Tel Aviv University

The Jaime and Joan Constantiner

School of Education

The Linguistic Consequences of Acquired Damage to Theory of Mind

Thesis Submitted for the Degree "Doctor of Philosophy"

By

Noga Balaban

Submitted to the Senate of Tel Aviv University

April 2010

This work was carried out under the supervision of

Prof. Naama Friedmann and Prof. Mira Ariel

The research was funded by a grant from The Israel Foundations Trustees (Doctoral Students Program No. 28)

Contents

Abstract

Theory of mind (TOM) is the cognitive ability that allows a person to distinguish between various mental viewpoints held by different people, and to relate mental states, intention and beliefs to behavior in various social situations. It allows a person to understand that others may hold knowledge and beliefs that are different from hers, and therefore it is a basic ability for every social interaction and it is crucial for developing communicating abilities. A deficit in TOM creates emotional and behavioral difficulties. One of the most central everyday activities that is dependent on this ability is taking meaningful part in discourse; In other words, understanding implicated and hidden meanings, presuppositions that are not stated explicitly, and nonliteral language.

Right brain damaged (RBD) patients have been found to have difficulties in various discoursive tasks. They have difficulties in producing coherent discourse, comprehending ironic and metaphorical expressions, grasping the gist of stories and punch lines in jokes. In light of these findings, the research in the past twenty years related between TOM and RBD. Imaging studies of brain-damaged patients and healthy individuals demonstrated this relation between TOM and intact right hemisphere.

Therefore, damage to the right hemisphere might be expressed in a speaker's disability to understand that her addressee holds a perspective that is different from hers. Accordingly, it can be expected that this impairment will affect linguistic abilities. The linguistic abilities that are expected to be affected are those that are dependent on the interlocutors' need to consider the knowledge each of them holds regarding the information mentioned during the discourse. This question, about the linguistic abilities of patients who suffer damage to TOM has not been thoroughly researched from this aspect before. We asked in this study in what way acquired damage to TOM affects specific linguistic abilities of RBD patients.

The research was conducted in two phases. The first stage identified individuals who suffer damage to TOM among a group of RBD patients. In order to identify such a group, we created a battery of tasks, comprised of various tasks that demonstrate different aspects of TOM (henceforth, the aTOMic battery). We used this battery to

test the performance of a group of 25 RBD patients and compared their performance in the aTOMic battery to that of a control group. The main finding from this phase of the research was that the group of RBD patients was heterogenic; 17 of the RBD patients were aTOMic, and 8 of them actually showed TOM ability which was no different that that of the controls.

In the second stage of the research we tested both groups of RBD patients, the aTOMic and the RBD patients who do not suffer damage to TOM, on various linguistic tasks that rely on the need to take the others point of view into consideration during conversation.

The linguistic abilities we tested were:

The ability to use the definiteness system to introduce new entities into the discourse. This task requires that the speaker monitors her addressees' acquaintance with the entities she introduces into the discourse, and to distinguish which are new to him. Adding the definite marker expresses the speaker's assumption that her addressee is familiar with the item presented. For example, the sentences: 'She submitted her paper', 'She returned the files', 'She left the room' are properly used in conversation only if the speaker knows that the addressee knows which 'paper', 'file' and 'room' she is referring to. This ability was tested using three different experiments, assessing both production and comprehension.

Another related ability we tested was the ability to use and comprehend reference terms. We tested whether there is a difference between the way aTOMic patients and RBD patients who are not aTOMic (*TOMers)* use the reference terms their language offers. Do they use them in a manner that allows their addressees to identify the referent they intended in each stage during the discourse? And as addressees, can they identify the referent the speaker is referring to? For example, opening a description of an interaction that takes place between two people with a pronoun (He promised to…) without an earlier introduction of the characters, does not allow the proper identification of the intended referent. This ability was tested using 6 different experiments. Two of the experiments tested the participants' ability to produce appropriate reference terms; 4 other tasks tested their metalinguistic understanding of the difference between the various terms.

Another linguistic ability we tested was the way the aTOMic and TOMers understand and use mental state verbs. The lexical meaning of mental verbs like *knew*, *thought*, and *remember* contains aspects regarding the mental understanding of the discourse situation. Understanding the way aTOMic and TOMer patients comprehend and use the lexical characteristics of these verbs is important to the understanding of the relation between TOM and linguistic abilities. This ability was tested using 4 different experiments.

In addition, we used another 4 experiments to test other linguistic abilities of these patients, such as the comprehension and production of relative clauses and Wh questions using sentence-picture matching and sentence elicitation tasks, and the mastery of Chomsky's Binding principles by pronoun comprehension tasks. These abilities are not related to TOM but they characterize other populations who suffer syntactic impairments, such as agrammatism. We added these tasks in order to assess whether the linguistic disabilities of aTOMic patients are related to some general linguistic failure. We also tested the aTOMics' and TOMers' ability to produce sentential complements which are considered by some as the precursor for TOM development in young children.

The results of the linguistic tasks show that the aTOMics succeeded consistently worse than the patients with intact TOM only in the linguistic tasks that were TOM-related. The aTOMics had difficulties in all the TOM-related linguistic tasks we presented, while the other group performed as well as the control group. On the syntactic tasks that were not related to TOM, and on the task of producing embedded sentences, both groups performed similarly to controls.

From these results we learn, first, that the people who suffer right brain damage do not form a homogenous group. Therefore there is a need to assess, for each individuals with RBD, whether his TOM is impaired or not. We hope the tools we created during this research can be used as clinical diagnostic tools and can aid the quality of treatment of this unique group of patients.

Second, we found that RBD patients might suffer specific linguistic impairment, although they did not suffer damage to the linguistic centers in the brain. Whether or

not a linguistic ability was impaired crucially depended upon whether or not it required TOM.

Therefore, we conclude that damage to TOM creates specific language difficulties. These difficulties are not accompanied by broader linguistic impairment, and appear only in linguistic abilities which are related to Theory of Mind.

1. Introduction

1.1. Theory of Mind

The aim of this study is to describe the effect of acquired damage to Theory of Mind on several linguistic abilities. Theory of Mind (TOM) is defined as the cognitive ability to relate behavior to mental states. It is the human capacity to attend to beliefs, intentions and desires as the basis for human action, both of our own and that of others (Astington & Jenkins, 1995; Happé, 1994, 1995; Malle, 2005; Scholl & Leslie, 1999).

TOM is fundamental to social cognition in at least two ways: First, from a *developmental perspective*, its attainment is the basis for social understanding and interacting. From a very early age we take part in interactions with other people, attending to human gestures and facial expressions. Very quickly the representations we create allow us to make sense of our own behavior and to predict future actions of others (Leslie, 1987; Meltzoff, 1999; Roth & Leslie, 1998; Wellman, 2004). Later on, when linguistic abilities mature, TOM is put to use in more sophisticated social situations, from interaction between peers (Slomkowski & Dunn, 1996) to teaching (Strauss, Ziv, & Stein, 2002; Ziv & Frye, 2004; Ziv, Solomon, & Frye, 2008). All of these are dependent on the basic ability to represent the mind of another.

Second, from an *ontological perspective*, TOM is essential because people cannot see the world without establishing connections between mental states and behaviors. We tend to attribute intentions and react to the intentions of others (Leslie, 1987; Meltzoff, 1995; Papafragou, 2002). We keep track and update our representations of these intentions and goals during interaction, and we adjust our representation of the situation accordingly. Even if full understanding of the other is forever beyond us, the tendency to try and make sense of the others' intentions, thoughts and feelings, is always present (Grice, 1975, 1989; Malle, 2005).

Deficits in TOM arise in various pathological states; Autism and Asperger Syndrom are critical examples and they are regarded as developmental disorders that include TOM deficits (Baron-Cohen et al., 1996; Hale & Tager-Flusberg, 2005; Rajendram & Mitchell, 2007). The present research will focus on TOM deficits in adults who suffered a brain damage to their right hemisphere, and as a consequence acquired

damage to TOM. Multiple brain imaging studies have shown a connection between performance in TOM tasks and specific brain locations, especially in the right hemisphere (Firth & Frith, 2003, 2006; Sampson, Apperly, Chiavarino, & Humphreys, 2004; Saxe & Wexler, 2005). Patients who suffer CVA in their right hemisphere (RH) show difficulties in various tasks which require understanding of social interactions (Happé, Brownell, & Winner, 1999; Shamay-Tsoory, Tomer, & Aharon-Peretz, 2005). These patients also show various difficulties in discourse production and comprehension. Their discourse lacks coherence (Johns, Tooley, & Traxler, 2008; Rehak, Kaplan, & Gardner, 1992) and they have difficulties in comprehending stories and cartoons that include knowledge gaps between characters (Happé et al., 1999). They also show insensitivity to prosodic cues (Pell, 1999, 2006). All these abilities are linked to the basic cognitive ability, to represent the mind of the other.

The current study was conducted in two phases. In the first phase we tested patients who suffered damage to various locations in their right hemisphere, on a number of tasks that assess various aspects of TOM. Our aim was to see whether these patients suffer from loss of TOM (which we will henceforth term *aTOMia*). We wanted to find out whether all patients who suffer damage in the right hemisphere also manifest a deficit in TOM. We also wanted to portray a comprehensive picture of their TOM, to see which aspects are more prone than others to be affected.

In the second phase of the study we tested these patients on linguistic capabilities that rely on a need to represent others' minds while engaging in discourse. We were interested in finding out whether a deficit in TOM resulted in a poorer linguistic performance, or not.

1.1.1. Defining the Problem

Patients who suffer damage to their right hemisphere are prone to several cognitive and motor disabilities that can create difficulties in different aspects of life. In some cases the brain damage has a direct effect on patients' physical and mental abilities and in most cases it has some indirect effect on their emotional well being. These patients' ability to communicate effectively is an important issue in every respect.

The topics we are interested in here are the linguistic abilities related to TOM. We asked whether patients who suffer aTOMia also show difficulties in linguistic tasks that rely on the ability to consider different viewpoints. One is the ability to properly introduce items into discourse using the definite or indefinite article. Another is the ability to use referring expressions of different kinds in a way that makes certain the listener will be able to identify the referent the speaker intended. A third important linguistic ability that relies on TOM is the ability to understand and put to use different lexical features of mental state verbs. We will also consider the linguistic abilities that were not affected by the damage to TOM. These will be the topics of our research of the linguistic abilities of patients who suffer aTOMia (aTOMic) but first in order to define the scope of this study, it is important to clarify how we perceive the relations between the different aspects of communication abilities of RBD patients.

Theory of Mind and communication

The interest in RBD patients' ability to effectively communicate, to comprehend spoken messages and produce them appropriately is an issue that has been treated from multiple viewpoints (Meltzoff, 1999; Myers, 2001).

Studies that examine *discourse abilities* show that RBD patients have difficulty in understanding and creating coherent narrative (Johns, 2008; Marini, Carlomago, Catagirone, & Nocentini, 2005). Some researchers highlight the important role of the right hemisphere in comprehending narrative (compared to understanding different kinds of unrelated linguistic materials) (Mar, 2004), in understanding central themes in discourse (Hough, 1990) and in efficiently understanding and directing the flow of conversation (Rehak et al., 1992). Other studies deal directly with communicative tasks and show a correlation between RBD and a disability to effectively participate in discourse. RBD patients were found to make inappropriate use of personal reference terms (Brownell, Pincus, Blum, Rehak, & Winner, 1997) and to produce incoherent texts when measured by their use of referential terms (Davis, O'neil-Pirozzi, & Coon, 1997). They were also unable to adjust the form of their requests to various social situations (Brownell & Stringfellow, 1999; Weylman, Brownell, Roman, & Gardner, 1989), to understand jokes (Coulson & Williams, 2005) and to recognize the difference between mistakes and deceits (Adenzato & Bucciarelli, 2008).

One of the important abilities that was studied regarding these patients' discourse capabilities is their ability to draw inferences during reading. Using priming paradigm, studies showed that these patients had difficulty in creating bridging inferences that helped guide the comprehension process of narratives (Beeman, 1993; Beeman, Bowden, & Gernsbacher, 2000). Still, other studies showed that if the cognitive load during online testing was reduced, RBD patients were able to draw such inferences (Tompkins, Fassbinder, Blake, Baumgaertner, & Jayaram, 2004), and they could also draw predictive inferences which were not obligatory for understanding narratives (Tompkins, Lehman-Blake, Baumgaertner, & Fassbinder, 2001). According to Grice (1975, 1989) these conflicting results are especially important to the understanding of RBD patients' communication profile because the ability to draw inferences is regarded as the basis of the broad ability to understand intentions.

Theory of Mind and Pragmatics

Several researchers focused on the *pragmatic competence* of RBD patients (Martin & McDonald, 2003; McDonald, 2000; Penn, 1999; Surian & Siegal, 2001). These studies present findings about RBD patients' low performance on different tasks that are predefined as pragmatic (e.g. overly literal comprehension, understanding humor and sarcasm, poor prosody understanding, idiom comprehension) and discuss the reasons underlying these disabilities. In their review Martin and McDonald (2003) explain these difficulties as resulting from either an inability to use context, or a difficulty to use different perspectives or as a result of rigid and concrete information processing. They do not present a decisive argument about which theory best explains the multi- aspect difficulties of RBD patients. Summing up their discussion they say:

> *"The notion of pragmatic language is broad reaching, encompassing a wide range of contextual influences on language meaning and a variety of modes of behaviour. Reflecting this, the causal basis of pragmatic language is difficult to define and indeed is unlikely to be unitary." (Martin & McDonald, 2003, p. 462)*

Part of the difficulty to define pragmatic deficiencies, can be ascribed to the intricate discussion within language studies about the definition of the scope of pragmatics

(Ariel, 2008). This difficulty has also affected the way the term was applied in clinical studies (Blake, Myers, & Tompkins, 2002; Penn, 1999; 2000; Perkins, 2005). To date, in some cases the different researchers' viewpoints converge and focus on the same ability, although the definitions of the topic in discussion may vary. For example, following Grice (1975, 1989) the ability to understand nonliteral language like irony or sarcasm is considered a pragmatic ability by most (e.g., Kasher, 1991). But it is taken as an example of social understanding by some (Shamay-Tsoory et al., 2005) and as a constituent of TOM ability by others (Happé, 1994, 1995).

While there is no reason to either subject or subsume one explanation to another, it is important to clarify the way we see the relation between pragmatics and TOM in the current research. As will be explained bellow, we follow the Gricean notion of comprehending intents, as the central component of TOM ability.

Grice (1989) defined the ability to communicate, both verbally and non-verbally, as the ability to *mind read*, the ability to infer communicative intentions from spoken utterances. Later on, theoreticians within the Relevance theory framework treated pragmatics as a sub-module within the TOM module dedicated to inferring speaker meaning from her overt expressed sentence meaning (Sperber & Wilson, 2002; Wilson, 2005). TOM is regarded and studied here as the cognitive ability uniquely dedicated to processing information regarding the mental realm (Happé, Winner, & Brownell, 1998; Happé et al., 1999). While this study was not designed to tackle the issue of modularity and the independence of TOM from other mental capacities (e.g., executive functioning), we support the view that TOM is an independent module dedicated to the processing of human intentions, desires and beliefs. This is based on theoretical explanations (Fodor, 1992; Sperber & Wilson, 2002; Wilson, 2005), developmental research (Fodor, 1992; Leslie, 2000; Scholl & Leslie, 1999), research about Autism and Asperger Syndrome (Baron-Cohen, Jollifffe, Mortimore & Roberstone, 1997), and evidence from brain scanning of healthy people (Frith & Frith, 2003; Saxe & Powell, 2006; Saxe & Wexler, 2005) and autistic patients (Mason, Williams, Kana, Minshew, & Just, 2008; Ozonoff & Miller, 1996). Although the question of whether pragmatics is an independent module (within the TOM module) is also debated in the literature (Kasher, 1991; Perkins, 2005; Sinclair, 1995), the idea that pragmatic ability is crucially depends on TOM ability is an important one for the current research. The ability to understand the speaker's knowledge and intents

conveyed in a sentence is considered here to be parallel to the ability to understand intent from witnessing behavior. Because the tasks we presented in the current research to assess TOM ability were a series of short stories and the data we gathered were orally given answers, our results might be subject to the criticism that we are in fact studying pragmatic abilities and not TOM. Because the current characterization of pragmatic ability is as part of TOM anyway, this concern is not a regarded as a problem.

1.1.2. Theory of Mind and linguistic abilities

The two fields, of language acquisition and TOM development are interrelated. A large amount of developmental studies were devoted to understanding the relations between the two abilities along the different stages of social development (Cutting & Dunn, 1999; Slomkowski & Dunn, 1996). An example of the complex connection is the ability to *join attention* with another person in attending to a third object. This ability appears during the first year of life (Carpenter, Nagell, & Tomasello, 1998) and later gives rise to other cognitive developments like peer play and shared pretend play which researchers consider the basis of social interaction (Leslie, 1987; Tomasello & Habrel, 2003). Joint attention is also one of the first milestones in the path to mature TOM (Wellman, Cross & Watson, 2001). During the first phases of acquiring language this basic ability is crucial for directing the child to attend to objects while they are named. This activity begins at the same time as the process of acquiring the mental lexicon (Happé & Loth, 2002).

Another important relation between language and TOM that has been thoroughly researched is the connection between TOM and complex sentences. Several researchers suggested that there is a unidirectional process of development between the two. Only once children are able to represent embedded proposition they become capable of representing false beliefs (J. de Villiers, 2003, 2007; J. de Villers & Pyers, 2002; Hale & Tager-Flusberg, 1999).

It is important to note here that the components that are crucial in development of TOM and those that take part in the mature cognitive ability are not necessarily identical. Theoretically, language development might be essential in the process of

TOM development and still be detached from its operation once it has matured. For example, different linguistic capabilities (semantic, syntactic and pragmatic) were discussed as preconditions for false belief understanding. But only research about adult theory of mind can answer the question of the importance of these abilities in mature TOM reasoning (Apperly, Samson, & Humphreys, 2009). Our interest is whether difficulty in TOM reasoning affects other linguistic abilities. In the second part of the research we will present the question whether some linguistic abilities can be compromised due to a deficit in TOM, although no direct damage to the brain areas that are responsible for language abilities was suffered.

In the next section we will describe the origins of TOM in normal and abnormal development. We will describe the development of methodologies to study TOM as well as the main findings accumulated in these studies. We will then review the findings about TOM and the right hemisphere, on-line experiments using brain imaging studies and findings from off-line experiments conducted with brain damage patients.

1.1.3. Development of Theory of Mind

Normal development

The research on TOM development originally focused on mastering *false belief* understanding. This issue attracted a great amount of attention because it was marked as a defining feature of the important ability to realize that beliefs are detached from reality. Wimmer and Perner (1983) first introduced the well known "Maxi and the chocolate" task that was designed to test this ability. Children were shown a doll figure, Maxi, hiding chocolate in a blue cupboard and leaving the room. When Maxi was out of the room a doll figure representing his mother came into the room and moved the chocolate to a green cupboard. After the mother left, Maxi returned to the scene and the children were asked: *Where will Maxi look for his chocolate?* In their pioneering study the researchers reported that consistent correct responses (e.g., that Maxi will look for his chocolate where he left it), appeared between the ages of 4 and 6 years. Younger participants, who were not yet able to detach their true beliefs from Maxi's false belief, tended to predict he would look for the chocolate in the place where it really was. The universality of this finding has been confirmed in a meta-analysis of more than 150 studies (Wellman et al., 2001), and the reliability of

different tasks testing this ability have also been confirmed (Hughes, Adlam, Happé, Jackson, Taylor, & Caspi, 2000). The significance of this finding is that children, approximately 5 years old, demonstrate an ability to suppress their own knowledge about a situation and base their prediction about someone else's actions on that person's knowledge, even though it contradicts their own, more accurate knowledge.

Another task that was presented by Perner and Wimmer (1985), in order to try and define stages along TOM development was *second order false belief*. In this task children were asked to describe one character's thoughts about another character's thoughts. Following the same methodology used in the first order false belief task, a short story was acted out using toy dolls. In the original story John and Mary were playing together in the park when an ice cream van appeared. Later on John and Mary were each informed separately that the ice cream van moved to a different location but John didn't know that Mary was also notified. The question presented was: *Where does John think Mary will go to buy ice cream*. The expected answer was that John thinks Mary will go to the park because John does not know that Mary knows the car has moved to a different location. This answer was provided only by children aged seven. A two year difference was found between children who understood first order false belief to those who understood second order false belief (Perner & Wimmer, 1985). Another recent study using a different second order false belief story showed that 7 years olds are able to give second order answers, but they do not always do so. They also give first order answers when second order ones are due (Hollebrandse, Hobbs, de Villers, & Roeper, 2007). Other studies showed that when different stories are used and the cognitive load is lessened by additional feedback and memory aids the age difference between children who understood first order false belief to those who understood second order false belief is erased (Sullivan, Zaitchik, & Tager-Flusberg, 1994).

These findings describe TOM as a naturally developing cognitive ability (Leslie, 1987; Sullivan et al., 1994; Surian & Siegal, 2001) that might be masked when using complex tasks in order to test it. Consequently, when turning to the study of acquired damage to TOM it is important to include tasks of different complexity and different modes (verbal and non verbal) to ensure that the preserved capabilities will be evident.

The inquiry about TOM development branched later to earlier development during infancy (Meltzoff, 1999; Wellman & Liu, 2004), and later development during childhood (Baron-Cohen, O'Riordan, Stone, Jones, & Plaisted, 1999; Happé, 1994). One of the interesting points along this development is the emerging ability to manipulate the content of thought of others using lies. This ability demonstrates the capacity to monitor other people's state of knowledge. In certain instances a person telling a lie produces false statements intended to preserve a false understanding, in others he intends to change a true belief or understanding.

In both situations the person lying must also conceal his own knowledge that contradicts the lie (Adenzato & Bucciarelli, 2008; Talwar, Gordon, & Lee, 2007). Studies about the development of understanding and producing deceits showed that children as young as 3 years old spontaneously lie about their behavior (peeking at an object when directed not to do so) (Polak & Harris, 1999). Later on, at the age they understand false belief, children lie according to the false belief they try to create. Older children, between 6-11 years gradually develop the ability to maintain their initial lie with other statement that back up the false belief they created. This ability was correlated with the performance of second order false belief (Talwar et al., 2007).

