef Perner (2005). Do Infants Really Understand False Belief? *Trends in Cognitive Science* 9, 462–463.

Ruffman, Ted, Lance Slade, and Elena Crowe (2003). The Relation Between Children's and Mothers' Mental State Language and Theory-of-Mind Understanding. *Child Development* 73(3), 734–751.

Ruffman, Ted, Lance Slade, Kate Rowlandson, Charlotte Rumsey, and Alan Garnham (2003). How Language Relates to Belief, Desire, and Emotion Understanding. *Cognitive Development* 18(2), 139–158.

Russel, James (1987). "Can We Say..?" Children's Understanding of Intensionality. *Cognition* 25(3), 289–308.

Russell, Phil, Judith Hosie, Colin Gray, Christie Scott, Norma Hunter, Joanne Banks, and Mairi Macaulay (1998). The Development of Theory of Mind in Deaf Children. *Journal of Psychology and Psychiatry* 39(6), 903–910.

Satoh, Ken, Akihiro Inokuchi, Katashi Nagao, and Takahiro Kawamura. (2008). Semantic Heterogeneity in Evidentials. *In New Frontiers of Artificial Intelligence, LNCS,* Volume 4914. Heidelberg, Berlin: Springer.

Saxe, Rebecca, Susan Carey, and Nancy Kanwisher (2004). Understanding Other Minds: Linking Developmental Psychology and Functional Neuroimaging. *Annual Review of Psychology* 55(1), 87–124.

Schank, Roger C. (1982*). Dynamic Memory: A Theory of Reminding and Learning in Computers and People.* Cambridge, MA: Cambridge University Press.

Schank, Robert C. and Robert Abelson (1977). *Scripts, Plans, Goals, and Understanding.* Hillsdale, NJ: Erlbaum Associates.

Schettino, Antonio, Leonor Romero Lauro, Franca Crippa, Simona Anselmetti, Roberto Cavallaro, and Constanza Papagno (2010). The Comprehension of Idiomatic Expressions in Schizophrenic Patients. *Neuropsychologia* 48(4), 1032–1040.

Schick, Brenda, Peter de Villiers, Jill de Villiers, and Robert Hoffmeister (2007). Language and Theory of mind: a Study of Deaf Children. *Child Development* 78(2), 376–396.

Schiffer, Stephen (1992). Belief Ascription. *The Journal of Philosophy* 89(10), 499–521.

Schroeder, Scott R. (2018). Do Bilinguals Have an Advantage in Theory of Mind? A Meta-Analysis. *Frontiers in Communication* 3(36).

Scott, Rose M. and Rene Baillargeon (2009). Which Penguin is This? Attributing False Beliefs About Object Identity at 18 Months. *Child Development* 80(4), 1172–1196.

Scott, Rose M., Zijing He, Rene Baillargeon, and Denise Cummins (2011). False-belief Understanding in 2.5 year olds: Evidence from Two Novel Verbal Spontaneous-Response Task. *Developmental Science 15(2), 181–193*.

Searle, John R. and Daniel Vanderveken (1985). *Foundations of Illocutionary Logic*. London, New York, New Rochelle, Melbourne, Sydney: Cambridge University Press.

Sellars, Wilfrid (1997). *Empiricism and the Philosophy of Mind*. Cambridge, MA: Harvard University Press.

Senghas, Ann, Sotaro Kita, and Asli zyrek (2004). Children Creating Core Properties of Language: Evidence from an Emerging Sign Language in Nicaragua. *Science* 305 (5691), 1779–1782.

Shatz, Marilyn, Sharon Silber, and Henry M. Wellman (1983). The Acquisition of Mental verbs: A Systematic Investigation of the First Reference to Mental State. *Cognition* 14(3), 301–321.

Siegal, Michael, Rosemary Varley, and Stephen Want (2001). Mind Over Grammar: Reasoning in Aphasia and Development. *Trends in Cognitive Science* 5(7), 296–301.