Another everyday practice that lends itself to testing of TOM is the act of teaching. Teaching as an informal event can be perceived as the opposite of lying. While lying is intended to cause a false belief, teaching is an effort to bridge over knowledge gaps, and in a successful act of teaching we are able to change in some way the point of view of the student towards events in reality or more abstract ideas. This act requires the teacher first, to appreciate there is a knowledge gap and secondly to have an intent to bridge that gap (Strauss et al., 2002; Ziv & Frye, 2004). The act of teaching is in some sense the opposite of maintaining a lie, but in both cases there is a need to attend, on line, to the mental processes of the partner in the interaction in order to diminish (in the case of teaching) or strengthen (in the case of lying) his false belief about the situation (Strauss et al., 2002; Talwar et al., 2007; Ziv et al., 2008). Studies regarding teaching showed that three years olds have some understanding of teaching but their reliance on demonstration indicated they were considering teaching to be a behavioral task, not a mental one. Five years olds taught very differently, they relied much more on verbal explanations, were attentive to the learners' questions and asked questions to check learner's understanding. This transition is parallel to the advances

in false belief understanding within preschool years (Strauss et al., 2002; Ziv et al., 2004; Ziv & Frye, 2008).

In the current study we explored RBD patients abilities in these TOM related activities of teaching and understanding social situations in which lies are told. We expect that if a patient has TOM deficit, his- or her reactions in these social situation will demonstrate this difficulty.

Abnormal development

Autism is a neuro-developmental syndrome that is characterized by deficits in social understanding and communication. The first signs of difficulty appear before the child is three years old. At this age, toddlers show inability or irregularity of word acquisition and an inability to use words to communicate, even though they may have the ability to recite passages they heard and also a tendency to repeat words spoken to them. Autism varies in degree of severity and pattern of difficulty. It is generally assumed that there is a *spectrum* of autistic disorders (ASD). Asperger syndrome (AS) is a milder variant and is distinguished from ASD by a lack of linguistic or cognitive delay (U. Frith, 2001). Besides delayed language acquisition, ASD children also show a lack of awareness to others, a lack in pretend play and a failure to respond to facial expressions and feelings of others (Lord, Cook, Leventhal, & Amaral, 2000). Baron-Cohen et al. (1996) in a large scale screening research showed that early developmental signs of mentalizing ability are good prediction measures for ASD: an inability to follow another person's gaze, lack of behavior of pointing at or showing objects of interest and not engaging in pretend play. Later in life these children develop narrow or unusual interests, they do not develop normal social ties with their peers and resist changes to their daily routine (Lord et al., 2000; Smith, 2008).

Since the mid 1980's many of the cognitive difficulties that ASD patients suffer are explained as lack of TOM ability. The first findings showed a significant difference between the performance of ASD children and normal children in understanding standard first and second order false belief tasks. These results were replicated with a different false belief task, the deceptive box. In this task children are shown a box of candy and asked to guess what is inside (the expected answer is candy). The

experimenter then opens the box, which actually contains pencils, takes the pencils out and shows them to the child. In the next stage the pencils are returned to the box and it is closed again. Now the child is asked what s/he thought earlier to be in the box. If the child answered that he earlier thought the box was filled with candy the researchers can conclude that he is able to reason about a false belief of his own normal developing children perform this task successfully at the same age the accomplish the standard false belief task (Hughes et al., 2000). ASD participants had difficulty in performing the deceptive box task (Rajendran & Mitchell, 2007). These findings led researchers to term these children's condition as "mindblindness", a complete inability to understand the mind (Baron-Cohen, Leslie, & U. Frith, 1985, 1989; U. Frith, 2001; Happé, 1995).

Later it was shown that able autistics can pass first and second false belief (Bowler, 1992).Researchers also noted that in the original studies only 20% of ASDs successfully completed the first order false belief task (Baron-Cohen et al., 1985; Happé, 1995). Baron-Cohen et al. (1999) tested a group of high functioning children with autism or Asperger syndrome and showed that although they were able to pass first order false belief tasks, they were impaired in their ability to understand more complex TOM tasks. In light of these findings the ASD and Asperger syndrome patients' difficulty was described as a developmental *delay* and not as a deficit. To date, it is widely agreed that along the autistic spectrum there is a shared difficulty to understand mentalizing, but its manifestation is dependent on the individual profile of the patient and the tasks used to test TOM (Baron-Cohen et al., 1999; Jollife & Baron–Cohen, 1999; Silliman, Diehl, Bahr, Hnath-Chisolm, Zenko, & Friedman, 2003; see Rajendran & Mitchell, 2007 for a review).

In order to adapt tests to meet the different disabilities, linguistic and cognitive, of individuals along the ASD spectrum, researchers devised different kinds of tasks that differed in their reliance on linguistic abilities and varied in the processing load they created. For low functioning children with autism researchers used a non verbal task. This was done by acting out the Maxi situation by two experimenters[1] (Colle, Baron-Cohen, & Hill, 2007).

[1] One of the experimenters was the 'hidder' the other the 'communicator'. The child was familiarized with the situation in which the communicator shows him which box to choose in

For high functioning children and adults with autism new tasks were devised. Happé (1994) was the first to present a battery of stories about domestic events in which *people say things they do not literally mean*. She titled them "strange stories" although they describe situations which are very familiar and mundane; they are experienced as strange only if the participant has difficulty in representing the intentions of the characters in the story. For example, one of the stories was about a boy who broke an expensive vase his mother treasured and when asked what happened, he blamed the dog for kicking it over. The participants were asked whether what the boy said was true (*no*) and why he said it (*in order to avoid punishment*). Other stories presented situations in which people say white lies or ironic remarks or respond in other biased ways to different situations. Findings showed that ASD children, who passed first order false belief, scored lower than younger normal controls and lower than mentally handicapped group, although the mentally handicapped participants received lower scores on verbal IQ measure than the ASD children (Happé, 1994). The same pattern of results, mentally based stories being the most difficult for able autistic children to understand, was also found using a revised set of stories which included a variety of topics (mental, human, animal, and nature) (White, Hill, Happé, & U. Frith, 2009).

The "strange stories" task is important not only because it presents the implementation of TOM ability to everyday events and varies the means in which researchers can assess TOM ability. It is revealing also because it asks for a justification in addition to the yes-no decision. Justifications are interesting because they can inform about the elements in the story to which the participants attend to while reasoning. The original false belief task only elicits inferences regarding another’s mind. The strange stories elicit reasoning about interpersonal relationships and therefore they are more informative regarding the participants' understanding of psychological and mental concepts (about the importance of reasoning: Amsterlaw, Lagattuta, & Meltzoff, 2009; Silliman et al., 2003).

---

order to receive the chocolate. In the test trial the communicator was out of the room while a switch was made. The child had to choose the correct box with the candy in it while ignoring the false belief of the communicator. The group of ASD children was compared to a group of SLI children who are, like the ASD group, late in language acquisition. Results showed that the ASD children scored significantly lower than the SLI group in understanding false belief but not in true belief and control tasks.

Another type of task that was presented in order to assess TOM ability in children older than 6-7 year olds and ASD patients is the understanding of embarrassing misunderstandings which are termed *faux pas* (Baron-Cohen et al., 1999). *'Faux pas'* is an awkward social situation in which a protagonist says something that is insulting to another character due to insufficient information about the situation (for example if someone makes an insulting remark about a third person and that person overhears the comment). Understanding these situations, i.e., the embarrassment the protagonist will experience when he finds out the missing information (that he was overheard), is a good measure for advanced TOM understanding. It requires an appreciation of the knowledge gap between the two characters in the story and the identification of the negative effect the statement causes to the participants. Research showed that this task was more demanding than understanding second order false belief. Normally developing children at the age of seven, who succeeded in the second order false belief, did not perform above chance level on the *faux pas* stories. Only 11 year olds succeeded in understanding the majority of *faux pas* stories presented. Children diagnosed with Asperger syndrome and high functioning autistic children scored significantly worse on the *faux pas* stories than on equally complex, but not mental, control stories. Both groups reached the maximum on control stories. The researchers concluded that although high functioning children with autism or Asperger syndrome can identify mental states of the characters individually, they can't unify this information and comprehend the *faux pas* situation and its impact on the different characters (Baron-Cohen et al., 1999). This kind of task (and some of the specific items from these studies) was later used to evaluate RBD patients' ability to understand social situations (Shamay-Tsoory et al., 2005).

Another important kind of stimuli that was used to assess TOM is the ability to understand pictures or cartoons that expressed mentally loaded situations. For example, Losh and Capps (2003) presented children diagnosed as high functioning Autistic or Asperger aged 8-14 years a picture story book (*Frog, where are you?* Mayer, 1969) and asked them to tell the story. Their narratives were compared to a control group matched on chronological age and verbal IQ. Results showed that children with autism or Asperger syndrome expressed less causal explanations both to behavior and to internal states (Losh & Capps, 2003).

The tasks used in the research of normal development and ASD allows a multi dimensional assessment that includes different kinds of reasoning tasks: yes/no decisions as in first and second order false belief, and justifications that allow an explicit analysis of the factors participants are attending to while reasoning. Using a battery of tasks that include a variety of stories and situations can, on the one hand, ensure that if a person has a problem in some aspect of TOM, that problem will be detected, and on the other hand it allows for different reasoning processes to surface, ensuring that there will be opportunities to demonstrate variable TOM ability.

1.1.4. Theory of Mind after brain damage

During the past 30 years language therapists, neuropsychologists and language researchers characterized by different research methods the mental and cognitive affects of damage to the right hemisphere. The different findings point to a deficit in the ability to communicate effectively and to understand communicative intentions. However, most of the studies find large variability among participants' performance and a unified explanation of these findings, as well as title for these difficulties is not in agreement (Joanette & Anslado, 1999; Martin & Mcdonald, 2003; Myers, 2001).

Apart from studying TOM and its relation to other mental abilities, sub-divisions within TOM were assessed separately. One distinction is between a basic stage of *attribution* of the relevant mental state (e.g., understanding that someone is angry or sad), and an advanced stage in which people *infer the cause* of that mental state (e.g., someone is sad and angry because he was betrayed by a friend). The ability to *attribute* mental states is assessed by tasks that show pictures of facial expressions, especially of the eye region, or by sounding short vocalizations, and asking participants to name or choose the mental state they perceive to see or hear. These studies show that high functioning ASD and AS have difficulty deciding which mental state they are seeing or hearing (voice: Rutherford, Simon Cohen & Wheelwright, 2002; eyes: Baron-Cohen, Wheelwright, Hill, Raste & Plumb, 2001). This ability is associated with the right hemisphere (Fournier, Calverley, Wagner, Poock, & Crossley, 2008; Sabbagh & Flynn, 2006).

In the current study we are interested in the more advanced part of TOM that allows a person to attribute specific mental content to another person and enables him to reason about those attributions. Advanced theory of mind is studied using different tasks in which participants are asked to explain the mental causes for events they are presented with. These tasks are more linguistically demanding but can help us gain an understanding about the way mental events are understood (Happé, 1994; Happé et al., 1999).

In the next section we will review studies that ask about the effect of damage to the RH on the ability to understand and reason about mental events. First, studies about specific brain locations activated during TOM reasoning will be described followed by off-line studies which demonstrate the same abilities.

The anatomical location of Theory of Mind

A central question in TOM research concerns the brain anatomical locations that are responsible for TOM reasoning. Several studies addressed this issue by scanning brain activity while participants perform different tasks. Participants are presented with (at least) two kinds of tasks: TOM tasks which demand representation of some mental aspect (intentions, motives, mental knowledge, beliefs etc.) and control tasks that are similar in all aspects but lack the TOM element. For example, Berthoz, Armony, Blair and Dolan (2002) presented short stories of embarrassing social situations to a group of healthy adults. The element that was manipulated was whether the violation of social norms was intentional or not. For example, one of the stories was about a woman named Joanna who was invited to her friend's house for a Japanese dinner. In the 'unintentional condition' Joanna chocked on a bite of the first course and spit it out while coughing. In the 'intentional condition' it was said that Joanna didn't like the food and spit it back out on her plate. In addition to the brain areas that were activated in both task, they also found significant differential activation. The brain areas that are active only during TOM tasks are considered responsible for the ability to reason about TOM (Bird, Castelli, Malik, U. Frith, & Husain, 2004; Siegal & Varley, 2002). Studies that used pet scans and f-MRI (Berthoz et al., 2002; review: Frith & Frith 2003; 2006) found consistent activation during TOM tasks in a number of brain locations: the temporal poles (TP) bilaterally, the posterior superior temporal sulcus

(pSTS) and adjacent areas of the parietal and temporal lobes (TPJ: temporo-parietal junction), bilaterally with greater effect on the right. Same areas were also activated, with some abnormal pattern in a group of ASD adults while attending to TOM tasks (Mason et al., 2008) However, the brain location Frith and Frith (2003;2006) maintained to be the key to mental reasoning is the medial prefrontal cortex (MPFC). They came to this conclusion by comparing results of studies that elicited mentalizing, to others, that elicited comprehension of non mental events. This comparison indicated that while activation in the temporal poles and the TPJ supply different background material necessary for successes in TOM reasoning[2], the MPFC is specifically responsible for the ability to anticipate and give prospective inferences regarding behavior (Amodio & U. Frith, 2006; Calarge, Andreasen & O'Leary, 2003; Frith & Frith, 2003, 2006).

Another way to elicit activation in brain areas that are needed for mentalizing is to engage participants in games that require attention and guessing the opponent's moves. Two studies (Gallagher, Jack, Roepstorff, & U. Frith, 2002; McCabe et al., 2001) used a game paradigm and compared brain activation in two conditions: during trials in which participants think they are playing against a computer versus others in which they think that there opponent is human. Both studies showed activity in the MPFC only when the participants believed that they were interacting with another person but not when playing against the computer.

The same brain locations were also activated during a TOM task presented as visual stimuli (Brunet et al., 2000; Gallagher, Happé, Brunswick, Fletcher, Frith, & Frith, 2000). Gallagher et al. (2000) presented TOM and non TOM stories and cartoons in an f-MRI study. They found activation in the same three brain locations discussed above, but the location that was activated only during TOM stories and cartoons was the MPFC. Another study that used cartoon stimuli was conducted by Brunet et al. (2000). They employed a pet scan while participants completed a three picture cartoon strip by choosing one final picture (out of three presented). Choosing the target

[2] The temporal poles, were activated during comprehension of different kinds of coherent and familiar stimuli. Comparing these findings with the finding that damage to the temporal lobes cause an inability to use information of advanced scripts, Frith & Frith (2003) concluded that these areas bilaterally generate the semantic and emotional context for understanding mental events. This material is crucial in understanding mental events and it is a perquisite to making predictions about future events or reactions but it is not unique to understanding mental stories and events. The importance of the TPJ is debated, see below.

picture meant the participant attributed an intention to the character in the comic strip. For example, they presented a strip of three pictures showing a man disassemble the bars from his jail cell window. The picture options for completing the story displayed the following scenarios (1) the man going to sleep in his jail bed (2) the man yawning near his bed and (3) the target – the man sitting on the bed tying the bed sheet into a rope. Selection of the target picture was interpreted as evidence that the participant created an implication regarding the characters intentions to escape jail. The brain areas that were active only when participants attributed these intentions were again the temporal poles, bilaterally, and the MPFC. The finding that non verbal material elicits activation in the same brain locations strengthens the claim that TOM ability is not language dependent. The same locations are activated in response to mental stimuli regardless of the mode they are presented by.

However, another line of research which designs tasks that differentiate various aspects of TOM presents different findings (Aichorn, Perner, Weiss, Kronbichler, Staffen, & Ladurner, 2008; Saxe & Powell, 2006; Saxe & Wexler, 2005). In these studies healthy adults read short passages, describing different attributes that are somewhat socially relevant like physical appearance of a character, some kind of physical sensation (e.g., being hungry) or background information (e.g., place of birth) and short passages about some content of thought (e.g., thinking about a flight being late). The researchers conducted an f-MRI scan of the participants' brains while they read the different passages. Although the same areas related to TOM were activated, only the activity in the R-TPG was increased measurably when participants read about content of thought. In contrast to the studies reviewed above the researchers conclude that the R-TPG is the location specifically dedicated to attributing thoughts (Aichhorn et al,, 2008; Saxe & Powell, 2006; Saxe & Wexler, 2005).

The debate about the specific contribution of each part of the neural system to TOM highlights the need for an assessment of the performance profile of RBD patients. Although the information gathered about the specific location of the brain damage can be informative about the neural circuitry that enables TOM, there is not enough information yet about the brain mechanisms which can allow us to deduce the TOM abilities of a specific participant from knowing his specific brain damage. Therefore it

is essential to collect such an assessment regardless of the specific location or etiology of the damage.

Performance of RBD patients on Theory of Mind tasks

Most of the research about brain locations of TOM was conducted on healthy individuals who were asked to perform different tasks while their brain activity was monitored.

A complementary line of research focuses on the performance of RBD on various tasks. Usually their performance is compared to healthy participants in order to understand in what ways the RBD's performance deviates from normal performance.

As explained earlier, some of the tasks devised in order to test TOM ability ask participants to judge certain situations. In these cases a one word decision or response informs the researchers about the mental processes the participant was employing. For example, in the false belief task there is a need to decide where the character will search for the item he wants. In other cases the participants are asked for a more lengthy verbal answer that demonstrates the mental concepts the participant is considering in his mental process. For example, in the 'strange stories' tasks, where participants are asked for a justification in addition to a yes/no answer. Each kind of task has its advantages and drawbacks. The "short answer" tasks are considered to be more genuine because the response is not dependent on linguistic abilities (Tompkins, Scharp, Fassbinder, & Meigh, 2008). On the other hand if there is no reason to expect a linguistic impairment, but a difficulty to explicitly reason about mental states or understand conflicting views appears, then it is important to reveal that difficulty and understand its sources (Happé, 1994; 1995). In order to gain a comprehensive understanding of the TOM ability a certain participant has, it is important to rely on both kinds of tasks.

The picture that emerges from the different studies reviewed here is a complex one. Surian and Siegal (2001) and Siegal, Carrington and Radel (1996) showed that RBD patients have difficulty in first order false belief tasks, if they were not aided by props or if the test question is not presented in a way that makes the point of the task clear.

Siegal et al. (1996) explained their finding that only 3 out of 11 RBD participants answered correctly to the misplacing false belief task, as a consequence of a pragmatic difficulty, not a conceptual one. They showed that when the experimenters explicitly pronounced that they are asking "where will the character look **first** for the target", and in doing so dismiss the inference that they are asking "where would it look in order to find it", more participants understood they are asked about intention and not about destinations. In this condition 5 out of 6 RBD patients answered correctly.

As mentioned in the section (1.1.1.) different studies devised illustrative tasks that are grounded in natural social interaction, for example, choosing the correct reference term for a third person (Brownell et al., 1997) or producing adequate requests in light of changing social situations (Brownell & Stringfellow, 1999). In both studies participants who suffered CVA to their right hemisphere performed significantly different than control participants. Brownell et al. (1997) presented two characters talking about a third party in various social situations. They asked RBD participants to choose which reference term a character speaking to another person would use to refer to a third party, either formal (Mrs./Mr.) or informal (a private name). The elements manipulated were whether the person choosing the reference term was well acquainted with the third party and whether he knew if his interlocutor was also acquainted with that third party. Only when the speaker was well acquainted with the third party and thought his conversation partner was also very familiar with him, was the choice of a first name reference appropriate. Results showed that the RBD were less sensitive to their hearers' acquaintance with the third party than controls. This was interpreted as an inability to consider the mutual information between the two interlocutors.

Another study which tested these patients' ability to understand social situations presented complex social situations and asked for a short judgment that provides information about the participants' appreciation of the situation. Winner, Brownell, Happé and Blum (1998) asked patients to judge utterances as either lies or ironic statements. Their decision should have been based on the information one character had about his interlocutor knowledge in a certain situation. For example, in one story, a boss saw his employee who called in sick that day at a hockey game. When the boss met him the next day he asked the employee if he had rested properly the day before.

The employee responded that he did rest all day. In half of the stories the employee was aware that the boss saw him at the game. In these cases the employee's last utterance was supposed to be judged as an ironic statement. In the second condition the employee was not aware that the boss saw him. In these cases the participants were expected to judge his utterance as a lie. In order to decide if the statement was a lie or an ironic comment the participants had to consider what the protagonist knew about the boss' knowledge of the situation. They had to consider second order knowledge, what one character knew about the knowledge of another character. RBD patients erred more than control participants, both in attributing lies and in understanding statements as ironic according to context. But it is important to note that not all RBD scored low and a few normal participants also performed poorly in this task (Winner et al., 1998).

Conflicting results appeared also in a different study which showed that RBD patients are not completely lacking the ability to understand and follow conversational intents. Cheang and Pell (2006) replicated a study first conducted by Brownell et al. (1983 in Cheang & Pell, 2006) that asked participants to choose an ending to a story. Each story was presented with four optional endings and the participants were asked to choose an ending that would create a joke and later to choose another ending that would make the story coherent but not funny. The RBD group showed large variability in the joke completion task. Some of the participants completed the task as well as controls (in a few cases even better) while others scored 2 standard deviations below the controls' mean score (Cheang & Pell, 2006). These results point to the importance of individual assessment of TOM abilities for people who suffered RBD.

As mentioned before, other tasks that were found to be difficult for RBD patients were understanding *faux pas* situations. Shamay-Tsoory et al. (2005) found correlation between the ability to understand *faux pas* situations and the ability to understand sarcastic remarks. Their interest was in the relation between social understanding and sarcasm. *Faux pas* served as a test of social TOM ability.

Another measure for the effect of TOM abilities on social communication is the way RBD patients understand and produce requests. Brownell and Stringfellow (1999) showed that these patients had difficulty in varying the form of their request for help (e.g., longer, shorter, more explanatory etc.) according to different social variables

like age or status of the person they are approaching. But, they did some adjustments to the changing social circumstances by adding in certain places the word "please". So, although the reactions which were consistently deviant from those of normals indicatated deficient TOM, this study also demonstrated that at least some of the RBD participants are in some respect aware of intentions and try to produce appropriate communicative messages, although not always successfully.

Other studies asked participants to express their understanding of mental causation in opposed to physical causation by asking them to explain an outcome of an interaction. Happé et al. (1999) presented RBD patients with two kinds of short stories and cartoons. In the test trials, assessing the mental state of the characters was crucial for understanding the gist of the story or the funny meaning in the cartoon. In the control trials, the understanding was based on some physical element in the story or cartoon. For example, a burglar that was running down the street was stopped by a policeman, but the policeman didn't know that the man had robbed a store, he stopped him in order to return a glove that fell from his bag. Once the policeman stopped the burglar he gave himself up and returned the jewelry he just stole. The participants were asked *why he did that*. A full explicit answer received a score of 2 and partial or implicit answers, a score of 1. RBD participants scored significantly worse than healthy controls and their reading time of the passages was longer. The same significant difference was found when these participants were asked to explain what was funny in a cartoon, based on some information gap between the characters. In the control stories and cartoons which described events that were caused by a physical element (for example, a burglar being caught due to an alarm ticking off) no difference was found between the two groups (Happé et al., 1999).