Slade, Lance and Ted Ruffman (2010). How Language Does (and Does Not) Relate to Theory of Mind: A Longitudinal Study of Syntax, Semantics, Working Memory and False Belief. *Develomental Psychology* 23(1), 117–141.

Smiley, Patricia and Janellen Huttenlocher (1995). Conceptual Development and the Child's Early Words for Events, Objects, and Persons. In Michael Tomasello and William Merriman (Eds.*), Beyond Names for Things*, 21–62. NJ: Lawrence Erlbaum Associates.

Snedeker, Jesse and Lila R. Gleitman (2004). Why is it Hard to Label our Concepts? In *Weaving a Lexicon*. Cambridge, MA: MIT Press.

Soames, Scott (2002). The Problem Posed by Propositional Attitude Ascriptions. In *Beyond Rigidity*. Oxford, NY: Oxford University Press.

Song, Hyun-joo and Rene Baillargeon (2008). Infant's Reasoning About others' False Perceptions. *Developmental Psychology* 44(6), 1789–1795.

Soproni, Kristina, dm Miklsi, Jzsef T. Topl, and Vilmos Csnyi (2002). Dogs' (canis familiaris) Responsiveness to Human Pointing Gestures. *Journal of Comparative Psychology* 116(1), 27–34.

Southgate, Victoria, Coralie Chevallier, and Gergely Csibra (2010). Seventeen-month-olds Appeal to False Beliefs to Interpret Others' Referential Communication. *Developmental Science* 13(6), 907–912.

Southgate, Victoria, Atsushi Senju, and Gergely Csibra (2007). Action Anticipation Through Attribution of False Belief by Two-year-olds. *Psychological Science* 18 (7), 587–592.

Spanoudis, George (2016). Theory of Mind and Specific Language Impairment in School Age Children. *Journal of Communication Disorders* 61, 83–96.

Sparrewohn, Roslyn and Pauline M. Howie (1995). Theory of Mind in Children with Autistic Disorder: Evidence of Developmental Progression and the Role of Verbal Ability. *Journal of Child Psychology and Psychiatry* 36(2), 249–263.

Sprong, Mirjam, Patricia Schothorts, Ellen Vos, Joop Hox, and Herman Van Engeland (2007). Theory of Mind in Schizophrenia: Meta-analysis. *The British Journal of Psychiatry* 191(1), 5–13.

Sprung, Manuel, Josef Perner, and Peter Mitchell (2007). Opacity and Discourse Referents: Object Identity and Object Properties. *Mind & Language* 22(3), 215–245.

Stalnaker, Robert (1984). *Inquiry*. Cambrdige, NY: MIT Press.

Steeds, Lucy, Karen Rowe, and Ann Dowker (1997). Deaf Children's Understanding of Beliefs and Desires. *Journal of Deaf Studies and Deaf Education* 2(3), 185–195.

Sterelny, Kim (2006). Language, Modularity and Evolution. In Graham MacDonald and David Papineau (Eds.), *Teleosemantics*, 23–41. Oxford, UK: Oxford University Press.

Steup, Matthias and Ram Neta (2018). Epistemology. In Stanford Encyclopedia of Philosophy, E. N. Zalta (ed). <https://plato.stanford.edu/archives/fall2020/entries/epistemology/>

Suddendorf, Thomas and Andrews Whiten (2003). Reinterpreting the Mentality of Apes. In Kim Sterelny and Julie Fitness (Eds.), *From Mating to Mentality: Evaluating Evolutionary Psychology*. Hove U.K.: Psychology Press.

Surian, Luca, Stefania Caldi, and Dan Sperber (2007). Attribution to Beliefs by 13- month-old infants. *Psychological Science* 18(7), 580–586.

Surian, Luca and Alessandra Geraci (2012). Where Will the Triangle Look for It? Attributing False beliefs to a Geometric Shape at 17 months. *British Journal of Developmental Psychology* 30(1), 30–44.