Recently Tompkins et al. (2008) suggested that there is a need to reevaluate the stimuli which were used in this study (Fletcher et al., 1995; Joliffe & Baron-Cohen, 1999; Russell, Schmidt, Doherty, Young, & Tchanturia, 2009). Their main criticism was that the control stories were very different from the TOM stories on a few important measures: the explicitness of the contradiction presented, the number of characters, shifts of viewpoints between the characters, syntactic complexity and cohesion.

The task Tompkins et al. (2008) used in their study was very different from the original Happé et al. (1999) task. The participants were asked to verify sentence probes. The appropriate probe for the TOM stories were derived from the correct answer to the test question (e.g., "why did the burglar admit?") used in a multiple choice version of the task that was developed for adults with aphasia (Tompkins et al., 2008). The accuracy and reaction time to the probes presented was measured. Comparison sentences that were unrelated to the preceding texts and filler probes were also presented (and were expected to be judged false). Results showed that RBD patients responded more accurately to TOM stories than control stories (but this difference was attributed to one specific item). The control group responded with equal precision to both types of stories reaching a ceiling effect. Response times showed a significant difference between groups (the non brain damaged answered more quickly than RBD). In addition, the answers to the TOM probes were faster than to the control stimuli in both groups. The researcher concluded that these results undermined the basic argument that RBD leads to deficiency in TOM. The most prominent difference between the studies, apart from the precise equation of stimuli, is the task presented. The task of justification elicitation in the Happé et al. (1999) study is very different from verifying the answer in the Tompkins et al. (2008) study. Referring to this difference the researchers say:

> *"These differences in assessment methods cannot account for the discrepant findings on RHD adults' relative ease of mental versus non-mental causal inferencing, because any potential advantage conferred by the more implicit method would have affected performance on both text types. However, the more implicit measure in this study was expected to improve RHD group accuracy overall.. Indeed, there was no accuracy performance gap between groups." (pp.55 -56).*

Although the results that show that RBD patients are sensitive to TOM stories, as much as they are to non-TOM stories or even more, are impressive the difference between the methods of research is still unresolved by this comment. In addition to the ability to comprehend mental situations, TOM also includes the ability to initiate reasoning and spontaneously attribute mental states to others, not only to identify them. Therefore an on-line measure of response time to TOM inferences cannot replace the evaluations of responses to open-end questions. In concluding their study

Tompkins et al. (2008) highlight the importance of individual differences found both in the Happé et al. (1999) study and in their own. These were not correlated to any demographic measure. They comment that a direct correlation between mental inference difficulty and lesion in the right hemisphere is not warranted. Future studies should consider not only site but the extent of brain damage as well as varied assessment of TOM.

This same concern was put forward earlier by Myers (2001). Her interest was to improve and enhance the clinical communication between therapists and researchers treating the difficulties RBD patients deal with:

> *"Research and clinical efforts in the area of communication impairments associated with acquired right hemisphere damage (RHD) are hampered by the lack of a definition and a universally accepted label for these deficits—one that is not solely dependent on lesion location. One of the most vexing problems for research on the nature (as opposed to the incidence) of RHD communication impairments is the fact that subjects are included in experimental groups based on site of lesion, rather than on the presence of the deficits under study. The situation is akin to conducting a study on the nature of aphasia in which subjects with left hemisphere damage (LHD) comprise the experimental group, regardless of whether or not they actually have aphasia." (Myers, 2001 p. 913)*

The present research was designed to meet these considerations. Because the aim of this study is to understand the linguistic consequences of damage to TOM, we cannot rely on the location of the brain damage found as a sole criterion and assume a TOM deficit. There is a need to first assess the participants' TOM ability in order to later correlate it with their linguistic performances.

The TOM assessment presented in the current research was designed to evaluate advanced TOM performance; the ability to understand, judge and reason about mental situations. We presented a battery of situations which are based on some mental occurrence and asked participants to judge and explain different aspects of them. This method, both versatile and linguistically demanding ensures that an accurate assessment of TOM is carried out.

1.2. Relevant Information and Appropriate Reference

One of the tasks speakers and hearers need to accomplish very frequently during conversation is the task of producing and comprehending referential terms. Three theoretical issues are important for explaining the way this task is achieved. The first is Relevance Theory (Sperber & Wilson, 1995) which offers a theoretical framework that defines the term 'relevance' and suggests a way this notion guides discourse production and comprehension. The second is the concept of Definiteness, especially the way speakers use the definite or indefinite articles to appropriately introduce referents into the discourse. The third is Accessibility theory (Ariel, 1990; 2001) which accounts for the way speakers use the variety of referential terms (including definiteness) their language offers. Each of these topics will be presented in the following discussion.

1.2.1. Relevance and Relevance theory

The ability to participate in a meaningful way in discourse depends on our ability to contribute information which is relevant to the situation and to our interlocutors. Grice (1989) described the human ability to participate in discourse as an expression of the ability to 'mind read', to understand the communicative intents of speakers. Understanding a communicative intent, he claimed, depends on the ability to understand two different kinds of information: the explicit meaning, encoded in the linguistic expressions, and deciphered according to lexical-semantic code and the implicit meaning. The implicit meaning is derived by the listener based on principles that guide discourse behavior. Grice assumed a conversation to be an act between partners who agree to participate in a coherent message exchange, this he called the 'cooperative principle'. This principle is realized by reference to four basic categories (which Grice named maxims): Quantity: your contribution to the discourse should be as informative as required, and no more nor less than that. Quality: don't say what you believe is false or you lack adequate evidence for. Relation: be relevant, and the fourth category: Manner: which states 'present the message in a brief, orderly and unambiguous fashion'. A crucial observation highlighted by Grice is that in certain circumstances the speaker breaches one of these guiding rules. In doing so she

generates an assumption that bridges between what was actually said to what she meant. This inference is the 'implicature', what the speaker is implying. An example from Grice (1989):

> *"Suppose A and B are talking about a mutual friend, C, who is now working in a bank. A asks B how C is getting on in his job, and B replies: Oh, quite well, I think; he likes his colleagues, and he hasn't been to prison yet. … A might reason as follows: (1)B has apparently violated the maxim "be relevant" and …I have no reason to suppose he is opting out from the operation of the Cooperative Principle. (2) I can regard his irrelevance as only apparent if, and only if, I suppose him to think that C is potentially dishonest. (3) B knows I am capable of working out step (2). So B implicates that C is potentially dishonest." (Grice, 1989, pp. 24-31).*

Grice (1989) proposed a divide between the explicit, semantic, verbal message to the implicated, pragmatic message. This allowed a convenient theoretical and practical division between the semantic representation, that results from processing the linguistic code and allows truth condition verification, to the pragmatic processes that follow the semantic representation and enrich it in some respect (for a critique of this position see Carston, 2002; Ariel, 2008).

Relevance theory (Sperber & Wilson, 1995) suggested different principles than those proposed by Grice for the way discourse understanding occurs. Grice was interested in describing how pragmatic processes aid and enrich semantic decoding. He considered the semantic processes to take place prior to and independent from the pragmatic processes. Relevance theory sees pragmatics through a different perspective. According to Sperber and Wilson, pragmatic ability is a *cognitive ability* and the study of this ability should be conducted within the framework of cognitive studies. According to their view, the process of understanding intentions is initiated by the listener. He comprehends the intended meaning of the speaker by obeying one principle the *'relevance principle'*, which is defined in terms of cognitive effects and processing efficiency. According to this definition, *relevance* is a property of inputs (i.e., utterance) to the cognitive processes. When an utterance is processed in context it can trigger contextual inferences (e.g., modify or reorganize assumptions). The relevance principle states that the speaker should only put the addressee to the

minimum cognitive effort in order to gain a sufficient amount of contextual inferences. The equilibrium point between these two contradicting goals is the point of 'optimal relevance' (Wilson, 2005; Sperber & Wilson, 1995). On the basis of this definition Sperber and Wilson (1995) propose two general principles. The first is the *'cognitive principle'* which stated that human cognition tends to be geared to the maximization of relevance. The second is the *'communication principle'* which states that every act of inferential communication is performed under the presumption of the speaker, that her utterance is the most relevant utterance compatible with her abilities and preferences.

How does the hearer bridge the gap between the linguistically encoded form and the full intended interpretation of the utterance? According to the *'communication principle'* the hearer decodes the linguistic message *and* enriches with inferences until the resulting interpretation meets his expectations of relevance (Wilson, 1999; Wilson & Matsui, 1998). According to Relevance Theory these two processes are not necessarily successive as Grice proposed. The researchers reformulated the division between the verbal explicit meanings to the pragmatic implicit intent. Sperber and Wilson (1995) distinguished between pragmatic processes that are part of the initial propositional representation, e.g. those that complete the literal message to a full propositional statement, and those that go beyond it. The verbally expressed utterance together with the initial pragmatic processes create what she termed the 'explicature' and the inference processes that goes beyond the propositional statement, the 'implicature'. The two processes together allow the hearer to comprehend the most relevant message the speaker is communicating (Carston, 2002).

The importance of the second stage, of comprehending implicatures, is widely accepted (Grice, 1989; Sperber & Wilson, 1995), but the proposal that pragmatic processes are involved in the initial stages of utterance interpretation, before the truth conditions can be evaluated, is more controversial and therefore will be demonstrated bellow (according to Carston, 2002; Wilson, 2005):

1. She gave him a chance.
2. The garden is near the house.
3. Everyone left the party.

In the first example the hearer has to assign reference to the pronouns 'she' and 'him' and to disambiguate the meaning of the word 'chance'. In sentence (2) the hearer must determine the meaning of the relative expression near'. In sentence (3) the hearer must determine the interpretation of the quantifier 'everyone'. The answer to all these questions can come only from pragmatic inference procedures; there is no coded arbitrary meaning that can be deciphered. Because these inferences are logically prior to determining the truth value of the proposition, we must conclude that pragmatic processes are involved in the basic structuring of the explicit verbal utterance (cf. Ariel, 2008; Carston, 2002; Wilson & Sperber, 1993).

So according to this analysis the understanding of a communicated message depends on two kinds of processes. The deciphering of the linguistic code is not expected to create difficulty to patients who suffer loss to TOM, but the second process, of creating inferences that rely on considering the speakers' intentions might be disrupted in such cases. According to this theoretical framework, decrease in TOM might result in a basic disability to comprehend propositions and not "just" their related inferences.

The question we pose here is whether patients who show decrease in TOM will experience difficulty in the two roles of discourse. As speakers, will they tailor their utterances to be the most relevant to their hearers? And as hearers, will they comprehend the most relevant message?

In the next pages we review studies that tested these abilities in patients on the Autistic spectrum (ASD) and with right brain damaged (RBD).

1.2.2. In Autistic Spectrum Disorders (ASD)

Difficulty in communication is the key feature of ASD. Happé (1993) was first to suggest that the communicative deficit in autism may be restricted to inferential communication, with the ability to use code communication remaining intact. In her study she tested three groups of ASD participants (between the ages of 9-38), patients who passed first order false belief tasks, patients who passed both first and second order false belief tasks and a group that did not succeed in either. They were all tested on their ability to complete different kinds of literal (simile) and figurative language expressions (metaphor and irony). The comparisons showed direct relationships

between the severity of the aTOMia, and the ability of the patients to understand figurative expressions. Participants who failed both false belief tasks were able to complete similes (e.g., X was like..) but not metaphors (e.g., X was..). ASD participants of the same age who passed first order FB tasks only were able to complete and comprehend metaphors but performed significantly worse than those who passed second order false belief, on understanding ironic statements. The researchers considered ironic statements to be more dependent on comprehending speaker's intentions than metaphors. Therefore, they interpret their results as showing a correlation between the ability to draw inferences and the severity of the aTOMia.

Another study that tested ASD patient's ability to comprehend intentions used a different task. Ozonoff and Miller (1996) presented autistic teenagers and adults (between the ages of 16 to 58) short vignettes composed of two sentences. The first was a somewhat ambiguous sentence and the second clarified the significance of the first. For example, one of the first sentences was "Jane hurried into the dentist's office" and the second: "She saw her purse on the table in waiting room". The participants were asked two questions about the passages: a memory question (e.g., Jane's purse was in the waiting room?) and a question that tapped their ability to infer about the topic i.e., about protagonists reasons for action (e.g., She had forgotten her purse when she left the office). The results showed that autistic individuals had little difficulty remembering facts but failed to correctly infer the main theme of the vignettes.

A different study which also showed that ASD are impaired in their ability to comprehend inferences compared the performance of ASD and SLI children. The participants were shown short dialogues that violated Gricean maxims of communication (Surian, Baron-Cohen, & Van der Lely, 1996) including maxims of truthfulness, quantity, relevance, and politeness. In each short conversation one speaker asked a question and two other characters each provided a reply to the question. One of the replies constituted a maxim violation while the other was a conventional answer. For example, to the question *"What did you do at school?"* the participants heard the answers *"We had a bath"* which was a violation of the maxim "be relevant" and the conventional answer *"We did some writing"*. The participants were asked to identify the reply that sounded "funny or silly". Relative to IQ-matched groups of children with SLI and normal children, autistic children performed more

poorly overall in this task, while the SLI children performed no differently from the control group. When looking at the different maxim violations, differences were found: the ASD group performed at chance on detecting violations of the maxims of quality and relation and scored somewhat better on quantity and politeness. The authors explain that the reason might be a preserved ability to draw some relevant contextual inferences on these last two kinds of violations. They state:

> *"We could not control the degree with which different items failed to be relevant, but it is quite plausible that the Maxim of Quality and Maxim of Relation utterances failed to achieve relevance to a greater degree than the ones violating the Maxims of Quantity and Politeness. In the latter cases, the addressee can draw an interpretation at the cost of some extra effort which decreases the relevance of the utterance... In contrast, for the items violating the Maxims of Quality or Relation the addressee cannot draw any plausible interpretation". (Surian et al., 1996, p. 65)*

This pattern of results, that ASD participants are able to draw some inferences and comprehend certain discoursive intentions, but not all, has been discussed in a recent review of 20 studies conducted in this field (Loukusa & Molianen, 2009). The different studies reviewed included a range of pragmatic inferences with different participants who were diagnosed with either ASD or Asperger Syndrome (AS), children and adults. The reviewers concluded that most studies show that these patients can draw some contextually relevant inferences although their ability was less developed than normally developing children and adults, therefore they claim that individuals with ASD and AS suffer deficiencies in pragmatic comprehension and inference abilities, but they do not have complete inability to draw them.

Next we turn to review research of RBD abilities in inferring intentions and their sensitivity to relevant information.

1.2.3. In Right Brain Damage (RBD)

The issue of RBD difficulties in communication and their ability to comprehend intentions and produce relevant messages has been studied by various methods. Myers (1979, republished in 2005) in her pioneering study of adults RBD patients interviewed 20 patients and presented important characterizations of their communication abilities. Her first description captures the main points under investigation here:

..that communication problems, when they exist, tend to be manifest by irrelevant and often excessive information and by literal treatment of questions and events. (p.1147)

She noticed that these patients had difficulties in constructing meaningful relevant answers to open ended questions like the ones used in clinical interviews, for example: *Tell me what happened to you and why you're in the hospital?* In contrary, they were able to appropriately answer questions that were clearly structured and required specific answers (e.g., Where do you live?). The explanation she offered was that in the case of structured questions, the sentence meaning is unambiguous and requires no additional inferences regarding the intent of the asker to fully understand what is required. By contrast, well formed responses to open-ended questions require inferences regarding the intentions of the asker: which kind of information she seeks and how much of it can be considered appropriate (Myers, 1978, In Sabbagh, 1999).

Recent examinations of RBD conversational characteristics (Blake, 2006; 2007) arrived at similar results. In one of the studies (Blake, 2006) RBD patients were read aloud short stories sentence by sentence and were asked to verbalize their thoughts after each sentence. Their speech was found to be different in a number of ways from that of normally aging patients. The RBD productions were rated as more tangential and egocentric. Tangential statements were defined as irrelevant, off topic or incorrect, centered on isolated details and statements that were digressions from the main topic. A statement was considered egocentric if the participant integrated himself or herself into the story or included personal remarks. Although this task and its analysis were not designed to test the ability to provide relevant information, the characterization the researcher arrived at were important to the current issue. They too showed that RBD patients discourse was different from that of normal participants in

their tendency to supply information which was classified by listeners (raters) as irrelevant and digressive (Blake, 2006).

Other studies added more detailed information about the difficulties of RBD experience when trying to fulfill their role in a conversation. For example, the way they produce indirect requests. Brownell and Stringfellow (1999) asked RBD patients and a group of control subjects to formulate requests in situations that differed on a number of variables. These variables were characteristics of the discourse situation (e.g., the kind of inconvenience the asker might cause (high/low), were the two interlocutors of different ages (yes/no), did they have similar social status (yes/no), ext.). The authors tested whether RBD patients adapted their request to the different contextual conditions. For example, did they offer more information to justify their request if compliance to their request create more inconvenience to the person they approach? They found that the RBD patients preferred to use the addition of a word "please" in requests that were more demanding. They did not use other linguistic tools, like varying the length of the explanations as controls did. The authors conclude that these patients were somewhat sensitive to the changes in the situation, but their ability to use language to mediate these differences was limited. In this case we see that although these patients considered the relevant features of the context their ability to produce the message that was relevant to the situation was limited.

A different study focused on another specific ability, the ability to appreciate jokes (Cheang & Pell, 2006). The participants were presented short stories and were asked to choose between four possible endings one that would create a funny ending to the story and one that would create a coherent, but not funny ending. Their performance on this task was compared to their performance on a different task designed to test their ability to identify communicative intents. In this task they were presented short stories and asked two kinds of questions, about the content of the story and about the intents of the characters described in it. The results showed that the RBD patients had difficulty in choosing funny endings compared to choosing coherent endings, and that they had no difficulty in answering informative questions but had difficulty at perceiving intents. According to the theoretical framework of Relevance Theory, success on the tasks in which RBD patients received lower scores required considering different point of views. Finding the funny ending demanded they compare the information the listener to the joke had and contradict it in some way.

Understanding intents clearly requires that the participant treat the information their interlocutor holds. The finding that these tasks were most difficult for RBD patients demonstrated that their main disability is in the realm of TOM, the ability to consider the speaker's intentions and to offer in turn information that fits these intentions. A similar difficulty, to attend to the more implicit and least salient information was also detected in a study of story comprehension (Frestl, Walther, Guthke, & von Cramon, 2005). In this study patients who suffered different kinds of brain damage (Right Brain, Left Brain and Traumatic Brain Injury) were shown short stories and were asked four kinds of questions about them. Two were about the main idea, one was related to what was explicitly stated in the text, and one had to do with information that was implied about the main idea. They were also asked about some details that appeared in the story which were not significant to the main idea. One question was about a detail that was explicitly mentioned, and one about a detail that could only be understood by implication. The RBD group had no errors on the questions that related to the explicitly mentioned main idea but erred much more (10 % more) than the other groups on questions that related to an implicit feature of the main idea. The researchers were able to draw implications about specific profiles of the different groups they studied. The RBD group was found to be characterized best by a difficulty in treating the information that needs to be inferred from a story and a high ability in comprehending the explicit main ideas and details.

These observations suggest that RBD patients, like ASD patients do not suffer global impairments in making conversational contributions; instead, their deficits are limited to situations which require them to fit their contribution to intentions which are not spelled out explicitly (Sabbagh, 1999). As speakers, they show difficulty in producing a message that reflects their apprehension of the context (Brownell & Stringfellow, 1999).

The tasks we presented the RBD patients were designed to test whether they are able to comprehend and produce relevant communicative messages according to the context they were presented using the appropriate linguistic means. What are the appropriate means will be described in the following chapters that introduce the topic of definiteness and accessibility.

1.3. Definite and Indefinite in Discourse

The topic of definiteness is a vast issue in the study of linguistics and philosophy of language (see Abbott, 2005). The aspect we focus on in this study is the way speakers manipulate definiteness to introduce new referents or elements into their discourse. This ability is important to the current study because its proper function depends on the speaker's perception of her addressee's recognition of the referent (Ariel, 1990; 2001; Prince, 1981). The issue at stake here is whether the speaker correctly assesses her hearer's ability to identify a referent from the context they share (from general knowledge, from the speech situation or from the prior discourse).

For example, suppose two people are talking in a street corner while suddenly a very fast car goes by. It would be felicitous if one of them will say: 'Did you see *that car?*', or 'Did you see *that?'* in reference to the car. The reason a speaker can use a definite NP this way is that it is very salient in the context and the speaker can assume it is very accessible to the hearer. But if two people are standing in front of two cars (or more) and one says: 'I love *that car*', it will be infelicitous, because the speaker cannot expect the hearer to know which car she is referring (unless some kind of gesture is made, or further description is offered). So, the difference between felicitous and infelicitous use of definite article in introducing a new item into the discourse depends on whether the speaker accurately evaluates her hearer's ability to identify it correctly from the context (Ariel, 1990; 2001; Prince, 1981).

The crucial link of this aspect of definiteness to TOM is evident. The speaker has to assess the hearer's state of knowledge concerning the item or referent in question in order to choose the appropriate form for introducing it. One quote that demonstrated the close link between definiteness and TOM is from Paul Chirstophersen writings, a Danish philosopher who discussed this feature of definiteness, a long time before the term 'Theory of Mind' came into use:

> *.. the speaker must always be supposed to know which individual he is thinking of; the interesting thing is that the 'the'- form supposes that the hearer knows it too"* (Chirstophersen 1939, 28 In Abbott, 2005).

The distinction presented in this quote, stresses what we can learn from the appearance of a definite article in conversation, namely, that its appearance supposes

the hearer is acquainted with the noun that follows. But notice that this conclusion cannot be drawn directly about the knowledge of the hearer. Rather, the use of a definite article tells us what *the speaker* thinks is known or unknown to the hearer during conversation. In other words, the speaker evaluates what is the mutual knowledge or shared knowledge between her and her addressee in order to appropriately use the definite article (Clark and Marshall, 1981).

The question is what were the sources for creating mutual knowledge about different referents that allowed the appropriate use of definite articles? Hawkins discussed the major uses of the definite article according to Chirstophersen (1939) and Jespersen (1949) (Hawkins, 1978). He identified three main sources of shared knowledge. One was the *background knowledge* – two interlocutors who know each other share before the discourse even begins. This can be personal knowledge or communal knowledge. For example, if you know your colleague has moved into a new office (and she knows you know about it) you can ask: *'Do you like the office?'* If two interlocutors share cultural knowledge, for example, that in every city in Israel there is a central bus station, a person can felicitously ask *'Where is the central bus station?'* without asking first if there is such a station in the city.

A second source of shared knowledge is the speech situation. The definite can be used appropriately if the referent is visible and salient to the hearer (and usually both) like in the example brought above of the fast car going by. The definite can also be used felicitously in cases when the referent's existence can be easily inferred in the situation, for example a sign near a zoo entrance saying: *'No feeding the flamingos'*.

The third source of shared knowledge is the linguistic context itself and the referents introduced during the discourse. The *anaphoric use*, for example: *I bought a dishwasher but the machine doesn't work.* Presenting *'a dishwasher'* sets up 'a shared discourse set' which can later be identified as the referent of 'the machine'. The same holds if we repeat a noun that has been previously presented as an indefinite NP. For example: *'My class is planning a few celebrations for the end of the school year. We will be having a party and a field trip. I'm happy about the party but I'm not sure I'll join the trip*'.

Another kind of anaphoric use is associative anaphoric use. For example, *'A car just went by and the exhaust fumes made me sick'*. According to Hawkins (1978) the

generic relationship between cars and exhaust fumes allow the NP *a car* to be a trigger to the associate definite NP 'the fumes'.

In the study of discourse the shared information which we have discussed so far is complemented by the *new information* which is brought to the discourse. This information, too, should be known to the speaker to be new to the hearer for it to be introduced appropriately. Prince (1981; 1992) made the distinction between two kinds of *old and new information.* 'Hearer old or new'; between what the speaker assumes her addressee knows and doesn't know in general. Part of this knowledge is the information that is given in the immediate context, but part of it is information known by the speaker to be known to her hearer from her background knowledge of him. The other distinction is between 'Discourse old or new'. Information which has been introduced to the current discourse is old, given information, and *discourse new* is information that has not yet been mentioned in the current discourse. The important possibility this explanation introduces is that there can be information which is known to the hearer (i.e., hearer old) but is new in specific discourse (i.e., discourse new). Prince (1992) found, using corpus analysis, that the distinction between definite and indefinite articles agrees better with the division between 'hearer old or new' and not between 'discourse old or new'.

These distinctions regarding the way speakers introduce new entities into discourse is an important intersection point between grammar and pragmatics that is especially important for understanding the connections between TOM and language. When the speaker introduces a new entity into the discourse, she must combine her linguistic knowledge about the definite system of her language and awareness to the accessibility of the referent in the mind of her interlocutor (e.g., whether it is 'hearer old or new'). This task is done without conscious reflection; it is part of the ongoing monitoring of the discourse situation (Ariel, 1990; 2001). The aTOMic participants in our study can therefore demonstrate the importance of TOM to the task of introducing new entities. We expect that if this task relies on the need to monitor on line the salience of the different referents in the mind of the hearers these participants will have difficulty in performing it.

Next we review research about the development of this ability in young children and its relation to TOM acquisition.

1.3.1. Acquisition of definiteness

The question we will focus on is whether children show sensitivity to the distinction between discourse new and old information, by using definite and indefinite articles and appropriate descriptions.

Studies show that young English speakers tend to add definite articles when it is unwarranted, until the age of three in spontaneous speech and in experimental settings (Schaeffer & Hacohen, 2003; Schaeffer & Matthewson, 2005). The context in which this is found in natural discourse is when a child is producing a definite NP while his interlocutor has no knowledge of the item being referred to, for example:

Sarah: where's *the black tape*?

Mother : what black tape? (Brown 73, In Schaeffer & Hacohen, 2003).

The researchers explain that these cases occur because the child has not yet acquired the concept of *'non shared knowledge'* and therefore she is choosing the article according to her state of knowledge regarding the item and not her hearer's. To test this assumption Schaeffer and Hacohen (2003) analyzed spontaneous speech of preschool SLI children. These children suffer different grammatical impairments but they are not susceptible to TOM or pragmatic difficulties. The researchers assumed that the SLI children will err only if the ability tested relies on grammatical knowledge (and not on TOM and pragmatics). They searched a corpus of spontaneous speech for two kinds of infelicitous use of the definite/indefinite articles. One was of the tendency to drop articles and use bare nouns instead (a grammatical error made by young children) and the second, to overgeneralize the use of definite articles. Results showed that SLI children tended to drop the article in 13% of their utterances, much more than normal children at that age (1%) and more than 2 year olds (9%). But they *do not* over generate the indefinite article (0%) just as normal children their age (0%), and less than the 2 years old (16%) (Schaeffer & Hacohen, 2003).

Karmiloff-Smith (1979) conducted a series of experiments to describe the acquisition of definite and indefinite articles in French speaking children between the ages of 3-10. One of her interesting studies presented a situation where an appropriate identification of the item referred to needed in addition to the definite article a description that would allow the items identification. The researcher showed children

between 3:3 – 11:7 year old, groups of four objects and then hid one of them. In some of the foursomes there were different items (e.g., doll, book, brush, cup), in others two of the items were identical (e.g., watch, pencil and two identical blue toy cars) and in others two of the items were similar (e.g., cow, plate, one blue fish and one yellow fish). The items were first shown to the child and then he was asked to close his eyes while the experimenter hid one of the items. Then the child was asked by the experimenter: *'What did I do?'* or *'What did I hide?'*. The researcher expected that if one of the two identical objects was hidden then the answer would be an indefinite NP (e.g., a blue car) but if one of two similar objects were hidden than the answer will be a definite and a described NP (e.g., the blue fish). In this last case the description was used as a determiner and it was essential to the proper identification of the item. The results showed that the ability to properly use the indefinite to refer to one of two identical objects, as expected, grew from 50 – 85% correct from 3 – 11 years (apart from 5 year olds who only used the indefinite in 40% of the cases). Showing that even the youngest children chose the appropriate article significantly more than in the other cases where the indefinite was a possible but not an essential choice. In using the descriptor as determiner results showed that three year olds do use it, in 29% of their answers, but children over 5 use it above 50% of the cases, and 11 year olds in 89%. Regarding these abilities the author says:

> *"The overall results of this experiment show that at no age do children have difficulty in using the indefinite article in its nominative function.. from 6 years particularly the grammatical addition of modifiers was rather consistent where relevant…Whilst percentages were fairly low it should be mentioned that both 3 and 5 year olds nonetheless make a distinction between items where modifiers were relevant and those where they gave redundant information. Those small children, who did not use modifiers were relevant, could nonetheless always recall the colors of the hidden objects if required to do so. This again confirms that the small child does not consistently understand the determiner function of modifiers, but rather their descriptor function." (pg. 85)*

This conclusion points to a dissociation between the use of the definite (in this case the determiner function of modifiers) and indefinite as essential grammatical particles and their use in discourse as markers of the (non)giveness of different items the speaker ascribes to her hearer.

Another revealing study tested children's sensitivity to mutual knowledge and its affect on the way they use definiteness, asked children (between 5-11 years old) to elicit *narratives* using a picture book (Kail & Hickmann, 1992). The same story book was used in two situations. Each child told the story in one of the two situations. The participant either looked at the picture story while the experimenter was seated next to her, and they were looking at the pictures together. Or, in the second situation, another adult was involved, the child was asked by the experimenter to tell the story to that second collaborator after the child blindfolded him. After the child finished telling the story, the collaborator recited the story back and the child helped him when necessary. The researchers expected that if the children were sensitive to the difference between the situations they would vary their use of definite or indefinite. They would use more definite articles when introducing items there was mutual knowledge of, when the child and the experimenter saw the pictures together. Results showed that there was a strong effect of the situation in all ages. In all ages the first mentions were more frequently done by indefinite than definite but, in the *mutual knowledge* situation children used more definite first mentions than in the situation of no mutual knowledge and more indefinites first mentions in the *no mutual knowledge* situation. There were also considerable differences between the age groups. Six year olds used both definite and indefinite to introduce new entities in the situation of *no mutual knowledge* in the same frequency. When they used indefinite determiners they frequently used them as deictic labels – this form did not occur in other ages. Nine year olds differentiated best between situations, using more definite when there was mutual knowledge and indefinites when there wasn't. Eleven year olds relied most heavily on discourse conventions. They were less affected by the change of situations. They used more indefinites to introduce new entities even if they shared mutual knowledge with their addressee. These results suggest that the pragmatic principles guiding the use of definiteness in narrative continue to develop after the age children first acquire false belief understanding and begin to consider the state of mind of their interlocutors (Kail & Hickmann, 1992).

In Hebrew definiteness marking is done by the marker '*ha*' as a nominal prefix that does not inflect. It is used in agreement both on nouns and the adjectives modifying them. There is no special article or prefix to designate an indefinite NP. The ability of Hebrew speaking children, between the ages of 2-5, to produce definite articles in

appropriate contexts was tested in a toy elicitation task (Avram & Armon-Lotem, 2005). The toys, performing some action, were presented in different ways to create a comparison between situations of *mutual* or *no mutual knowledge*. Results showed that the children made definite omission errors until the age of 4, when their English contemporaries already used the definite appropriately. Five year olds were found not to omit or over generate definite articles in Hebrew, like adults. The researchers explained the difference between languages as resulting from a less regular definite system in Hebrew (Avram & Armon-Lotem, 2005).

In summary, the syntactic realization of definiteness is different in different languages, but for the most part, the ability to use them properly in discourse was found to be related to the ability to attend to the availability of mutual knowledge. The understanding of situations of mutual knowledge or the lack thereof naturally depends on TOM. The speaker must take into consideration what is known to her addressee in the context of their conversation and what is not. The application of this awareness to linguistic differentiations between the appropriateness of introducing an entity with definite or indefinite marker is evident in discourse already at the age of 4 (Schaeffer & Hacohen, 2003; Schaeffer & Matthewson, 2005) and is regularly applied in an adult like manner to narrative towards the end of elementary school years (Kail & Hickmann, 1992).

1.3.2. Lack of TOM and the use of definite/indefinite articles

Research about the relations between lack of TOM in the Autistic spectrum disorder (ASD) or Asperger Syndrome (AS) population and the use of definite and indefinite articles in discourse has not been researched directly as far as we know. But different researchers have treated these patients' ability to create cohesive links during conversation. One of the measures of coherence was their ability to introduce items appropriately into the discourse. For example Fine, Bartolucci, Szatmari and Ginsberg (1994) analyzed the discourse of individuals with autism or Asperger syndrome. The results showed that the autistic participants did use anaphoric expressions in discourse no less than the other groups (Asperger and control) but they tended to refer more often to the speech situation than to previous text in the conversation. The Asperger syndrome group also referred to prior conversation but made errors in their use of

cohesive ties, primarily making unclear references that were difficult to interpret. This showed an irregular pattern of relating to shared knowledge. Another research that examined different kinds of coherent ties also showed a high rate of inappropriate cohesive ties made by ASD children (Baltaxe & D'Angiola, 1992). The examples given for incomplete and erroneous ties showed an inappropriate use of articles when presenting items into the discourse but the researchers did not present this linguistic inability as relating to lack of TOM or pragmatic difficulties. For example they gave the utterance (3) below the title of 'errors of reference', and the utterance (4) the title 'incomplete tie', but it was not clear what differentiated between the two who seemed to be two examples of inappropriate introduction of referents:

(3) Who gave you this book? *He*

(4) The boy delivered *it* to class.

The researchers stated that neither of these pronouns was recoverable from context (Baltaxe & D'Angiola, 1992).

The group of RBD that were found to be aTOMic in the current study were tested in such a specific aspect, their ability to appropriately use the definite article ('ha') and indefinite in different contexts and in their ability to use descriptors as determiners. Success in these tasks showed that the participant can appropriately identify an item from the context it appeared in, and could properly use different markers to assure her hearer will be able to identify the intended referent.

1.4. Accessibility and reference

As introduced in the preceding chapters in order to be able to contribute meaningful information regarding referents in discourse the interlocutors should consider the information they have in common and the information which is new (to the hearer and/ or to the discourse). But the differentiation between what is given and what is new can also be considered as a continuum between complete familiarity to complete unawareness.

According to Ariel (1990; 2001) every natural language offers means to mark the accessibility of each referent mentioned in discourse. The speaker marks the different referents mentioned along the discourse in accordance to their accessibility in the mind of her hearer (to her best estimate). When a referent is not highly accessible to the addressee he needs a relatively large amount of information in order to retrieve it from his mental representation. In this case the speaker will choose a *full name* and sometimes add a relevant description to insure correct identification. These are low accessibility markers. When the referent is highly accessible, if it the topic of the discourse or has been mentioned very recently, the speaker will mark it with a high accessibility marker, like pronouns or gaps.

In order to choose the appropriate marker, in each phase of the discourse, the speaker has to monitor her hearers' mental model of the discourse and attend to the shifts of attention to the different referents mentioned. She does this without explicit consciousness to this task, during on-going discourse. This monitoring of the addressees' knowledge is expected to decrease when lack of TOM ability takes place.

Ariel's definition of the term accessibility is grounded in the theoretical framework of Relevance Theory (Sperber & Wilson, 1995). As presented in chapter 2.1 Relevance theory assumes that the accessibility of different contextual components changes during discourse processing. In accordance to the *relevance principle,* the hearer will always try to invest the minimal amount of cognitive effort in order to receive the maximal amount of contextual inferences that bridge between new information and prior discourse postulations. The point of 'optimal relevance' is the point where the speaker balances between the two goals. This is also the point the speaker is aiming at

when choosing the reference term. She is trying to choose the term that will demand the least effort to process while achieving the highest point of accuracy.

The accessibility of the referents during discourse is dependent on a number of factors:

a. The distance between the referent's prior mention during the discourse. From analyzing different kinds of texts Ariel (1990) showed that speakers tend to mark the referent with lower accessibility markers as this distance grows.
   When a referent is mentioned the amount of information needed to correctly choose the intended referent is the largest. In successive mentions, the speaker takes into account the distance from the earlier mention.
b. Whether the referent is the discourse topic or not. The more salient the referent is in the discourse, the more it will be referred to by a low accessibility marker. Research showed (Broadbent, 1973 In Ariel, 1990) that when either of two entities can plausibly be referred to by the pronoun 'it', the subjects tend to relate it to the referent in subject position, which is the more salient position. It was also found that the number of times an entity is referred to helps the addressee determine what is the discourse topic (Levy, 1982, in Ariel, 1990). It was also demonstrated that different populations (adults, preschool children and aphasic patients) use this preference to match a pronoun referential term with an antecedent in the subject position ('subject rule') as a default rule in identifying the antecedent. The research showed that if other rules that contradict the 'subject rule' are presented (e.g., parallelism) subjects who err, err more in preferring the antecedent in subject position than in object position. In other words, if the subject fails to apply the contradicting rule his choice of antecedent will not be random, rather, he will prefer the noun phrase in the subject position as the antecedent of the pronoun (Zuckerman, Vasic, Ruigendijk, & Avrutin, 2002).

c. The point of view of the speaker. When there is a change in the speaker's point of view to that of the antecedent, the probability for a low accessibility reference increases.

d. The level of competition between different entities in the discourse. The more referents there are the higher is the chance they will be referred to by low accessibility markers (Clancy, 1980 In Ariel, 1990).

The speaker keeps track of these criteria during discourse. She chooses the appropriate reference terms according to her (unconscious) evaluation of saliency of the discourse representation in the mind of her interlocutor. This term will allow the identification of the referent she points to with minimal cognitive effort (Ariel, 1990; Sperber & Wilson, 1995).

In the same way that the accessibility of a referent is a relative term, so is the accessibility rate each referential term denotes. For convenience, Ariel (1990) divided the reference terms to three groups: low, intermediate and high accessibility markers. Note that each group includes a verity of markers which do not indicate the same degree of accessibility, and cannot therefore replace one another in the discourse. It is certainly possible to grade the markers in each group, as Ariel (1990) did, and create a sequence of terms that includes the terms from all three groups.

The main point that accessibility theory makes is that such a scale can be traced in every natural language after a careful analysis of the way speakers use the variety of referential terms in that language.

1.4.1. Development of appropriate use of referring expressions

Developmental studies support the connection between the ability to attend to others perspectives and the ability to use referential terms. These studies show that the acquisition of personal pronouns is preceded by the ability to grasp the existence of two different (physical) viewpoints. In a longitudinal study that followed children, from the age of 1:6 until they had acquired all pronouns, the researchers found that before children acquired first and second person pronouns they could coordinate between two visual perspectives. This was demonstrated by asking the children to hide or show different objects to the experimenter which had a different visual perspective then they had. The research also showed that acquisition of third person pronoun was achieved in parallel to the ability to coordinate between three visual perspectives (Richard, Girouard, & Décarie, 1999).

As was discussed earlier (chapter 1.1 TOM), the ability to grasp different and converging visual perspectives is considered as a sign of emerging TOM. This takes place when children, approximately 9 months old, can engage in joint attention with

another person. The study by Richard et al.'s (1999) about pronoun acquisition showed that there was a connection between these cognitive abilities and pragmatic behavior, of coordinating a pronoun with the viewpoint.

The acquisition of pronouns is a first step in fulfilling the discourse task to coordinate between changing viewpoints and application of the proper personal pronoun. Matthews, Lieven, Theakston and Tomasello (2006) showed that 2:0 year old children are not affected by the perceptual availability of an event for their addressee when choosing a reference term. At that age they ignored the addressee's perspective when answering the open question "what happened?' At the age of 3-4 years children showed some sensitivity to their addresses' viewpoint and tended to include lexical nouns (and the verb) in their answer when the addressee could not see the event referred to. Only 4 years olds tended to give the appropriate pronoun response when their addressee could see the event, while 3 year old answered the same or gave a verb only response in such condition. The correspondence between using the higher accessibility markers and the situation of mutual knowledge of the situation seems to follow the use of lower accessibility markers when there is a need to present new information.

Prior discourse also affects the way children choose reference terms. Campbell, Brooks and Tomasello (2000) showed that when children (2;6-3;6) answered specific questions like "What did X do?" they preferred to use null reference or some pronouns and when asked generic questions like "What happened?" they tended to answer with pronouns and they also used nouns. This showed that they are sensitive to the prior mention of an entity in the discourse context. In a subsequent study (Matthews et al., 2006) that reduced the difference between the two context questions described above, the researchers presented the question "what happened?" either with a prior mention of the name of the actor (“Was that the Clown? Oh! What happened?”) or not (“That sounds like fun! What happened?”). This contrast allowed checking whether children were sensitive to the prior context or to the actual appearance of the full name in the preceding question. They found that 3 and 4 years old used fewer full nouns and answered with a pronoun-verb construction if the name of the character was mentioned before the question. When the character's name was not mentioned they replied with a noun—verb construct. Even 2 year olds responded differently to the different contexts. At that early age they were significantly more

likely to name the referent (in a noun-alone response) if the name of the character was not mentioned with a full noun before their question. This showed that the sensitivity to prior context was evident at a very early age, before full acquisition of the pronominals was achieved. It also signaled children's grasp of mutual knowledge and their employment of this knowledge to the choice of the appropriate accessibility marker.

In a corpus analysis of children's use of referring expressions Gundel, Ntelitheos, and Kowalsky (2007) demonstrated that 3 years old children were capable of using referring expressions in a way that suggested that they are sensitive to the attention state of their interlocutors. They used personal pronouns almost exclusively when the referent is clearly in focus and most definite article uses were for entities that were often not in focus. Older children's ability to use and choose appropriately different kinds of referring expressions was tested using a story elicitation task. Hickmann, Kail, and Roland (1995) asked French children 6, 9 and 11 year old to narrate a story using a picture book (*Frog where are you*, Mayer, 1969) in two situations. In the first situation the experimenter and the child had mutual knowledge of the pictures. They sat next to each other and the experimenter asked them to tell her the story. In the second condition they were asked to tell the story to a naive addressee who sat in front of them and was blindfolded, therefore had no access to the picture being described. They found different patterns of responses across the ages showing that the development of this ability takes place long after the initial acquisition of pronouns. At age 9 and 11 children used more referring expressions to denote the human protagonist in the situation of no mutual knowledge. No such difference was found at 6 year olds[3]. These results demonstrated that the early acquired cognitive ability to apprehend states of mutual attention and diverging attention was evident in language from an early age (around 2:0) but the full mastering of the referential system that the language allowed, and/or the ability to use it appropriately in discourse situations, matures much later (age 9-11).

[3] Older children were more affected by co- reference than by episode boundary, choosing more pronominals if an earlier mention of the human protagonist appeared earlier in the clause, disregarding episode and frame boundaries.

The next question refers to the connection between TOM and the ability to assign reference during discourse. Is TOM necessary for applying the guidelines detailed in Accessibility theory (Ariel, 1990, 2001) for choosing the appropriate referential term in discourse? An important source for evidence to answer this question is research of populations that suffer damage to TOM, namely individuals diagnosed with Autism or Asperger syndrome, and patients who suffer RBD. Both will be considered next.

1.4.2. In Autistic Spectrum Disorder (ASD)

An irregular use of pronouns in the ASD population was reported in studies that analyzed their spontaneous speech (Baltaxe, 1977; Tager-Flusberg, 1989). It was found that young autistic children made reversal errors (they referred to themselves as 'you' instead of 'me'). This tendency indicated a very basic misunderstanding of the effect the change of point of view had on the choice of reference terms. This spontaneous phenomenon was tested in a planned research that compared ASD children and adolescence to a group of mentally retarded individuals matched on chronological age and vocabulary knowledge (Lee, Hobson & Chiat, 1994). In order to assess production, they were asked who could see an object in a picture when that picture either faced the experimenter or themselves. Comprehension was assessed by asking them to identify themselves and the experimenter in a photograph. Both groups succeeded in both comprehension and production of appropriate pronoun (me/you) in both kinds of tasks. These results showed that ASD who accumulated the vocabulary knowledge of about 5 year olds (chronological age of 14:9) appropriately vary the referent they choose. It was also found that ASD participants sometimes inappropriately chose to use their own name or the experimenter name instead of using a pronoun. This showed that they disregarded the fact that the knowledge about the identity of the referents was *mutual knowledge* in the discourse situation and/or they did not apply this appropriately to choose the reference term (which should be a pronoun, because it was very accessible to both interlocutors). Contrary to the interpretation that was given for the observations of their spontaneous speech (Tager-Flusberg, 1989), this planned study demonstrated that ASD children have an ability to identify the true referent of pronouns. But, their ability to choose between name and pronoun differed from that of the control group.