Sutton, John (2002). Cognitive Conceptions of Language and the Development of Autobiographical Memory. *Language and Communication* 22(3), 375–390.

Symons, Douglas K. (2004). Mental State Discourse, Theory of Mind, and the Internalization of Selfother Understanding. *Developmental Review* 24(2), 159–188.

Symons, Douglas K., Candida C. Peterson, Virginia Slaughter, Jackie Roche, and Emily Doyle (2010). Theory of Mind and Mental State Discourse During Book Reading and Storytelling Tasks. *British Journal of Developmental Psychology* 23 (1), 81–102.

Tager-Flusberg, Helen (1995). 'Once upon a ribbit': Stories Narrated by Autistic Children. *British Journal of Developmental Psychology* 13(1), 45–59.

Tager-Flusberg, Helen (2000). Language and Understanding Minds: Connections in Autism. In Simon Baron-Cohen, Helen Tager-Flusberg, and Donald J. Cohen (Eds.), *Understanding Other Minds: Perspectives from Autism and Development Cognitive Neuroscience* (2 ed.)., 1–45. Oxford, UK: Oxford University Press.

Tager-Flusberg, Helen and Kate Sullivan (1994a). A Componential View of Theory of Mind: Evidence from Williams Syndrome. *Cognition* 76, 59–89.

Tager-Flusberg, Helen and Kate Sullivan (1994b). Predicting and Explaining Behavior: A Comparison of Autistic, Mentally Retarded and Normal Children. *Journal of Child Psychology and Psychiatry* 35(6), 1059–1075.

Talwar, Victoria and Kang Lee (2008). Social and Cognitive Correlates of Children's Lying Behavior. *Child Development* 79(4), 866–881.

Tan, Jancita and Paul L. Harris (1991). Autistic Children Understand Seeing and Wanting. *Develoment and Psychopathology* 3(2), 163–174.

Tan, Li Hai, Alice H. D. Chan, Paul Kay, Pek-Lan Khong, Lawrance K. C. Yip, and Kang-Kwong Luke (2008). Language Affects Patterns of Brain Activation Associated with Perceptual Decision. *PNAS* 105(10), 4004–4009.

Tardif, Twila (1996). Nouns Are Not Always Learned Before Verbs: Evidence from Mandarin Speakers' Early Vocabularies. *Developmental Psychology* 32(3), 492–504.

Tardif, Twila, Susan A. Gelman, and Fan Xu (1999). Putting the "Noun Bias" in Context: a Comparison of English and Mandarin. *Child Development* 70(3), 620–635.

Tardif, Twila, Catherine Wing-Chee So, and Niko Kaciroti (2007). Language and False Belief: Evidence for General, not Specific, Effects in Cantonese-speaking éreschoolers. *Developmental Psychology* 43(2), 318–340.

Tardif, Twila and Henry M. Wellman (2000). Acquisition of Mental State Language in Mandarin and Cantonese-speaking Children. *Develomental Psychology* 36(1), 25 – 43.

Tardiff, Twila, Marilyn Shatz, and Letitia Naigles (1997). Caregiver Speech and Children's Use of Nouns vs. Verbs. A Comparison of English, Italian and Mandarin. *Journal of Child Language* 24, 535 – 565.

Tavano, Alessandro, Silvia Sponda, Franco Fabbro, Cinzia Perlini, Gianluca Rambaldelli, Adele Ferro, Stefania Cerruti, Michele Tansella, and Paolo Brambilla (2008). Specific Linguistic and Pragmatic Deficits in Italian Patients with Schizophrenia. *Schizophrenia Research* 102(1 – 3), 53 – 62.

Thierry, Guillaume, Panos Athanasopoulos, Alison Wiggett, Benjamin Dering, and Jan-Rouke Kuipers (2009). Unconscious Effects of Language-Specific Terminology on Preattentive Color Perception. *Proc Natl Acad Sci* USA 106(11), 4567 – 4570.