The study conducted by Lee et al., (1994) was planned to test only two kinds of referring expressions that matched two different individuals in the world. Recently, a broader research was conducted to assess ASD ability to differentiate between different kinds of referring expressions that corresponded to one antecedent. Arnold, Benneto, and Diehl (2009) tested ASD children and adolescents on their ability to appropriately use pronouns and null NPs as reference terms. The participants were divided to two age groups, young (9;8-12;9) and older (13;1-17;8). A control group of typically developing adolescences (matched on IQ, age and gender) was assessed also in order to create a baseline. All the participants were shown a cartoon video and were asked to describe what was going on in it. The researchers found that all the participants used pronouns and zeros more often for entities that had been recently or prominently mentioned, what Ariel would call "high accessibility contexts". A difference from baseline was found only in the younger group of participants with Autism. They tended to use fewer pronouns and zeros and more nouns than the older participants with Autism and the control participants of both age groups. This effect was most pronounced in reference to things that had been mentioned, but not in the previous clause. The writers could not explain this difference as a consequence of a shorter memory span since the young and older group of ASD had the same memory span and the older group did not score differently than the control group. The explanation they offered is in line with the theoretical framework offered by Ariel (1990), that the younger ASD participants did not take into account the availability of the reference in the mind of their interlocutor and tended to give more information than needed. They add that this explanation is consistent with earlier research (i.e., Baltaxe, 1977)[4]. Baltaxe (1977) observed the discourse of ASD young adults (14-21) and found they produced full NP when pronoun was more appropriate as in (1) bellow and to repeat a pronoun when a zero was more appropriate, as in (2):

(1) *Examiner: Have you ever seen a lovely lady that you thought you might like to have as your wife?*

Participant: : No, I haven't seen <u>a lovely lady</u> like that but I am going to keep meeting lots of nice looking nice <u>lovely ladies</u> close to my age hopefully.

[4] However, it contrasts with Tager-Flusberg's (1995) data, in which children with autism failed to use full noun NPs when an item was first mentioned. Arnold et al. (2009) explained that this might be because they were using pronouns deictically.

(2) *Examiner: How do you account for the fact that the team has dropped down almost to the bottom?*
Participant: They're a lousy team this year. They'll fall from world champions. They have won the world championship. They have won thirteen world championships.

Arnold et al. (2009) results also fitted the pattern found by Lee et al. (1994) who showed ASDs preferred to produce a proper name rather than a pronoun although they were referring to themselves or to their interlocutor. The writers speculated about the reasons for the older ASD group success in the task. One possible explanation they raise for their results is that their participants were not presented with a mentalizing task so theoretically this specific group might not have a TOM deficit. Another possibility is that the task of choosing a reference term in accordance with the interlocutor's state of knowledge is in some way easier than tasks used to evaluate TOM, especially those designed to show the ability to represent the content of thought of another person. An important factor that affected the ASD performance is the cognitive load during production, participants in all groups produced fewer pronouns in utterances that were not fluent (indicating some level of production difficulty), or while under the load of planning a longer utterance. In sum the researchers stated:

> *"while our findings do not preclude the possibility that speakers sometimes choose expressions on the basis of the addressee's needs, they are also consistent with a growing body of evidence that production choices are often more sensitive to speaker- internal constraints" (Arnold et al., 2009, pg. 143)*

The fact that cognitive load lessens the use of pronouns does not stand in opposition to the claim that the major factor affecting the choice of an appropriate referent term is the consideration of its accessibility in the mind of the interlocutor. According to relevance theory (Sperber & Wilson, 1995) and accessibility theory (Ariel, 1990) the task is dependent on investing cognitive effort that will allow the correct identification of each antecedent.

1.4.3. In Right Brain Damage (RBD)

We do not know of many studies that tested RBD patients' ability to use reference terms, and none that have tested RBD patients ability to choose referential terms in accordance to the accessibility of the different items in the discourse, or in the minds of their interlocutors. There are a number of studies that used these patients' ability to produce referential expression as a measure of their ability to create cohesive texts in text elicitation tasks. For example, Davis et al. (1997) measured cohesion by computing the ratio between the number of *'coherent ties'* between refereeing expressions and the overall number of referring expressions. In other words, they grouped together under the title of *coherent ties* different types of referring expressions (personal pronouns, demonstrative pronouns and lexical reference signaled by '*the*' before a noun previously mentioned). This measure did not differentiate between referential terms that were used appropriately or not (for example, if pronouns were used in cases where it was not clear who they are intending to refer to) it also did not include important referential options (null reference). The participants were tested in three tasks, two were picture elicitation, one while the pictures were in front of them (and in view of the experimenter too). The second was for the same picture stories that were moved away. The third task was a task of re-telling different stories. Results showed that the RBD group was impaired in referential cohesion ratio, in the retelling task but not in a picture elicitation tasks compared to matched controls. It is important to note that the range of cohesion ratios of the RBD in both elicitation tasks was very high (between 0.18-0.91 in one and 0.4-0.97 in the second). These results demonstrated again the wide diversity of ability within this group of patients.

Another study tested RBD ability to narrate coherent stories showed participants a 9 minute video and asked them to tell a naïve listener the story they saw as completely as possible (Uryase, Duffy, & Liles, 1991). The analysis was conducted by comparing each narrative to a target story devised by the experimenters. The group of RBD was compared to (an aphasic group) and a group of non brain damaged. The cohesion of the texts were tested by identifying *cohesive markers* (e.g., personal reference, demonstrative reference etc.) and the adequacy of each marker was classified as complete, incomplete, incomplete tie, error or error tie. The criteria for this classification were not specified. The results showed the RBD's retellings were less

cohesive than the normal control group. They produced a smaller proportion of complete ties and a greater proportion of incomplete ties.

A different study (Marini et al., 2005) tested RBD (and left brain damaged who were non aphasic) and a group of healthy controls on three kinds of tasks: retelling a story, a picture story elicitation task when the pictures were presented in order and another elicitation where they first had to arrange the pictures by order and then to tell the story. The measure for cohesiveness the researchers used was the rate of *'cohesive errors'*, these were defined:

> *"A cohesive error was scored each time a cohesive function word was used in the wrong way, an ambiguous coreference was established or whenever number and/or gender agreement over utterance boundaries could not be detected. Furthermore, also the presence of unfinished utterances whose meaning was continued in the following utterance was considered as a cohesive error" (Marini et al., 2005, pg. 49)*

In contrast to the two studies mentioned above, in this study the RBD group was found to have a lower cohesion rate score in the picture elicitation tasks compared to healthy controls but not in a retelling of a story. We can conclude that the different tasks and the different measures used affected the overall picture of results. It seems that the RBD group had difficulty in producing coherent narratives compared to normals but there was a large heterogeneity within this population. It is also evident that the understanding of these difficulties will improve if we consider not only location of brain damage but also the ability to create and use mental reasoning.

Why should we test a-TOMic patients on the ability to use referential terms? The theoretical framework presented above, of the pragmatic ability, to match between a certain entity and a suitable reference term in accordance to its accessibility in the mind of the hearer, calls for attention to TOM. As discussed in the preceding chapter, patients after RBD lose some of their ability to attend to, and reason about, the mental state of their interlocutors. The question we raise here is whether this loss expresses itself in the task of choosing and interpreting referential terms. As speakers, a reduced ability to represent their hearers' mind should make it difficult for them to choose the appropriate reference term in accordance to its saliency in the mind of their hearer, in

each phase of the discourse. As hearers, their task is to decipher to which of the different entities mentioned in the discourse the speaker is referring to. Difficulty in matching reference terms to entities will reflect difficulty in understanding the speakers' intentions.

1.5. Mental State Verbs

Mental state verbs (MSVs) are the most prominent verbal tool for expressing mental and psychological processes. They are an important issue in this study because using them appropriately demands coordination between their semantic, syntactic, and pragmatic properties. Semantically, MSVs describe internal states and events like *thinking, knowing, forgetting,* and *guessing*, usually transitory states rather than long term attributes. Many of them are polysemies (e.g., *know*) and the meaning of different MSVs are tied to one another (Booth, Hall, Robison, & Kim, 1997). Syntactically, these verbs usually appear in complex sentences, they take sentences as complements (S-comp) more frequently than other verbs (for instance, more frequently than movement verbs), and therefore their acquisition is considered an important step in children's language acquisition (Nixon, 2005). Pragmatically, they take a verity of roles, as a device to draw attention (e.g., *know what*..), as filler words (e.g., *you know*) as hedges (e.g., *It's going to rain, I think*), etc. (Booth et al., 1997; Field, 1997). In many utterances, MSVs are used to express an attitude (of certainty or surprise) about the content of the embedded clause. For example in the utterance: *I didn't know you liked red*, the speaker is probably stating her stance in regard to the information in the complement, and not referring to a specific content of thought. Some linguists claim that this is by far their main function in discourse (Field, 1997; Thompson, 2002).

The use of MSVs in sentences and discourse offers interesting opportunities to study their linguistic characteristics and in turn, these characteristics can be used experimentally to inform us about the minds that put them to use. We focused on three characteristics of these verbs: (a) Factivity (P. Schultz, 2003) (b) the certainty they communicate, and (c) the time frame these verbs assume.

An inadequate comprehension or production of MSVs will show that these characteristics are not included in the mental representation of the verbs in the minds of the participants in the study. This would mean that important parts of the messages communicated by using these verbs are not getting across. In the case of patients whose TOM abilities are impaired, it will show a connection between the ability to use MSV and the ability to represent the intentions and meanings interlocutors are

communicating. Next, I will describe the three lexical characteristics and review experimental research concerning them.

1.5.1. Lexical-Semantic attributes of MSVs

Factivity

The quality of being 'factive' or 'non-factive' is usually ascribed to different MSV (following Kiparsky & Kiparsky, 1971; e.g., Field, 1997; Nixon, 2005; Spanoudis, Natsopoulos, & Panayiotou, 2007). According to these researchers, factive MSVs are predicates that presuppose the truthfulness of their complement. One of the tests they use to divide between factive and non-factive verbs is the consistency of their presupposition under negation. For example:

(1) She knew that [it was raining].

(2) She didn't know that [it was raining].

In both (1) and (2) the presupposition that *'it was raining'* holds. This presupposition can also be termed the shared background knowledge, which the speaker and hearer assume to be true (Smith & Wilson, 1979, In Eisele, Lust, & Aram, 1998). Another test is the consistency of the same presupposition in questions. For example:

(3) Did she know that it was raining?

The question in (3) is about the subject's knowledge, not about the fact, taken for granted, that it did rain.

In contrast, the truth conditions of sentences with non-factive predicates (in 4-5) do not involve the truthfulness of their complement; they only refer to the mental model of the speaker:

(4) She thought that it was raining.

(5) She didn't think that it was raining.

And there is no presupposition/assumption about the state of affairs in the world following from the question in (6)

(6) Did she think it was raining?

Initially, the quality of being factive or not was attributed to the verb alone. Kiparsky and Kiparsky (1971) listed verbs that either belonged to one of the two group. Attention was also given to the type of complement that appears in the predication but the characteristic was ascribed to the verb alone and not the whole predication (Eisele et al., 1998). Recently, P. Schulz (2003) revised the treatment of factivity and ascribed an equal responsibility to both predicate and complement in achieving a factive or non factive reading. She started out with demonstrating that the same predicate can give rise to a factive and a non factive reading:

(7) I forgot to buy bread.
(8) I forgot that I bought bread.

The first, (7), is nonfactive, but (8), with the same matrix verb, is factive. This example shows that the embedded clause takes an equally crucial part in achieving the factive reading. P. Schultz (2003) revises the definition of the predicates and considers them as *potentially factive*. She explains that the factive reading arises when two conditions are satisfied. The first is that the predicate has a *potential factive* reading. The second condition is that a certain relation between the time of the utterance and the time of the event described in the complement takes place. More precisely, there has to be *"at least one moment in time such that the topic time of the complement clause preceded or overlaps with the topic time of the matrix clause."* (P. Schultz, 2003, p. 20) In other words, a factive reading demands *the potentially factive* matrix predicate embed an S-comp that describes an event that occurred before the time the utterance was produced, as in (8). If the S-comp is in future tense a factive reading cannot be achieved, as shown in (9) (P. Schultz, 2003).

(9) John forgot that Mary will be in Berlin.

This case, P. Schultz (2003, p. 20) explains, is shorthand for something like *John forgot that Mary planned to be in Berlin*. But the predication as appears in (9) is not factive. If a non-finite construction is embedded, as in (7), the predication is read as non-factive. In these cases, the truth of the complement is *implied*, not presupposed. This is the major difference between the factive and non-factive predications. In

factive predications, the information in the complement is presupposed and its truth value is fixed. In non-factive predications, the truth value changes according to the matrix verb. For example:

(10) I remembered to take my coat.
(11) I forgot to take my coat.
(12) I didn't remember to take my coat.

In (10) the complement *I took my coat* is implied to be true. When the matrix verb is negated, lexically (as in 11) or syntactically (as in 12) the complement is false (for further discussion about nonfinite complement clauses see P. Schultz, 2003, pp. 21-29).

The ability to differentiate between factive and non-factive predications can serve as an indication for sensitivity to the information status of different complements, what information speakers commit to, and what information they only imply. This sensitivity is an intersection between the lexical characteristics of the predication and its communicated meaning. This intersection can prove to be very telling because it can inform about the stance of the lexical and mental features of these verbs. If only the lexical characteristics of these verbs are taken into account, then we expect that patients with TOM deficit will not show any difficulty in using them appropriately. But if computation about the mental state of the speaker in choosing the different verbs is taken into account we expect that this group of patients will show difficulty in differentiating between them.

Certainty

MSVs also signal the certainty the speaker ascribes to the information in the complement. While factivity is a dichotomous notion, a scalar differentiation can be made along a *certainty continuum*. For example, within the potentially-factive predicates, *know* signals a higher level of certainty than *believe* and within the non-factive predicates *sure* express a higher certainty than *think* (Nixon, 2005). Linguists interested in understanding the fine connections between speaker stance and the way it is communicated in discourse show, using of corpus natural discourse, that the most frequent use of MSVs is to signal the stance of the speaker towards the message

communicated (Field, 1997; Thompson, 2002). An example for the details speakers can communicate using different (potentially) factive verbs is given by Field (1997). She demonstrated that *affective* and *epistemic* factive predicates can give rise to different understandings of the stance of the speaker towards the message she is communicating, in addition to the stance the grammatical subjects hold towards the information in the complement. For example, in the sentence: *He knew that what she had told him was a lie*, the use of the epistemic MSV *'know'* communicates a high degree of confidence about the reality of the information in the complement (*that what she had told him was a lie*), on the part of the grammatical subject. Additional information is understood from the use of an affective MSV. From the sentence *He was amazed that what she told him was a lie*, the verb *amazed*, communicates not only that the grammatical subject (*he*) came to know about the lie, it also tells that the speaker uttering the sentence thinks the agent in the sentence was surprised by this realization (Field, 1997). This is an example of the effect MSVs have on the message communicated. We will be interested in testing whether a decrease in TOM affects the ability to communicate and understand the range of certainty these verbs imply. We expect that patients with TOM deficit will show decreased ability if, in considering the certainty these verbs communicate, there is a need to access the mental stance of the speaker towards the situation described. If only semantic information is taken into account, these patients should perform as well as their peers.

Time frame

Some MSVs semantically define the time of their complement's occurrence. For example, *hope*, *promise*, and *plan*, cause an expectation for an event that will happen in the future, whereas other verbs, like *regret*, *find out*, *be sorry*, indicate an event that already occurred. An inappropriately tensed complement creates an infelicitous sentence (although a rich context can make it appropriate in certain cases). For example the sentence in (13) is felicitous, whereas the sentence in (14) is not.

(13) Yoav hoped that he will get to the party on time

(14) *Yoav was sorry that he will get to the party on time

We do not know of a study that described this lexical-semantic property. Studying it is interesting because, unlike factivity and certainty, this attribute does not add a

communicative meaning. Insensitivity to it might mean that the representation of the verb's lexical-semantic attributes are inaccessible to these patients, and not "only" the characteristics of MSVs that are crucial to understanding the predications these verbs appear in.

Next I will review the relation between the MSVs and TOM during the acquisition of the two and the relation between them as reported in studies of ASD and other populations that show a TOM deficit.

1.5.2. Development of TOM and MSV acquisition

The processes of language acquisition and TOM development are interrelated in a number of ways. In this section we will focus on the lexical semantic aspects that relate MSVs to TOM. The syntactic relation, namely the dependency, if exists between complex sentences and TOM will be reviewed in the next section (1.4).

The most prominent connection MSVs have with TOM is that they are the vocabulary of the mental realm. The acquisition of MSVs is the acquisition of the ability to organize the knowledge and understanding of the mental world, one's own and that of others. Their acquisition is considered a special challenge especially because the mental processes are abstract and because each mental verb shares aspects of its meaning with other mental verbs, yet is also distinct, and often has multiple meanings (Booth et al., 1997; Montgomery, 1997, 2002). The way children and adults use MSVs allows researchers to understand possible ways of organization of the mental field and the development of this organization (Papafragou, Cassidy, & Gleitman, 2007; Schwanenflugel, Henderson, & Fabricius, 1998; Schwanenflugel, Fabricius, & Noyes, 1996; Tager-Flusberg, 2000).

A second connection between TOM and MSVs is that the developmental courses of both seem to proceed in parallel. Roughly, the TOM ability of 3 year olds allows them to understand that others' actions are directed by their intentions. But, at this age, children are still unable to predict another character's action, if he holds a conflicting mental picture about reality than they do, and therefore fail in false belief tasks. Four

year olds start to show the ability to coordinate two conflicting views of the same reality, and succeed above chance in false belief tasks, and 5 year olds reach perfect performance. MSVs start to appear in the discourse of young children in the second year of life. First, they are used only as conversational devices and at the end of the third year they are used in reference to another person's thought (Naigles, 2000). By age 4 they start to appreciate the different lexical qualities these words carry. They begin to understand that *know* marks a statement as more reliable than *guess* or *think* and by 5 years this distinction is well understood (Moore, Bryant, & Furrow, 1989). This task is especially complex since the input from adults can be very misleading at times. For example, low certainty makers (e.g., *think*) are sometimes used in certain contexts to state high certainty, For example, a father might say: *I think it's time for you to go to sleep,* to communicate a very high certainty message, maybe even an order: *It's time for you to sleep*! (Naigles, 2000).

Significant correlations were found between measures of TOM and the use of cognition verbs in spontaneous speech (Astington, 1998) and between success in TOM tasks and the performance on tasks tapping the certainty distinction between *think*, *guess*, and *know* (Moore & Furrow, 1991; J. de Villiers & Pyres, 1997). The process of understanding certainty continues into elementary school years. Schwanenflugel et al. (1998) showed different scenarios to elementary school children (between 8-12 years old, third and fifth graders) and adults and asked them to choose from a given list *"the words that described how you might use your minds in each situation"* (p. 514). The researchers calculated correlations between the verbs chosen in each situation. They assumed that if the children consider a particular lexical property (e.g., certainty), words that share the same value on this continuum (high or low) will appear together. The results show that this ability developed within the ages that participated in the study. The older children differentiated between MSVs high and low in certainty (and between MSVs that describe input and output), as did adults. So, whereas the ability to understand certainty emerges at the time TOM is developing, it also continues to develop in elementary school years. It will be interesting to consider an acquired loss in this ability due to a general difficulty in understanding intentions.

A third connection between TOM an MSVs inherently concerns the issue of factivity. Both TOM and MSVs are based on representations at two independent levels. TOM

allows a separation of two representations: the facts about some event in reality and the opinions or others people's knowledge of those facts as well as between the opinions of two people. In parallel, the understanding of MSVs depends on the separation between the state of affairs or facts in the embedded clause and the opinion about them, presented by the matrix verb. For example, understanding the utterance *I think it is raining outside*, requires an understanding of the independence of two truth values. This is not the case for other sentences in which the truth of the matrix clause is the same as that of the embedded clause (e.g., *I saw that it was raining outside*). The same independence between levels of representation holds for TOM. It is the realization that thought exist abstractly and they are detached from reality (Moore & Furrow, 1991; Naigles, 2000). Studies showed a correlation between the ability to understand complex sentences and succeed in false belief tasks (de Villiers & de Villiers, 2000). This issue will be discussed further in section (4.4).

Findings reviewed in this section show that the lexical-semantic properties of MSV are relevant in meaning and the age of their acquisition is similar. But it is relevant to note here that there is an important difference between the process of acquisition and the event of acquired loss, the second must not be considered a reflection of the first. The process of acquiring TOM and the ability to create complex sentences might be dependent in the acquisition process but detached from each other in the adult function (Apperly et al., 2009) and therefore will not necessarily correlate after mental loss. Next we review the connection between the ability to create complex sentences and loss in TOM.

1.5.3. TOM and MSVs in Autism and brain damage

Autistic spectrum disorder (ASD) children and adults have difficulty in performing different TOM tasks. The question here is whether they also have difficulty with the mental vocabulary. Baron-Cohen et al. (1986) showed that compared to control subjects, children with autism provided fewer mental state terms in their narratives for a sequence of pictures depicting a simple false belief scenario. This finding might demonstrate a linguistic problem in using these verbs or it may demonstrate these children's inability to perceive the mental aspects presented in the pictures. Another study that tested semantic knowledge of mental terms asked participants diagnosed as autistic to identify words that name mental activities (e.g., *"something the mind can*

do") from a list of words. They were also shown a different list and asked to identify words related to the *body*. The ASD group was compared to a group of moderately mentally handicapped children with no Autism. The ASD group scored significantly worse on identifying mental words. Only 4, out of the 15 children tested, received the best score, identifying 6 out of 8 mental words presented. The same children had no difficulty identifying words related to the body (Baron-Cohen, Ring, Moriarty, Schmitz, Costa, & Ell, 1994). The results show a dramatic interaction between group and word kind but it is important to note that while the mental words that appeared in the list were *verbs* (e.g., think, dream, know, pretend), the words relating to the body were *nouns* – names of body parts (e.g., hand, eye, face nose). It is reasonable to assume that the words on the second list were much more familiar to the participants.

Should we expect a distinction between the ability to understand the presupposition and the implication of mental predications as a result of damage to TOM? We know of two studies that showed such dissociation. One tested high-functioning children with autism (Dennis, Lazenby, & Lockyer, 2001), the other tested children who suffered early focal brain damage either to their right or to their left hemisphere (Eisele et al., 1998). Note, that in the second study the children's TOM ability was not assessed directly, rather, a connection between the right hemisphere and pragmatic ability was assumed. Dennis et al. (2001) tested high functioning ASD children's ability to understand *presuppositions* by asking them to judge the truth value of factive predication. They were presented with factive (e.g., *Karen knows that the door is shut*) and non factive predications (e.g., *Karen thinks that the door is shut).* The matrix verb was either affirmative (e.g., *know*), semantically negative (e.g., *sorry*) or grammatically negated (e.g., *did not know*) complemented with affirmative or negative clauses. They were asked to answer *yes/no* or *maybe* to verification questions[5]. Their ability to draw implication was tested in a similar way by presenting them predication with implicative (e.g., *remember, forget, manage*) and non implicative (e.g., *want*) mental predicates, in affirmative or negative. They were asked to judge their truth value. Results showed that the high-functioning ASD participants understood the presuppositions similarly to a matched aged control group, but their ability to draw implications was significantly worse. This meant that they performed

[5] The one question they give as example is a verification of the matrix verb and not the complement it is: *Does Karen know the door is closed?*

differently on factive and nonfactive predicates, showing a differentiation between presuppositions and implications. Another study showed similar results: Eisele et al. (1998) tested children (4 to 17 years old) who suffered left or right focal brain damage at least 2 years prior to the research. They were presented factive and nonfactive construction using three verbs: *know*, *forget*, and *remember*. The predications were either factive (e.g., *Max remembered that he locked the door*) or nonfactive that can create only implications (e.g., *Max remembered to lock the door*). The matrix verbs were either affirmative or negative. The participants were asked two verifications of truthfulness, one of the complement (e.g., *Did he lock the door?*) and the second, of the matrix predication (e.g., *Did Max remember?*). In order to receive a score for correct understanding the participants had to answer both questions correctly. Results showed that children who suffered left hemisphere damage erred more in answering to sentences that included negation, both syntactic (e.g., *not remember*) and lexical (e.g., *forgot*). It seems they had a general disability to compute the scope of negation in the two kinds of negative sentences. As a consequence, they did not understand both implication and presuppositions. In contrast, participants with right hemisphere damage showed difficulty only on the lexical negated predication, a difficulty which resulted in a disability to understand implications, not presuppositions (Eisele et al., 1998).