Tillas, Alex (2014). How Do Ideas Become General in their Signification? *The Baltic International Yearbook of Cognition, Logic and Communication* 9, 1 – 35.

Tillas, Alexandros (2015a). Language as Grist to the Mill of Cognition. *Cognitive Processing* 16(3), 219 – 243.

Tillas, Alex (2015b). On the Origins of Concepts. In Christoph Kann, David Hommen, and Tanja Osswald (Eds.), Concepts and Categorization, pp. 169 – 196.

Tomasello, Michael (1995). Joint Attention as Social Cognition. In Chris Moore and Philip J. Dunham (Eds.), *Joint Attention: Its Orgins and Role in Development*, 103 – 130. New York, NY: Lawrence Erlbaum Associates.

Tomasello, Micheal (1996). Piagetian and Vygotskian Approaches to Language. Human Development 39, 269 – 276.

Tomasello, Michael (2000). The Social-Pragmatic Theory of Word Learning. *Pragmatics* 10(4), 401 – 413.

Tomasello, Michael (2009a). *Constructing a Language*. Cambridge, MA: Harvard University Press.

Tomasello, Michael (2009b). The Usage-Based Theory of Language Acquisition. In Victoria La trabe university (Ed.), *The Cambridge Handbook of Child Language*. Cambrdige, MA: Cambridge University Press.

Tomasello, Michael (2014). *A Natural History of Human Thinking*. London, England: Harvard University Press.

Tomasello, Michael and Hannes Rakoczy (2003). What Makes Human Cognition Unique? From Individual to Shared to Collective Intentionality. *Mind&Language* 18(2), 121 – 147.

Tucker, L. (2004). *Specific Language Impairment and Theory of Mind: Is Normal Language Development an Essential Precursor for on Time Theory-of-Mind Development?* Unpublished manuscript, University of Western Australia

Uncapher, Melina R., Benjamin J. Hutchinson, and Anthony D. Wagner (2011). Dissociable Effects of Top-Down and Bottom-Up Attention During Episodic Encoding. *Journal of Neuroscience* 31(35), 12613 – 12628.

Van Cleave, Matthew and Christopher Gauker (2013). Linguistic Practice and False-Belief Tasks. *Mind & Language* 25(3), 298 – 328.

Varley, Rosemary and Michael Siegal (2000). Evidence for Cognition Without Grammar from Casual Reasoning and 'Theory of Mind' in an Agrammatic Aphasic Patient. *Current Biology* 10(12), 723 – 726.

Vauclair, Jacques (2002). Categorization and Conceptual Behavior in Nonhuman Primates. In Mark Bekoff, Colin Allen, and Gordon M. Burghardt (Eds.), *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*, 239 – 246. Cambridge, MA: MIT Press.

Vicker, Beverly (2009). *Social Communication and Language Characteristics Associated with High Functioning, Verbal Children and Adults with Autism Apectrum Disorder.* Bloomington, IN: Indiana Resource Center for Autism.

Vinden, Penelope G. (1996). Junn Quechua Children's Understanding of Mind. *Child Development* 67(4), 1707 – 1716.

Vinden, Penelope G. (1999). Children's Understanding of Mind and Emotion: A Multiculture Study. *Cognition & Emotion* 13(1), 19 – 48.

Vygotsky, Lev S. (1962). *Thought and Language*. Cambridge, MA: MIT Press.

Wassman, Jrg, Birgit Truble, and Joachim Funke (2013*). Theory of Mind in the Pacific: Reasoning Across Cultures*. Heidelberg: Universittverlag Winter Heidelberg.

Weisberg, Deena and Alison Gopnik (2013). Pretense, Conterfactuals, and Bayesian Causal Models: Why What Is Not Real Really Matters. *Cognitive Science* 37(7), 1368 – 1381.

Wellman, Henry M. (1990). *The Child's Theory of Mind.* Cambridge, MA: MIT Press.