These results are important because they show that the theoretical differentiation between two kinds of predications, those that have a factive reading and those that are non-factive, has psychological reality. Early damage to the left hemisphere was correlated with grammatical difficulties, whereas the early damage to the right hemisphere was correlated with an inability to understand implications rather than presuppositions. The same differentiation between presuppositions and implications was also meaningful between the performance of high functioning ASD children and their age matched controls.

Another research (Spanoudis et al., 2007) tested school children who were diagnosed (by their teachers or by explorative testing conducted by the researchers) as having linguistic difficulties on their ability to understand MSVs. The researchers administered Bishop's CCC (Children's Communication Checklist, Bishop, 1998) to assess the pragmatic ability of the group, and 4 tasks of MSV understanding as a measure of their mentalizing ability. Their aim was to see if the children who were

reported as having language difficulties and communication/pragmatic difficulties[6] on Bishop's CCC would score differently on the MSV tasks than children who had the same language difficulty but it did not co-occur with pragmatic difficulty. Two of the tests were semantic (to define mental words from context) and two were lexical (choosing an implication or a presupposition that fitted a factive or a non factive predication). The general result was that the children who received the lower scores on the Bishop's CCC pragmatic assessment also received the lowest scores in all the MSVs tasks, but they were not significantly worse than the children who only had language difficulties. The age-matched control group of typically developing children, with no language difficulties, scored significantly better than the two groups, on all task. These results show that a profile of pragmatic difficulty can be intertwined with certain linguistic abilities[7] and cannot be easily differentiated from other linguistic problems (Spanoudis et al., 2007).

Concerning the property of certainty, Ziatas, Durkin, and Pratt (1998) found that ASD children performed significantly worse than children with Asperger syndrome, children with specific language impairment (SLI) and normally developing children on a task that asked to differentiate between high and low levels of certainty. A correlation was found between the performance on a false belief task with the MSV task for the ASD group and the Asperger syndrome group. Majority of the last group passed both and a majority of the former, failed in both. Similar findings were reported by Tager-Flusberg (2000). She presented a certainty task that compare the verbs *know* and *think* and a standard location change false belief task to three groups. A group of autistic children, a group of mentally retarded children who were of similar age and verbal ability, and healthy preschoolers. Her main findings replicated those reported by Ziatas et al. (1998). She found a significant correlation between the language and theory of mind tasks for all three groups. She also calculated the number of children in each group that either failed or passed each of the tasks and found that for preschoolers and the mentally retarded children, almost one third of each group

[6] For a child to be assigned to the experimental group her verbal intelligence measured by the WISC-III standardized version in Greek should differ from the mean VIQ at least by 1 standard deviation (SD), whereas her score on performance IQ had to be within the normal range.(pg.491-2)

[7] It could be that the linguistically impaired group encountered a different difficulty than the PLI group, and the same low scores do not reflect the same difficulties.

passed the false belief task while failing the language task, but the autistic group either failed or passed both[8]. This finding was explained as resulting from a closer connection between language and TOM in the ASD understanding.

In light of the intertwining connections between TOM and MSVs and the earlier studies showing problems in understanding for ASD and right hemisphere damaged children we see the ability to use and understand MSVs as another important intersection that may allow us to learn about the connection between TOM and language.

[8] Only one out of 16 passed FB and failed language.

2. General Method

2.1. Participants

A group of 25 right brain damage (RBD) patients took part in the study, 8 of them female and 17 male (See Table 1 for a detailed description of the participants' background). Their mean age was 53 years (ranging between 25- 65 years SD = 11). Twenty of the participants were native Hebrew speakers. The five that were not native speakers of Hebrew spoke the language for at least 57 years (Dov for 60 years, David for 59 years, Dror and Rachel for 58 years, and Sara for 57 years). One participant had elementary school education (8 years), 18 of them received high school education, and 6 received academic education. Twenty four of the participants were right handed and one left handed. Twenty four of the patients suffered one incident of CVA in their right hemisphere, 2 of them suffered recurrent infarctions (David and Yigal). One of the CVA patients also suffered covernoma in her Rt-frontal lobe (Gila) and another suffered vertobasiliar meningioma (Sara). One other patient was surgically treated for removal of parieto-frontal-temporal tumor (Sachar). Eighteen of the participants suffered left Hemispatial *Neglect*, a neurological phenomenon characterized by a difficulty to attend to the left side of the visual field. All the participants were tested at least 2 months post onset.

The patients were included in the study on the basis of their having a lesion in the right hemisphere (only). We did not consider reports about their TOM abilities prior to the inclusion in the study.

In addition, because most of the tests used in the study were new and were created specifically for the research question, they had no previous norms. Therefore they were presented to a control group of 25 adults, with a mean age of 51.3 (ranging between 27-66). Twenty four of them were native Hebrew speakers, and one spoke the language for 60 years. One of the participants received elementary education, 6 high school education and 18, academic education. Part of the control group took part in every task. The specific number of control participants was given in the description of each task.

Table 1. Participants' background

	Gender	Age	Spoken Language	Education	Neglect	Hand-edness	Lesion site and Etiology	Mon. post Onset
David	M	65	Hebrew, Arabic	High School	yes	R	Recurrent CVA ischemic hemoragic transformations in frontal and parietal lobes	3
Tzipora	F	60	Hebrew	High School	yes	R	Ischemic infarct in the Right MCA territory	5
Dafna	F	50	Hebrew, English	High School	yes	R	Ischemic CVA Pones	4
Sara	F	63	Hebrew, Romanian	Elementary	no	R	Recurrent Right Parietal parasagittal meningiomas, Parietal craniotomy and CVA vertebra-basiliar stroke.	5
Abraham	M	52	Hebrew	High School	yes	R	Ischemic infarct in the Right MCA territory	2 years
Sason	M	65	Hebrew	Academic	yes	R	Right parietal ischemic infarct	9
Arye	M	48	Hebrew	High School	yes	R	Ischemic stroke with hemorrhagic transformations Right MCA	8
Dror	M	64	Hebrew, Arabic	High School	yes	R	Ischemic infarct involving Right fronto-temporal-parietal areas.	6
Jacob	M	51	Hebrew	High School	yes	R	CVA- Ischemic stroke (Talamus, Internal Capsule)	
Yigal	M	54	Hebrew	High School	no	R	Recurrent right ischemic infarct involving frontal areas and the Corona Radiata.	6
Dov	M	65	Hebrew, English, Polish, Idish	Academic	yes	R	Ischemic infarct in the territory of the Right MCA.	9
Daniel	M	55	Hebrew, Arabic	High School	no	R	Ischemic infarct in the territory of the Right MCA, complete block of Rt. ICA	2

	Gender	Age	Spoken Language	Education	Neglect	Hand-edness	Lesion site and Etiology	Mon. post Onset
Sachar	M	36	Hebrew	High School	yes	R	Craniotomy for removal of Right parieto- frontal-temporal tumor.	7
Rachel	F	62	Hebrew, English, Arabic	High School	no	R	Right Frontal ischemic infarct.	10
Simon	M	58	Hebrew	High School	yes	R	Ischemic infarct in the territory of the Right MCA.	2
Gdalia	M	58	Hebrew, Arabic	High School	no	R	Ischemic infarct in the Right internal capsule.	5 years
Gila	F	56	Hebrew	High School	yes	R	Right Frontal covernoma – Craniotomy and evacuation of intra cerebral Rt. Frontal Hematoma CVA Ischemic Infarction in Right Capsular Putaminal and Right Thalamic regions	2
Yaron	M	46	Hebrew	High School	yes	R	Ischemic infarct involving the Right Corona Radiata.	7
Tzvi	M	25	Hebrew	Academic	Hemi-anopsia	R	Ischemic infarct in the territory of the Right MCA, with brain edema and pending herniation.	7
Sigalit	F	62	Hebrew, English	Academic	Hemi-anopsia	L	Right occipital infarct. Spontaneous intraparenchimal hemorrhage	2 years
Danny	M	54	Hebrew	Academic	no	R	Ischemic infarct in the territory of the Right MCA.	3
Moshe	M	59	Hebrew	High School	no	R	Subacute infarct in the RH, adjacent to the internal capsule, caudate body, across to the Globus Pallidus	10
Sharon	F	38	Hebrew, English	High School	yes	R	Ischemic infarct in the territory of the Right MCA.	5
Ahuva	F	32	Hebrew	High School	no	R	Ischemic infarct in the territory of the Right MCA.	4
Ayal	M	47	Hebrew	Academic	no	R	Ischemic infarct in the territory of the Right MCA. Right periventricular infarct.	7

Memory tests

The RBD participants were tested on three memory tests in order to assess whether they suffer loss in these abilities. They were administered working memory tests from the FriGvi battery (Gvion & Friedmann, 2008; Friedmann & Gvion, 2003). These tasks were presented to 18 RBD patients. See the first column in Table 2 for detailed list of the patients who participated in this task.

Table 2. The participants who took part in each of the tasks

	Tasks																			
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
David		*						*	*	*										*
Tzipora	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Dafna	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*		*	*
Sara	*	*																		*
Abraham	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*		*	*
Sason		*	*	*				*					*	*	*	*			*	*
Arye	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*		*	*
Dror	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Jacob		*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*		*	*
Yigal		*	*				*	*	*	*	*	*	*	*	*				*	*
Dov	*	*						*	*	*	*	*	*	*	*					*
Daniel		*	*	*			*	*	*	*	*	*	*	*	*	*		*	*	*
Sachar	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*		*	*	*
Rachel		*							*			*			*				*	*
Simon		*	*	*	*	*	*	*	*	*	*	*	*	*	*	*			*	*
Gdalia		*	*	*																*
Gila	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Yaron	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*		*	*
Tzvi	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*		*	*
Sigalit	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*		*	*
Danny		*																	*	*
Moshe	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*		*	*
Sharon	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Ahuva	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*		*	*
Ayal	*	*	*	*		*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Controls	-	14	10	7	8	8	7	8	9	9	7	12	9	12	10	-	7	-	-	14

(1) Memory (section 2.1.1.)

(2) TOM (chapter 3)

(3) Production of the definite or indefinite (section 4.1.1)

(4) The use of descriptions in Discourse (section 4.1.2)

(5) The grumble test (section 4.1.3)

(6) Story retelling (4.2.1.1.)

(7) Production of appropriate differing descriptions (section 4.2.1.2.)

(8) Choosing between high and intermediate accessibility term (section 4.2.2.1.)

(9) Choosing between high and low accessibility markers (section 4.2.2.2.)

(10) Choosing between two kinds of low accessibility markers (4.2.2.3.)

(11) Choosing between NPs (4.2.2.4.)

(12) Factivity (section 4.3.1.)

(13) Factivity in context (section 4.3.2.)

(14) Certainty (section 4.3.3.)

(15) Matching semantics to tense of complement (4.3.4.)

(16) Comprehension of relative clauses (section 4.4.1.)

(17) Production of relative clauses (section 4.4.2.)

(18) Comprehension of Wh questions (section 4.4.3.)

(19) Comprehending binding principles (section 4.4.4.)

(20) TOM and sentential complements (section 4.4.5.)

Three kinds of memory tasks were presented. Two were: *Word and nonword spans*: Word and nonword lists were presented orally at a rate of one item per second. The participants were asked to recall the items serially. Each span test included 6 levels, of 2-7 words or nonword sequences, with 5 sequences per level. The memory span of each participant was defined as the maximum level at which at least 3 sequences were fully recalled; half a point was given for success in 2 out of 5 sequences (e.g., a participant who recalled three 3-word sequences and two 4-word sequences had a score of 3.5). The *word span* test included sequences of 2-syllable unrelated words, and the *nonwords span* included 2-syllable nonwords, constructed by changing a single consonant in real words.

The third test was *Matching word order span*, which assessed input phonological working memory span. The participants heard two lists of 2-syllable unrelated words and were asked to judge whether the order of the items in the two lists was the same. The task included 6 successive levels, 10 items in each level. The first level was composed of lists of two words, and the last of two lists of 7 words. Each level was composed of 5 matching and 5 non-identical pairs. The sequences were presented at a one word per second rate. The span level was defined as the maximal level at which the participant performed correctly on at least seven pairs.

Table 3 presents the scores of the participants that were tested on these memory tasks compared to the score of age matched controls (Gvion & Friedmann, 2008).

The results show that the majority of the participants scored no differently from their age matched controls. Two patients scored significantly lower on the word span test (Dafna and Yaron) but they scored no differently from age matched controls on the other two tests. Three patients scored significantly lower than their age matched controls on the matching word span test (Abraham, Jacob, and Gila). Two of them (Abraham and Gila) scored within the normative range in the other two tests and the third (Jacob) was not tested on the other two tests.

Table 3. Working memory scores of RBD patients compared to aged matched norms

	Word span	Norm	SD	Non word span	Norm	SD	Matching Words	Norm	SD
Tzipora	4	5.05	0.64	3.5	3.15	0.34	5	6.1	0.87
Dafna	*4	5.54	0.45	3	3.29	0.45	6	6.8	0.42
Sara	4	4.86	0.78	3	3.4	0.46			
Abraham	5	5.05	0.64	4	3.15	0.34	*4	6.1	0.87
Arye	5	5.54	0.45	3.5	3.29	0.45	6	6.8	0.42
Dror	4	4.86	0.78	3	3.4	0.46	5	6.27	0.79
Jacob							*4	6.1	0.87
Dov	4.5	4.86	0.78						
Daniel							6	6.1	0.87
Sachar	4.5	5.57	0.75	3	3.46	0.54	5	6.33	0.98

	Word span	Norm	SD	Non word span	Norm	SD	Matching Words	Norm	SD
Gila	4	5.05	0.64	3	3.15	0.34	*4	6.1	0.87
Yaron	*4	5.54	0.45	3	3.29	0.45	6	6.8	0.42
Tzvi	6	5.39	0.67	4	3.29	0.47	7	6.4	1.07
Sigalit	6	4.86	0.78	3.5	3.4	0.46	7	6.27	0.79
Moshe	5	5.05	0.64	3	3.15	0.34	5	6.1	0.87
Sharon	6	5.57	0.75	4	3.46	0.54	6	6.33	0.98
Ahuva	5	5.57	0.75	3.5	3.46	0.54	5	6.33	0.98
Ayal	5	5.54	0.45	4	3.29	0.45	6	6.8	0.42

*Significantly poorer than the age-matched control group, $p < .05$, using Crawford and Howell's, 1998 t-test

2.2. General procedure

The patients were recruited from three rehabilitation centers. The study was approved by the Helsinki ethical committee of each establishment and by the ethical committee of Tel Aviv University. Each participant gave his written consent to participate in the study.

The tests were administered to each patient separately in a quiet room in the rehabilitation center where he/she was treated or in his/her home. Each session lasted between 30 min to an hour, depending on the participant's ability to remain attentive and breaks were given whenever necessary. All the sessions were audio recorded and transcribed later. The number of meetings varied depending on the participant's willingness to continue taking part in the study. Due to this restriction not all the participants took part in all the tasks reported on in the study. A detailed list of the participants that took part in each task is reported in Table 2.

2.3. Statistical analysis

For comparing between the different groups of participants in their performance on the same task we used the Mann Whitney test. It is a nonparametric test for the significance of the difference between the distributions of two independent samples.

For comparing the differences within the groups, between two conditions, we used the Wilcoxon signed ranks test, a nonparametric test for the significance of the difference between the distributions of two non-independent samples involving repeated measures or matched pairs.

To test whether an individual patient score was significantly different from the control group average we used the Crawford and Garthwaite (2002; see also Crawford & Howell, 1998) *t*-test for the comparison of a single subject to a group, with an alpha level of 0.05.

Other statistical tests that used were reported together with the results of the different tasks.

In most of the comparisons we had a prior hypothesis about the direction of the difference between the groups or conditions, therefore the *p* values reported are a result of a one- tailed comparison unless specified differently.

3. Theory of Mind after Right Brain Damage

The relations between TOM and damage to the right hemisphere have been discussed in the introduction (1.1). In this chapter we describe the *aTOMic battery*, the test battery that we devised in order to test the participants' TOM abilities, and its findings.

3.1. Method

Participants

The participants were 25 brain damaged patients and 14 control participants. See Table 2 for detailed list of the patients who participated in this task.

Materials:

aTOM battery

The *aTOM battery* included 8 categories. Each category assessed a different aspect of the TOM ability. Two items were presented in each category. The use of the two items ensured that the results we found for each category were not arbitrary. It also assured us that if a certain patient suffered a decrease in TOM, the problem will be detected.

The categories:

1. False belief: based on the original false belief task (Wimmer & Perner, 1983) we presented passages describing a change of location of an item. The change takes place without the protagonist knowledge. We asked the participants two questions: a memory/reality question: *where is the item*? and a false belief question: *Where will she/he look for that item?* One item (A) was about a misplaced book the other (B) about a misplaced garden tool.
2. Second-order false belief: we asked about one character's belief about a second character's state of knowledge. In order to reduce the cognitive load of presenting two scenarios[9], we presented one story and asked two second-order false belief questions

[9] In an earlier version of the battery we presented two stories, one of the items was translated from Hughes et al. (2000; also appeared in Astington, Pelletier, & Homer, 2002). The second was written according to it. The translated item was later discarded and replaced because the group of control participants answered it without reasoning about the second character's state of knowledge. The item that we wrote to match the first was also discarded because it was obscure; a considerable percentage of the control participants erred in its memory questions.

about it. The story was presented in two stages. First we described a state of affairs between two characters. The story was about a mother who planned to give her son a puppy as a surprise birthday gift. To keep it a surprise she told him he will get some other gift. While she was out, the son discovered the puppy and understood that it was his real birthday gift. Next, we asked the participant two questions: a memory question and a comprehension question. Then we continued the story and introduced a third character, a friend of the son. Each of the three characters held different information regarding the others' thought and knowledge about the birthday gift. All three characters knew what was the planned as a surprise, but the mother thought her son did not know what it was, and the son's friend didn't know that the mother was hiding what the present was. Next we presented a yes/no question about the second order belief of the sons' friend (A) and asked the participant to explain their answer, and another second order belief question about the mother's thoughts (B) and again we asked for an explanation. Finally, we asked two other memory questions to ensure that the story was understood. The item was adapted from Sullivan et al. (1994).

3. Knowledge gaps: two stories presented similar interactions between two characters. The first character attempted to steal jewelry or cheat on a test. The second character approached the first *not knowing of the attempt*. The first character (unaware of the first's naïve reason for approaching him) admitted the act. We asked two questions: *Why did the second character admit?* (because he misunderstood) and *Was the second character aware of the first's motive*? (no). The item about the burglar who gave himself up (A) was translated from Happé et al. (1999). The second, about a girl who confessed she was about to cheat on a test (B) was written by us along similar lines.
4. Instruction: we presented 4 events of teaching (e.g., of tying shoelaces, learning to read, learning a card game and learning a sports game). In two items the teacher mistakenly thought that her pupil knows the act to be taught (Teacher thinks Pupil Knows – TPK), when in fact she didn't know. In the other two items the teacher mistakenly thought the pupil didn't know what she intended to teach, when in fact she did know (Teacher thinks Pupil Doesn't Know- TPNK). In all cases we asked: *Will the teacher attempt to teach?* and *Why?*. In the TPK items a correct answer was that the teacher will not attempt to teach (A) a sport game or (B) a card game because she thinks the pupil already knows. In the TPNK a correct answer was that the teacher will try to teach (A) to tie laces or (B) read, because she thinks her pupil doesn't know. These items provide an opportunity to learn whether the participants decide if

the teaching takes place according to the state of knowledge of the teacher about the pupil or according to the pupil's actual state of knowledge about the task to be learned. The items were adapted from Ziv and Frye (2004).

5. *Faux pas*: two tasks tested the ability to appreciate that a persons' rude behavior resulted from her/his lack of knowledge about the situation. One of the items (A) was about a remark one friend made about an urgent need to replace curtains which were actually brand new, translated from Baron-Cohen et al. (1999). The second (B), was about a complaint a young man made about a noisy drum lesson, not knowing that his cousin was one of the participants in the lesson. This item was written according to the first and resembled it in length and complexity. We presented three questions: a memory question, a yes/no question about the point of the story and an explanation to the second answer.
6. Surprise: in these two items one of the characters acted as if she was surprised although she had found out about the plan to surprise her. We asked: *Were the things she said true? Was she really surprised?* (no) and *Why did she say it?* (not to cause disappointment). Both items, (A) about a trip and (B) about a party were based on the surprise stories presented by Happé (1994).
7. Empathy: two stories presented a person who approached his/her friend to say he/she was sorry about the friend's failure in a test, before that friend found out about his/her failure. We asked four questions. The first was a memory question. The second was the empathy question: *Why was X sorry?* The last two questions were second-order false belief questions about the emotional situation. One was a yes/no question, *What did X think when he approached the friend? That she already knew about her failure or that she didn't know about it?* The second was a justification. One of the items (A) was about a girl who was not chosen for a sports racing team, was translated from Baron-Cohen et al. (1999). The second, (B) about a failure in a university course was written along similar lines.
8. Cartoon: Two cartoons were presented. The key to understanding them was grasping the figures thoughts and intentions. In one (A) a viciously smiling shark was holding out a stick that looked like a hand waving for help, while a man standing on a nearby rock is getting ready to jump in. In the second cartoon (B) a rooster is strangling a hen and next to them a small elephant is hatching from an egg. First we asked the participants to describe what they saw in the picture. Then we asked what they thought the cartoon was about and if they found it funny in any way. For the first

cartoon an appropriate description would be one that mentioned the shark's intentions to lure the man to jump in the water. For the second cartoon an appropriate answer would be one that referred to the rooster's understanding that his partner cheated him. The cartoons did not include any words.

Procedure

The items were presented in a semi-random order. They were read to the participants in normal rate and intonation and were reread again if necessary. The participants were encouraged to follow the text while the experimenter read the items aloud. Later they were given a chance to read it themselves if they wished to do so. The task was presented in one or two sessions, according to the participant's needs.

3.2. Results

Coding: The memory questions were used to verify that the participant listened and understood the events in the story. If an incorrect answer was given, the item was read again or a clarifying remark was added. The yes/no TOM questions were coded as either correct or incorrect (1/0). The justifications were coded twice. First, a *general cognitive score* (0/1) was given to each answer, whether the justification was appropriate or inappropriate (see below). This score was used for calculating the final score for each item. The final score was calculated by multiplying the score to the yes/no question by the general justification score. Therefore, a score of (1) was given only if the participant gave a correct yes/no answer as well as an appropriate justification.

In addition to the general cognitive scoring, each justification was graded more specifically. This scoring was based on the parameters used by Tager-Flusberg & Sullivan (1994a) and Happé et al. (1998).

As explained above, the justifications were *generally* graded as *appropriate* or *inappropriate*. The appropriate justifications included those that were either *explicit and specific* about the mental attitude of the characters or only a *general and implicit* mental justification.

The *inappropriate* justifications were also classified into one of two categories: answers that mentioned *some mental aspect* of the situation, even though it was an irrelevant one, and those that concentrated only on *physical aspects of the situation and irrelevant answers* (e.g., *the story doesn't say*…). See Appendix for examples of justifications given by the participants.

For example, in one of the two *'surprise stories'* a man is planning to surprise his girlfriend with tickets to a trip abroad. He tells her they will be going to a city in the their own country. His girlfriend finds the receipt and understands he is planning a surprise for her. Nevertheless, when they approach the airport she reacts as if she didn't know about it and cries out – "Wow, I can't believe it, I'm so surprised, we're going abroad, that's wonderful!" The participants were first asked if it was true what the girlfriend said: *'Was she really surprised, or not?*' and then they were asked for a justification: *Why did she say that?*

Examples of the four kinds of justification (1-4) are presented below:

(1) Appropriate and explicit: *She didn't want him to know she knew about the tickets* (Avraham)

(2) Appropriate but general: *So he won't feel bad* (Sigalit).

(3) Inappropriate but mental: *Because she was happy about it* (=the trip). (David)

(4) Inappropriate and physical: *I suppose (because) he didn't go to Jerusalem, he drove to the air-port.* (Sason)

3.2.1. Data analysis

The *aTOM battery* included 8 categories of abilities that represent different aspects of the TOM ability (false belief, second order false belief, knowledge gaps, instruction, *faux pas*, surprise, empathy and understanding cartoons). Each category was tested by two items.
Table 4 details the average score on each item in the *aTOM battery* for the RBD group and the control group.

Table 4. Percentage of correct responses of the RBD patients and control group to two items in each task

	False belief		2nd order FB		Know. Gaps		Instr. TPK		Instr. TPNK		Faux pas		Surprise		Empathy		Em. 2nd order		Cartoon	
	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B
Rt. CVA	84	84	50	78.6	56	88	33.3	25	35	85	52	44	72	68	76	88	60	68	36	36
n =	25	25	14	14	25	25	21	20	20	20	25	25	25	25	25	25	25	25	25	25
Control	100	100	78.6	100	100	100	78.6	86	71	93	85.7	100	100	92.9	100	100	100	100	100	86
n =	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14
p =	.21	.21	.1	.17	.27	.01	.01	.002	.04	.36	.04	.002	.08	.1	.11	.27	.02	.05	.0006	.006

From these scores we can see, first, that the control group scored better on all the tasks, although they did not reach ceiling on all of them. Using Mann-Whitney ranked order test, we tested whether there were significant differences in the performance of the two groups in each task (The *p* values of each comparison are presented in the bottom line of Table 4). A significant difference was found in both tasks in the category of *instruction- TPK*, when the teacher was expected not to teach because she thought her student already knows what she was about to teach (a sport game: U = 213.5, $p = .01$, a card game: U = 225, $p < .01$). A significant difference was also found in both items in the *faux pas* category (making an ignorant rude remark about a drum lesson: U = 273, $p < .01$, and about new curtains: U = 234, $p < .05$). In both items in the *empathy second order false belief* task (when someone was showing empathy after a friend's failure to make the race team: U = 245, $p < .05$, or pass a course test: U = 231, $p = .05$), and in both of the *cartoons* (shark: U = 287, $p < .001$; and rooster: U = 262, $p < .01$). In the category of *instruction TPNK*, where the teacher was expected to teach because she was not aware that her student already knew what she was about to teach, two different results were found. One of the items (teaching to read) was easy for the RBD group (item B), and they received a high score, which was not different from that of the control group (U = 151, $p = .36$). The other item (A), about teaching how to tie shoe laces turned out to be unclear to the participants in all groups; although a significant difference was found between the groups, even the mean score of the control group was considerably low, as 4 of the 14 control participants erred on this task. Therefore this item was discarded from the battery and from further analysis.

In the *knowledge gap* category there were two different results for the two items. On item A (burglar), no significant difference was found (U = 196, $p = .27$). In item B, (cheat in a test) the RBD scored significantly lower than the control group (U = 252, $p = .01$).

In all the other items, the control group received higher scores than the RBD group, but a significant difference between the groups was not detected. In the *false belief* tasks, for both items: U = 203, $p = .21$. In second order false belief, item A (what will the boy answer his friend) U = 126, $p = .10$; item B (what does the mother think) U = 119, $p = .17$. In the *surprise* category, item A (party) U = 224, $p = .08$; item B (trip) U = 218.5, $p = .10$. In both *empathy* category: item A (running team) U = 217, $p = .11$; item B (university course), U = 196, $p = .60$.

Following this first analysis, we considered the individual scores of the RBD patients in all the items of the *aTOM battery*. The individual scores are presented in Table 5. As explained above, each score was calculated by multiplying the score on the *yes/no question* by the *general score* of the justification answer. According to this procedure only participants who answered correctly and justified their answer appropriately received a score of 1.[10] As shown in Table 5, the group of RBD patients is not homogenous. The scores of the RBD patients ranged between 18.8% correct to 100%. This difference led us to divide the group of RBD patients to two: *aTOMic participants:* participants who scored under 70% were considered *aTOMic,* namely, individuals with a TOM impairment. These patients who suffered RBD showed a considerable decrease in TOM reasoning (n = 17). The other group was *TOMer* participants: participants who received average score higher than 80%, who were considered *TOMer,* namely individuals with normal TOM. These TOMer patients (n=8) showed unimpaired TOM abilities in spite of their right brain damage. All the participants in the *aTOMic* group erred in at least one whole category (two items). The participants in the *TOMer* group succeeded in at least one item in each category. The average score in the *aTOMic* group was 46% (SD = 16%), the average score of the *TOMer* group was 92% (SD = 7.5%). We also found that all the participants in the aTOMic group performed significantly poorer than the control group ($p < .05$, using Crawford and Howell's, 1998 t-test) and that all the TOMers performed as well as the control group. For detailed information see the first column in Table 6.[11]

[10] The empty cells are the cases where an answer was missing, to one or both of the questions. The considerably large number of empty cells in response to the '*second order false belief*' task is a result of the change of the item in the course of the research. See above section 3.2.

[11] It should be noted that one of the control participants formed an outlier in his group, as his performance in the TOM tests was significantly poorer than the rest of the group. Without his inclusion in the test, the aTOMic participants' performance would have been even more distinct from the controls'.

Table 5. RBD patients' performance on the aTOMic battery (blanks represent unanswered questions)

	False Belief		2nd order FB		Knowledge gaps		Inst TPNK	Inst TPK		Faux pas		Surprise		Empathy		Em 2nd ord		Cartoons		Total
	A	B	A	B	A	B	B	A	B	A	B	A	B	A	B	A	B	A	B	%
David	0	1			1	0		0	0	0	0	0	0	1	0	0	0	0	0	18.8
Tzipora	1	1	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	26.3
Dafna	1	0	0	0	1	0	0	0	0	1	0	0	0	0	1	1	1	0	0	31.6
Sara	1	0			0	0	1	0	0	0	0	1	1	1	1	0	0	0	0	35.3
Sason	1	1			1	0				0	0	0	1	0	1	0	0	0	0	35.7
Abraham	0	1	0	0	1	0	0	0	0	1	0	0	0	1	1	1	1	0	0	36.8
Dror	1	1	0	1	0	0	1	0	0	0	0	1	1	1	0	0	0	0	0	36.8
Arye	1	1			1	0	1	0	0	0	0	0	0	0	1	1	1	0	0	41.2
Jacob	0	0	0	1	1	1	1	0	0	0	0	0	1	1	1	1	1	0	0	47.4
Yigal	1	1			1	1				0	0	1	1	0	1	0	0	0	0	50.0
Daniel	1	1			1	1	1	0	0	0	0	1	0	1	1	0	1	0	0	52.9
Dov	1	1	0	1	1	1	0	0	0	1	0	0	0	1	1	1	1	0	1	57.9
Sachar	1	1			1	1	1	0	0	0	0	1	1	0	1	0	0	1	1	58.8
Rachel	1	1			1	0				1	1	1	1	0	1	0	1	0	0	64.3
Simon	1	1			1	1				0	0	1	1	1	0	1	1	0	0	64.3
Gdalia	1	1			1	1	1	0	0	1	1	1	1	1	1	0	0	0	0	64.7
Gila	0	0	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	0	0	68.4
Yaron	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	0	84.2
Sigalit	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	89.5
Moshe	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	94.7
Sharon	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	94.7
Tzvi	1	1	0	1	1	1	1	1		1	1	1	1	1	1	1	1	1	1	94.4
Danny	1	1			1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	100.0
Ahuva	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	100.0
Ayal	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	100.0
%Average	84	84	50	79	88	56	85	33	25	52	44	68	72	76	88	60	68	36	36	
SD (%)	37	37	52	43	33	51	37	48	44	51	51	48	46	44	33	50	48	49	49	
n	25	25	14	14	25	25	20	21	20	25	25	25	25	25	25	25	25	25	25	

Table 6. The individual performance of the RBD patients compared to the control group ($p < .05$, using Crawford and Howell's, 1998 t-test)

	1	**2**	**3a**	**3b**	**4**	**5a**	**5b**	**6**	**7**	**8a**	**8b**	**9a**	**9b**	**10a**	**10b**	**11a**	**11b**	**12**	**13**	**14**
David																				
Tzipora																				
Dafna																				
Sara																				
Abraham																				
Sason																				
Arye																				
Dror																				
Jacob																				
Yigal																				
Dov																				
Daniel																				
Sachar																				
Rachel																				
Simon																				
Gdalia																				
Gila																				
Yaron																				
Tzvi																				
Sigalit																				
Danny																				
Moshe																				
Sharon																				
Ahuva																				
Ayal																				

*Dark blue cells—Significantly different from the control group

*Light blue cells—Not significantly different from the control group

*The participants in the TOMer group appear on shaded background

(1) aTOM battery (chapter3)

(2) Production of the definite or indefinite (section 4.1.1)

(3a)The use of the definite marker in Discourse (section 4.1.2)

(3b)The use of descriptions in Discourse (section 4.1.2)

(4)The grumble test (section 4.1.3)

(5a) Story retelling: with corrections analysis (section 4.2.1.1)

(5b) Story retelling: without corrections analysis (section 4.2.1.1)

(6) Production of appropriate differing descriptions (section 4.2.1.2.)

(7) Choosing between high and intermediate accessibility term (section 4.2.2.1.)

(8a) Choosing between high and low accessibility markers in conjunctions (section 4.2.2.2.)

(8b) Choosing between high and low accessibility markers in embedded sentences (section 4.2.2.2.)

(9b) Choosing between two kinds of low accessibility markers: the less informative option (section 4.2.2.3.)

(10b) Choosing between NPs: the short NP option (section 4.2.2.4.)

(11b) Factivity: Negated non-factives (section 4.3.1.)

(13) Certainty (section 4.3.3.)

(9a) Choosing between two kinds of low accessibility markers: the more informative option (section 4.2.2.3.)

(10a) Choosing between NPs: the detailed option (section 4.2.2.4.)

(11a) Factivity: Negated p-factives (section 4.3.1.)

(12) Factivity in context (section 4.3.2.)

(14) Matching semantics to tense of complement (4.3.4.)

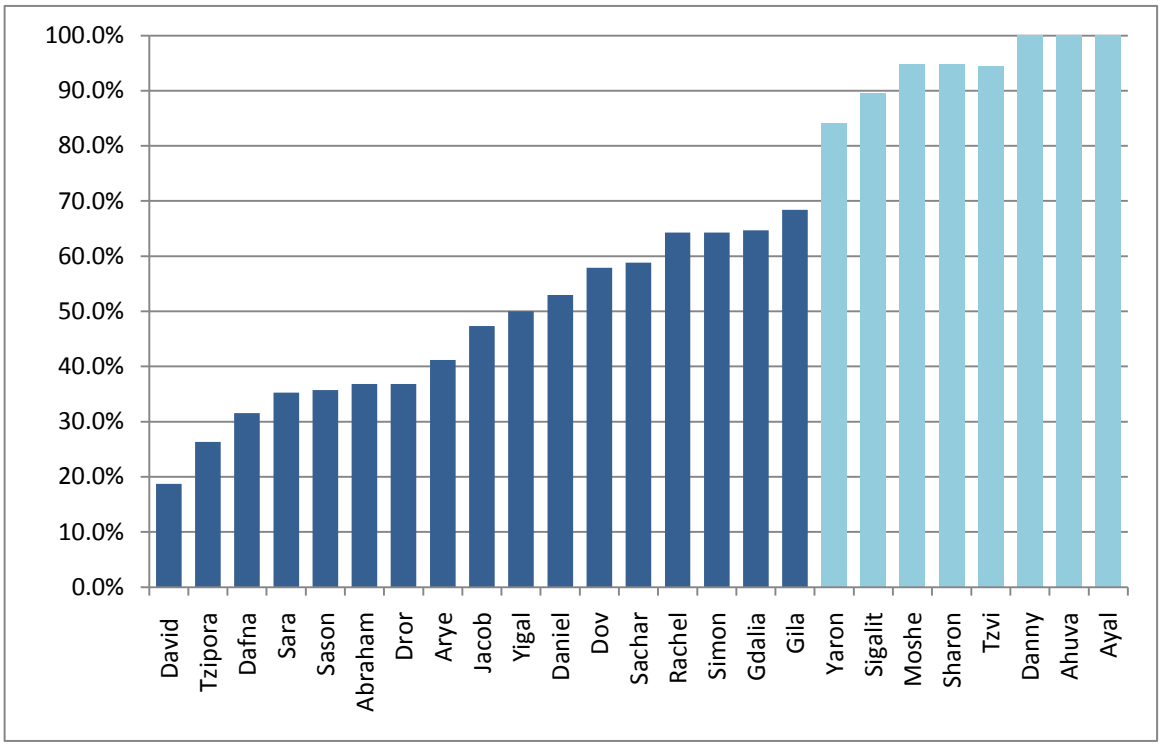

Figure 1. The performance of the RBD group in the *aTOM battery*

Figure 1 presents the average score of each participant, the dark blue bars represent *aTOMic* participants, and the light blue bars represented *TOMer* participants.

Before testing the differences between the three groups (*aTOMic*, *TOM*er, and *control*) we checked if the scores of each couple of items that were designed to tap the

same aspect of TOM, were in fact equal. We compared the scores of the whole group of RBD on each couple of items using 2-tailed McNemar test. (The Instruction TPNK was not included in this analysis because one of the items was discarded due to a low performance of all groups including the controls). The results are presented in Table 7.

Table 7. McNemar comparisons for differences between two items in each category (two tailed *p* values)

False belief	.62
Second Order False belief	.62
Knowledge Gaps	.01
Instruction TPK	1
Faux pas	.62
Surprise	1
Empathy	.50
2nd order Empathy	.48
Cartoons	.48

In 8 of the categories no significant difference was found between the two items (false belief, second order false belief, instruction TPK, *faux pas*, surprise, empathy, second order empathy and cartoons). The results of the McNemar comparisons indicate that these couples can be collapsed to a single score that will indicate the participants' performance on the category. A significant difference was found in one category- *knowledge gaps*. The answers to the two items included in this category was significantly different ($p < .01$). One of the items was easy for all participants, only three of the a-TOMics failed to answer it correctly. The participants that erred in the easy item also erred on the other item. For the following analysis, we averaged the scores between couples in which no significant difference was found. The category *knowledge gaps* was represented by each of the two scores.

Next, we compared the scores of the *aTOMics*, *TOM*ers, and *controls* on each category, and the three items that were kept apart, using Mann Whitney rank order test. The results were presented in Table 8.

A significant difference was found in all categories between the aTOMic group and the TOMer group and between the aTOMics and the control group. Together there were 8 categories and one additional item (from the knowledge gaps category) that detected differences between the aTOMics and the two other groups. The aTOMics scored significantly lower than the TOMers and the controls in these 9 comparisons. In two items the difference was not significant. One of the items is the Knowledge gap item (B) (a burglar who gave himself up because he thought he was caught), and one item in the instruction TPNK, (B) (about a teacher that was expected to try and teach a pupil reading because she didn't know the pupil already knew how to read). In both cases the aTOMic group did well and the items failed to distinguish between them and the two other groups.

In all the categories and individual items the TOMer group scored no differently than the control group.

Table 8. Comparisons between the performance of the aTOMic, TOMer, and control groups on the various TOM items

	False Belief	**2nd or. FB**	**Knowledge Gaps**		**Instr. TPK**	**Inst. TPNK**	**Faux pas**	**Surprise**	**Empathy**	**2nd ord. Empathy**	**Cartoon**
aTOMic vs. TOMer	0.09	0.01	0.25	0.04	0.19	0.002	<0.001	0.02	0.02	0.01	<0.001
aTOMic vs. Control	0.05	0.004	0.21	0.003	0.23	<0.001	<0.001	0.01	0.01	0.003	<0.001
TOMer vs. Control	0.49	0.14	0.49	0.33	0.41	0.19	0.31	0.41	0.49	0.49	0.49

*significant differences are shadowed.

3.2.2. Justifications

As explained above, we exacted the participants to justify their yes/ no choices in 7 out of the 8 categories of the aTOM battery.

Approximately 10% of the scores were judged by two experimenters, there was agreement on approximately 95% of the scores. Differences were resolved in discussion. A detailed description of the items, expected justifications and examples is given in the Appendix.

Table 9 presents the percentage of each kind of justification[12] given by each group of participants. The data in each row add up to 100%.

Table 9. Percentage of different kinds of justification given by the three groups

	Appropriate		Inappropriate	
	Specific and Explicit	General and Implicit	Mental-inappropriate	Physical or Irrelevant
aTOMic	38.4	8.9	18.2	34.5
TOMer	78.9	9.9	7.2	3.9
Control	84.1	8.6	3.3	4.1

As can be seen in the Table above, the aTOMic group gave less appropriate justifications (average = 47.3%) than the two other groups (TOMers – 88.8% and control 92.7%). Accordingly, the percent of inappropriate answers was highest for the a-TOMics. The percent of appropriate and specific answers was different for the aTOMic than the other two groups, but the difference between TOMers and controls was small. Interestingly, all three groups gave almost the same percent of right, but implicit or general answers.

When giving an inappropriate justification, the a-TOMics gave considerably more physical or irrelevant answers than answers with some mental consideration. In the

[12] The false belief category is not included because no justification was asked for in its items.

TOMer group the pattern was opposite, when inappropriate, the larger percent of justifications included a mention of some mental aspect. The control group gave the same percent for the two kinds.

3.3. Summary and Discussion

We assessed the TOM abilities of a group of 25 RBD patients and a 14 control participants, using an *aTOM battery*, which included 18 different items that represented 8 categories, each category represents a different aspect of TOM.

The first important finding is that the group of RBD was far from homogenous. Their overall scores ranged between 18.8% and 100%. According to this finding the group was divided into two, 17 of the patients were defined as aTOMic, patients who suffered a considerable loss to their TOM ability. These patients scored below 70% in the whole aTOM battery. Eight other patients scored between 84% and 100%, were classified as TOMer, participants who do not suffer lack of TOM although they did suffer right brain damage.

These results show that right brain damage does not automatically imply damage in TOM. Although different studies showed RBD did have a variety of difficulties that are TOM-related (e.g., Gallagher et al., 2000; Happé et al., 1999) other studies showed that some RBD patients can pass these tasks as well as their control group (Siegal et al., 1996; Tompkins et al., 2008). The results presented here demonstrate that besides the need to control the materials tested and to explicitly define which aspect of TOM is evaluated (Tompkins et al., 2008), there is a need to directly assess TOM ability and not presume its presence or absence as a side effect to the brain damage.

The aTOM battery enabled a broad assessment of TOM abilities; Most of the items presented differentiated between the aTOMics and TOMers: false belief, second order false belief, instruction TPK, *faux pas*, surprise, empathy, and mental cartoons. No difference was found between the TOMer group and the control group. These results show that an assessment of TOM using a considerably large variety of tasks can allow

this important distinction, between the aTOMic and TOMer groups. This dissociation we found is important both for future research and for clinical purposes.

The analysis of the participants' justifications also showed important findings: the majority of the justifications given by all groups were *explicit and specific* but the aTOMics gave considerably less answers of this kind than the other two groups. The percent of *appropriate but general or implicit* justifications was equal for all groups. This may reflect that in a fixed amount of situations the participants felt that some general mention of reason will be enough to explain the situations; this may have resulted from the phrasing of some of the items that generated this kind of justifications.

The aTOMic group also gave a considerable amount of *inappropriate answers*, compared to the two other groups. Most of the answers in this category were *physical or irrelevant* showing that these participants tended to talk about the state of affairs in reality and not about the way the situation was mentally understood. They mistook questions regarding the content of thought to be questions regarding the situation presented in the items. For example, one of the *faux pas* items described a girl who started to attend drum lessons that take place near her cousin's house. When the whole family met during the weekend, the cousin complained aggressively about the new music classes that started near his house. He said that all the students there are awful and that he is considering filing a complaint to the municipality. We asked whether the complaining speaker knew his cousin was one of the students in the drum class and after a yes/no answer was given, we asked for a justification. Patients from the aTOMic group answered different kinds of physical irrelevant answers:

(5) Tzipora: *He must have heard them making all that noise*.

(6) Dror: *If he is complaining he probably made some kind of inquiries about it.*

(7) Dov: *He must have asked. He heard the drumming; he asked questions, I guess he got answers, from the story itself you can't tell*.

These answers demonstrate the patients' ability to discuss the physical situation but not the mentally-held information each character has about it.

In the following chapters we will present the different linguistic abilities tested. According to the findings gathered in the TOM assessment, we will analyze and compare the language results between the two groups of RBD patients, *aTOMic* and *TOMer*.