Wellman, Henry M., David Cross, and Julanne Watson (2001). Meta-Analysis of Theory-of-Mind Development: The Truth about False Belief. *Child Development* 72(3), 655 – 684.

Wellman, Henry M. and Candida C. Peterson (2013). Deafness, Thought-bubbles and Theory of Mind development. *Develomental Psychology* 49(12), 2357 – 2367.

White, Aaron S., Valentine Hacquard, and Jeffrey Lidz (2018). Semantic Information and the Syntax of Propositional Attitude Verbs. *Cognitive Science* 42(2), 416 – 456.

Whorf, Benjamin Lee (1956). *Language, Thought, and Reality: Selected writings of Benjamin Lee Whorf.* MIT Press.

Willet, Thomas (1988, January). A Cross Llinguistic Survey of the Grammaticalization of Evidentiality. *Studies in Language* 12(1), 51 – 97.

Wilson, Deirdre and Dan Sperber (2002). Pragmatics, Modularity and Mind-reading. *Mind & Language* 17(1 – 2), 3 – 23.

Wimmer, Heinz, G.-Jrgen Hogrefe, and Josef Perner (1988). Children's Understanding of Informational Access as Source of Knowledge. *Child Development* 59 (2), 386 – 396.

Wimmer, Heinz and Joseph Perner (1983). Beliefs about Beliefs: Representation and Constraining Function of Wrong Beliefs in Young Children's Understanding of Deception. *Cognition* 13, 103 – 128.

Winawer, Jonathan, Nathan Witthoft, Michael C. Frank, Lisa Wu, Alex R. Wade, and Lera Boroditsky (2006). Russian Blues Reveal Effects of Language on Color Discrimination. *PNAS* 104(19), 7780 – 7785.

Wittgenstein, Ludwig (1953). *Philosophical Investigations*. London, UK: Macmillan Publishing Company.

Wolff, Philipp and Kevin Holmes (2010). Linguistic Relativity. *Cognitive Science* 2, 253 – 265.

Woodward, Amanda L. (2003). Infants' Developing Understanding of The Link Between Looker and Object. *Developmental Science* 6(3), 297 – 311.

Woolfe, Tyron, Stephen C. Want, and Michael Siegal (2002). Signposts to Development: Theory of Mind in Deaf Children. *Child Development* 73(3), 768 – 778.

Xia, Tiansheng, Guiping Xu, and Lei Mo (2019). Bi-lateralized Whorfian Effect in Color Perception: Evidence from Chinese Sign Language. *Journal of Neurolinguistics* 49, 189 – 201.

Yirmiya, Nurit, Marian Sigman, and Daphna Zacks (1994). Perceptual Perspective-Taking and Seriation Abilities in High-Functioning Children with Autism. *Develomental Psychopathology* 6(2), 263 – 272.

Yoshida, Hanako and Linda B. Smith (2005). Linguistic Cues Enhance the Learning of Perceptual Cues. *Psychological Science* 16(2), 90 – 95.

Zaitchik, Deborah (1990). When Representations Conflict with Reality: The Preschooler's Problem with False Beliefs and False Photographs. *Cognition* 35 (1), 41 – 68.

Zhou, Ke, Paul Kay, Veronica P.Y. Kwok, Tiffany N.M. Ip, and Li Hai Tan (2010). Newly Trained Lexical Categories Produce Lateralized Categorical Perception of Color. *PNAS* 107(22), 9974 – 9978.

Ziatas, Kathryn, Kevin Durkin, and Chris Pratt (1998). Belief Term Development in Children with Autism, Asperger Syndrome, Specific Language Impairment, and Normal Development: Links to Theory of Mind Development. *Journal of Child Psychology and Psychiatry* 39(5), 755 – 763.

Zipoli Caiani, Silvano (2010). Enacting Communication. Evidence Supporting an Embodied Theory of Language. *RIFL* (Rivista Italiana Filosofia del Linguaggio) 0(2), 152 – 176.

Subject index

https://doi.org/10.1515/9783110748475-010