4. Linguistic abilities

After establishing the status of TOM abilities in our participants, we moved to the main research question of this study – whether a deficit in TOM leads to specific deficits in language. For this aim we compared the two groups, the a-TOMic and TOMer on a variety of linguistic tasks that will be presented bellow. If such differences exist it will demonstrate that the classification we found in TOM, using the aTOMic battery, is significant for performance of linguistic tasks.

In the first part of this chapter we will present the tasks that relate to the participants' ability to distinguish between different levels of relevance and accessibility of information during discourse. To test this ability, we presented a variety of tasks. First we will present the tasks that tested the ability to use the definite and indefinite markers in introducing information which is new to the addressee (section 4.1). Later we will present the tasks that tested the participants' ability to comprehend and make use of the different kinds of referring expressions, which signal different levels of accessibility of items (in the mind of the addressee) at different stages during the discourse (section 4.2).

Next we will review the tasks that tested the participants' ability to comprehend and produce mental state verbs (section 4.3). Studying the use of these verbs offers an opportunity to test the way deficit of TOM affects the way these verbs, which are the most prominent tool for discussing mental states and situations, are perceived. We tested these patients' command of the semantic and lexical properties of these verbs.

In the last section (4.4) we will review syntactic tasks we presented to the patients in order to assess their syntactic abilities. Some of the tasks aimed at identifying syntactic difficulties that are common in other populations which suffer lack or decrease in their syntactic abilities (e.g., agrammatic aphasia patients and SLI children). Other tasks were administered because they directly relate to the discoursive abilities tested in the first sections.

4.1. Accessible or not: Comprehension and production of the definite article in discourse

4.1.1. Production of the Definite or Indefinite

In this test we presented RBD patients with short passages that ended with a Noun Phrase (NP) that could be read either as a definite NP or an indefinite NP. In half of the passages the same NP was mentioned before its appearance at the end of the passage. In those passages the NP should have been read as a definite NP because it was already accessible to the reader.

Participants

The participants in this test were 19 RBD patients, 12 aTOMic and 7 TOMer. A group of 10 control patients was also tested. See Table 2 for detailed list of the patients who participated in this task.

Materials

In this test we examine the subjects' ability to distinguish between the conditions at which they should use definite (i.e., ones prefaced by *ha* 'the') or indefinite NPs. We chose to test them on the definite marking of oblique NPs marked by either *be* 'in' or *le* 'to'. Some of the prepositions in Hebrew merge with the following definite article. For example, the combination 'in the' is pronounced as *ba* instead of *be+ha* 'in + the'. The same change occurs for the preposition 'to' *le-*, which is pronounced *la-* when preceding a definite noun. The crucial point for this task is that this phonological change is not reflected in writing, because of the underspecification of vowels in the Hebrew orthography. Thus, while the spelling of *be* ('in' preceding an indefinite) and *ba* ('in the') is identical, their pronunciations are clearly distinct. The same goes for the preposition *le* ('to' preceding an indefinite) and *la* ('to the').

For example:

(1) Yonatan halax le-seret.

Yonatan went to-movie

Yonatan went to a movie.

(2) Yonatan halax la-seret.
Yonatan went to-the-movie
Yonatan went to the movie.

The aim of this study was to test the participant's knowledge of the appropriate use of definite and indefinite noun phrases. We tested this by assessing their ability to select the suitable pronunciation of the preposition before an NP, according to the context in which that NP was presented.

We presented 20 short passages that ended with an NP with a bound preposition (either 'to' or 'in') (i.e., the target NP). In half of the passages the target NP was introduced earlier in the passage, and the mention at the end was its second appearance in the passage (see example 3 below) calling for a definite determiner in the second occurrence. In the other 10 passages the target NP appeared only as the last word in the passage, and this was its first mention (see example 4 below). It was therefore expected to be read as an indefinite.

(3) Maya just finished high school and she is working as a waitress in a café. She hoped to earn more than 200 NIS in a shift but so far she hasn't been earning even 150 NIS. Maya told this to her older sister who works in a fancy restaurant. Her sister tells her: why don't you come to work with me **in the restaurant (ba-misada)**?

(4) Nati loves to eat French cuisine. His mother suggested taking him for dinner for his birthday. When she told him about it he said: That's great! You know I'm always ready to go to eat in **a restaurant (be-mis'ada)**.

As explained above, the appropriate pronunciation of the target NP's prefix in example (3) is with the vowel 'a' (i.e., definite contexts). In example (4) the appropriate pronunciation of the NP's prefix is with the vowel 'e'. The two vowels are easily distinguishable in Hebrew.

Procedure

The passages were typed in bold letters. The last sentence in each passage was underlined. The experimenter and the participant sat next to each other so both could

follow the words as they were read. The experimenter read the beginning of the passage and the participant was asked to read the last sentence in each passage that ended with the target NP. One practice trial was presented at the beginning. The reading was recorded and later transcribed.

Results

Coding: when the participant read the target NP's preposition with the required definiteness marking, according to the context preceding it, the reading was scored as (1). If the participant read the NP with the wrong definiteness marking, it was scored as (0). The score of each participant was the total number of correct readings in each of the two context types.

Pretest: the contexts were read by 30 Hebrew speaking healthy participants (between the ages of 15-65). The contexts in which the target NP's prepositions were read appropriately by less than 90% of the control participants were removed from the final analysis. Two of the *indefinite contexts* were discarded according to this criterion. So in the final analyses there were 8 indefinite contexts and 10 definite contexts.

The average scores of the three groups in the two kinds of contexts appear in the Figure 2 below.

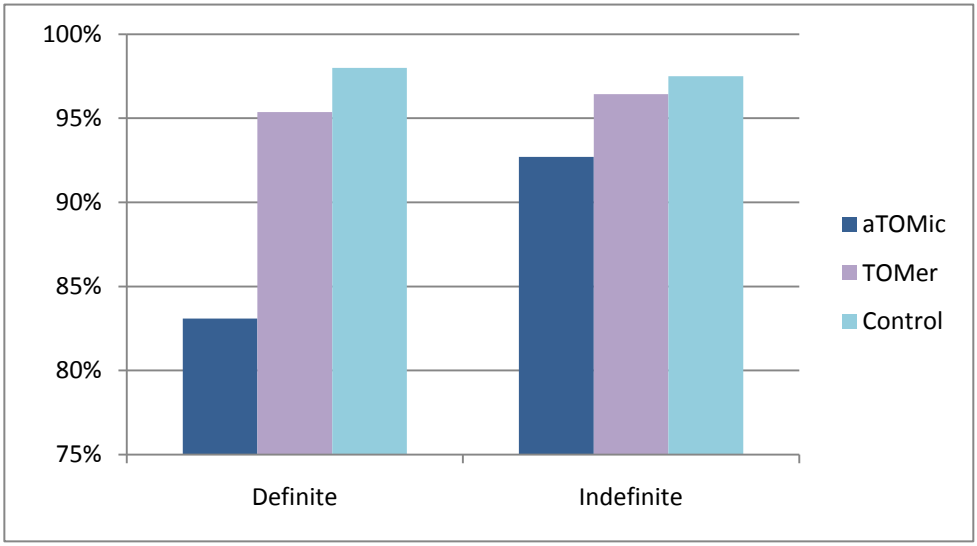

Figure 2. Percentage of correct reading of definite and indefinite preposition prefixes

The results show that the aTOMics scored worse in the definite than in the indefinite contexts. They made more errors of reading the preposition in the definite contexts as indefinite than vice versa. In these errors, they read the prepositions of the target NPs in contexts they were supposed to be read as definite, in the indefinite form, as though they were not introduced to their hearer before. Their reading in the contexts in which the target NP was not mentioned before was appropriate. The two other groups, the TOMers and the controls, read the prepositions appropriately according to whether or not the target NP was mentioned in the context before.

The *total score* of the aTOMic group (average = 86.8%, SD = 9.4%) was significantly poorer than the total score of the TOMer group (average = 94.3%, SD = 4.5%) (U = 61.5, p = .05). The aTOMics also performed significantly poorer than the controls (average = 96.8%, SD = 3.7%) (U = 98, p < .01). The difference between the TOMers and controls was not significant (U = 47, p = .13).

The performance of the aTOMics in the *definite contexts* (the contexts in which the definite reading *ba-* or *la-* was appropriate) (average = 83.1%, SD = 16.5%) was significantly lower than performance of the TOMers (average = 95.4%, SD = 5.9%) (U = 62, p = .05). They also performed significantly lower than controls (average = 98%, SD = 4.2%) (U = 95, p = .01). The difference between the TOMers and controls was not significant (U = 44, p = .2).

When the individual performance on the *definite contexts* was tested, the comparison of each RBD participant to the control group showed that 8 of the 12 aTOMics and 2 of 7 TOMers performed significantly poorer than the control group (p < .05, using Crawford and Howell's, 1998 t-test). For detailed information see Table 6.

In the *indefinite context* (where the indefinite reading *be-* or *le-* was appropriate) the aTOMics average score (92.7%, SD = 11.4%) was not significantly different from that of the TOMers (average = 96.4%, SD = 6.1%) (U = 49.5, p = .28). The aTOMics did not differ significantly from the controls either (average = 97.5%, SD = 7.9) (U = 77.5, p =.13). There was also no significant difference between the TOMers' and controls' average score (U = 40.5, p = .31).

Summary

The results of this production task showed that the aTOMics had a difficulty using the *definite* article (*ha*) in the appropriate contexts. We found that aTOMics tended to read the NPs with an indefinite marker significantly more than TOMers and controls, as though the target NP was not mentioned and introduced before to their listeners. The aTOMics reading of the *indefinite* did not differ significantly from that of the TOMers and controls.

As explained above, when an item is first mentioned it should be read as indefinite[13] (e.g., 'restaurant' in example (2) above). But, its second mention should be with a definite article that implies the speaker is referring to the same item that has already been introduced and their listeners have knowledge of it. The aTOMics showed a lower sensitivity to the state of knowledge of their addresses about whether or not the NP has been previously introduced, compared to the TOMers and controls. They tended to disregard the fact some items were already introduced and repeat the indefinite reading as if they were presenting the NP for the first time. It seems that as speakers they tend to see the shared information as new to their addressees.

The overall score of the aTOMic group was relatively high (86.8%). But note that there are only two possible lexical readings of the letter 'b' or 'l' in the context they were presented – either 'ba-' or 'be-' (or 'le-' 'la-'), so the random chance of reading the proposition correctly is (theoretically) 50%.

It is important to note that the TOMers scored the same as the controls. This finding points specifically to the TOM ability as the link between the linguistic ability and the cognitive deficit. The brain damage itself was not the cause for the lower scores in this task.

[13] Except cases where there is general knowledge of specific and unique NPs like 'the sun'.

4.1.2. Comprehending the use of the definite marker and descriptions in discourse

In this test we asked the participants to judge cases of appropriate and inappropriate use of the definite marker.

Participants

The participants in this test were 18 RBD patients, 11 of them aTOMic and 7 TOMer. A group of 7 control patients were also tested. See Table 2 for detailed list of the patients who participated in this task.

Materials

In order to further test the participant's ability to comprehend which use of the definite article (*ha-*) is appropriate and which is not, we used a judgment task. We tested two conditions: the first was whether they considered if an item was given (i.e., already presented) in the context or not. The second was whether they considered the unique identifiability of the referent, a condition required for definite NPs.

We presented 30 passages that ended with either a definite or an indefinite NP. The participants were asked to judge if the last sentence in the passage was felicitous or not and if it was not to offer a correction. The instructions were *"I will read you short passages. Some of them end with an appropriate sentence and some with an inappropriate sentence. I'm asking you to tell me if the last sentence is appropriate, if it is felicitous. If it's not, offer a correction".*

The test included 16 passages. In 8 of the passages the last NP in the ending sentence was presented as indefinite although that NP had been introduced earlier and was highly salient in the passage. The appropriate response to these sentences was to judge them as infelicitous and the appropriate correction was to add the definite prefix (i.e., infelicitous definite contexts). For example:

(5) Hedva is washing dishes in the kitchen. Her husband is standing in the living room. She asks him 'Is there anything on the table?' her husband answers: 'There is *a plate* with crumbs" Hedva asks him: *bring me a plate*.

In 4 other similar passages the last NP appeared with a definite prefix and the appropriate response was to judge them as felicitous (i.e., felicitous definite context).

In the other 8 passages the last NP in the ending sentence was presented with a definite article but its use was infelicitous because the passage presented more than one salient item of the same type. The appropriate answer was to judge these sentences as infelicitous and the appropriate correction was to add a description that will differentiate between the items (i.e., infelicitous description contexts). For example:

(6) Doron loves animals. He has all kinds of animals in his home. He has a Labrador dog, a Dalmatian and three cats. Today Doron is not feeling well so <u>he asks his son: can you take *the dog* out for a walk?</u>

In 4 other similar passages the last NP appeared with a defining description, the appropriate response was to judge these sentences as felicitous (i.e., felicitous description contexts).

There were also 6 filler passages that ended with felicitous endings. Together there were 14 passages that ended with felicitous sentences and 16 that ended with infelicitous sentences.

In an earlier version, out of the 6 filler passages, 3 were infelicitous due to morpho-syntactic errors (adding a propositional *le / be* in inappropriate conditions).

Procedure

The passages were typed in bold letters. The last sentence in each passage that was supposed to be judged was underlined. The experimenter and the participant sat next to each other so both could follow the passages as they were read. One practice trial was presented at the beginning. The reading was recorded and later transcribed.

Results

Coding: the judgment scores to felicitous and infelicitous sentences were graded as either correct (1) or incorrect (0). The corrections given to sentences judged as infelicitous were graded as appropriate (1) or inappropriate (0). The final score was calculated by multiplying the *judgment score* by the *correction score*. So, a final score

of (1) was given only if the participant judged an infelicitous sentence as such and gave an appropriate correction. The final score for the felicitous sentences was the judgment score.

Figure 3 shows the aTOMics, TOMers, and controls percentage of correct responses to the felicitous and infelicitous passages.

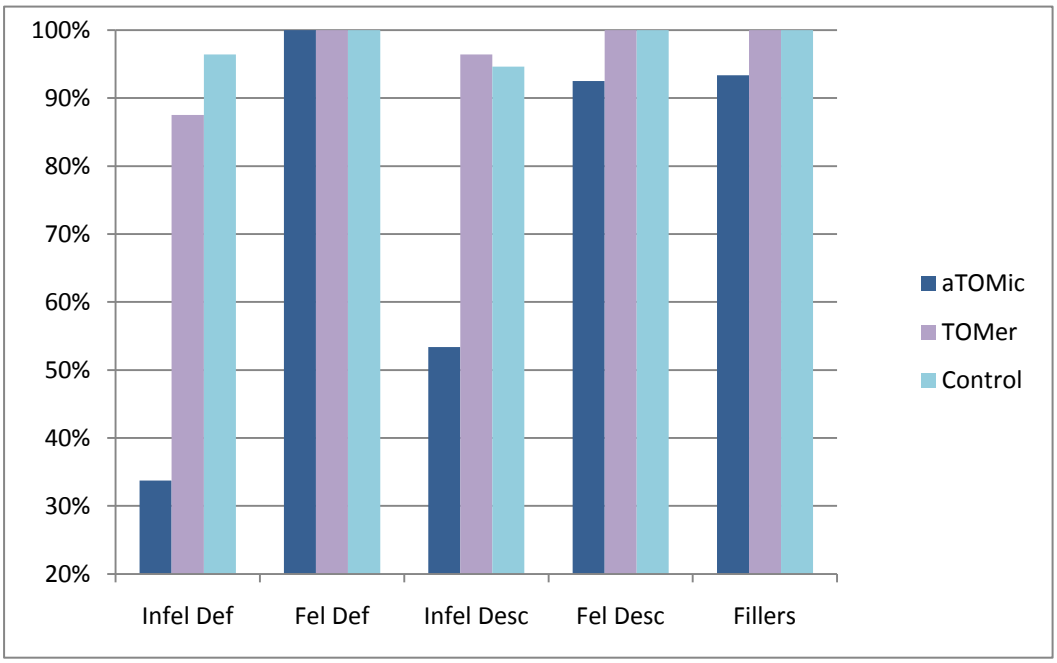

Figure 3. Percentage of correct responses to felicitous and infelicitous use of the definite prefix and descriptions.

The difference between the performance of the aTOMics and the two other groups tested is evident in the *infelicitous* contexts. The aTOMics missed the cases that required a correction. They tended to allow indefinite NPs in contexts where they should be definite, and were not sensitive to the need to add a description that can differentiate between two similar NPs.

The *total score* of the aTOMic group (average = 70%, SD = 15%) was significantly lower than the TOMer group (average = 96%, SD = 5%) ($U = 77$, $p < .001$). They also scored significantly lower than the controls (average = 98%, SD = 3%) ($U = 77$, $p < .001$). There was no significant difference between the TOMers' and controls' total scores ($U = 29.5$, $p = .28$).

In the *infelicitous definite* condition, the pattern of results was similar: the performance of the aTOMics (average = 33.7%, SD = 27.2%) was significantly poorer than that of the TOMers (average = 87.5%, SD = 14.4%) and controls (average = 96.4%, SD = 6.1%) (U = 74.5, $p < .001$; U = 77, $p < .001$ respectively). There was no significant difference between the TOMers and controls (U = 35, $p = .14$).

We also tested the *individual* performance on the infelicitous definite condition. We found that all 11 aTOMics and 2 out of the 7 TOMers performed significantly poorer than the control group ($p < .05$, using Crawford and Howell's, 1998 t-test). For detailed information see Table 6.

In the *felicitous definite* conditions no significant difference was found between the three groups, all scored 100% correct.

In the *infelicitous description* condition the aTOMics (average =53.4%, SD = 31.6%) also scored significantly lower than the TOMers (average = 96.4%, SD = 6.1%) and control participants (average = 94.6%, SD = 9.8%) (U= 67, $p < .01$; U= 66.5, $p < .01$, respectively). No significant difference was found between the TOMers and controls (U = 23.5, $p = .48$).

In comparing the *individual* performance on infelicitous description condition we found that 8 out of 11 aTOMics performed significantly poorer than the control group ($p < .05$, using Crawford and Howell's, 1998 t-test). All the TOMer participants performed as well as the control group. For detailed information see Table 6.

In the *felicitous description* condition no significant difference was found between the performances of the three groups. The aTOMics average score was 92.5% (SD = 11.5%). The TOMers and controls both scored 100% correct. The difference between the aTOMics' and TOMers' scores and that between the aTOMics' and controls' scores was identical (U = 52.5, $p = .11$).

No significant difference was found between the three groups in the *filler condition* also. The average score of the aTOMics was 93.3% (SD = 12.9%). The TOMers and controls both scored 100% correct. The difference between the aTOMics' and TOMers' scores and that between the aTOMics' and controls' scores was identical (U = 49, $p = .18$).

In an earlier version of this task, 6 aTOMic participants answered three filler items that were morpho-syntactically infelicitous, their average score on these items was 93% correct, suggesting that the difficulty discovered cannot be attributed only to the difficulty in answering 'no' in the judgment task.

Summary

The results showed that the aTOMic patients succeeded less than TOMers and controls on this comprehension task. The responses of the aTOMics to the infelicitous conditions indicate that they have a difficulty in identifying the contexts in which the definite article is warranted; when a definite NP denotes a referent that has already been introduced to context. They are also oblivious of the situation where two similar NPs in the context cannot be identified correctly if no further distinguishing description is presented.

In contrast to the aTOMics' low scores in the infelicitous conditions, their performance in the felicitous conditions did not differ from the performance of two other groups. These results point to the aTOMics' inability to consider the state of knowledge of the interlocutors described in the passage regarding the items mentioned. They show a tendency to disregard the need for adding a definite in contexts where both interlocutors know to which item the speaker is referring to (e.g., the/a plate in example 5). They also tend to disregard their hearers' need for a precise description that will allow them to correctly identify a specific item from context (e.g., one of the dogs in example 6).

The high rate of answering 'yes' response cannot explain these results because in an earlier version of this task a sub group of the aTOMics tested, had no difficulty in responding 'no' in cases of ungrammaticality (e.g., morpho-syntactic ungrammaticality).

In sum, it seems that when aTOMic patients encounter a task that demands that they choose an appropriate linguistic token (definite *ha* /or add an appropriate description) in a situation where they have to consider two conflicting viewpoints in a single situation, their performance decreases dramatically. Their basic difficulty in representing the content of the other person's state of knowledge is demonstrated here

in an inability to appropriately judge the use of the definite prefix or the need for a description. In this task too, the TOMers scored as well as controls, showing again the importance of defining the exact difficulty different RBD patients suffer.

4.1.3. Appropriate use of the definite article (*ha-*): The Grumble test

Participants

The participants in this test were 14 RBD patients, 8 of them aTOMic and 6 TOMer. A group of 8 control participants were also tested. See Table 2 for detailed list of the patients that participated in this task.

Materials

The aim of this task was to test whether the participants were sensitive to the use of definite article in an inappropriate context. According to Bertrand Russell's classic work on denoting phrases (Russell, 1905), the definite form presupposes uniqueness, i.e., the existence of one and only one entity that meets the descriptive content of the NP. Therefore, the use of this form when two similar items are presented is a breach of this presupposition and should not be accepted by the participants.

We presented 12 pictures, in each there were three figures: two figures of the same kind (e.g., two elephants) and a third, different one (e.g., a boy). See example bellow (Fig. 4). We asked the participants to point to a referent or an item in the picture according to the experimenter's instruction. There were two target pictures for which we asked to point to a referent that appeared twice, using a definite NP (which breached the uniqueness presupposition). For example, while presenting figure 4 we asked *'show me <u>the elephant</u>'*.

Figure 4. An example for a picture presented in the target items.

Our test was whether the participant would react in some way that would show she cannot cooperate with the experimenter's request. Any kind of grumble, refute, disapproval, or dissatisfaction was accepted as an appropriate reaction.

For the other 10 pictures presented, we gave 4 kinds of instructions:

(a) To show *the item* when only <u>one item</u> of that kind appeared in the picture.

(b) To show *an item* when <u>no item</u> of that kind appeared in the picture.

(c) To show *an item* when <u>one item</u> of that kind appeared in the picture.

(d) To show *an item* when <u>two items</u> of that kind appeared in the picture.

Instruction (b) was important in order to verify the participants were willing and able to react in a way that would show they disapprove of the instruction given by the experimenter.

Procedure

The items were shown in the same order to all participants. The instructions were given in the following order:

a, c, b, a, target, c, a, target, c, a, c, d.

Four of the participants answered an earlier version of this task in which more items of each kind were presented.

Results

Coding: correct responses (1) to the target items were any signs of dissatisfaction from the instruction posed.

The average correct responses to the target pictures are shown in Figure 5.

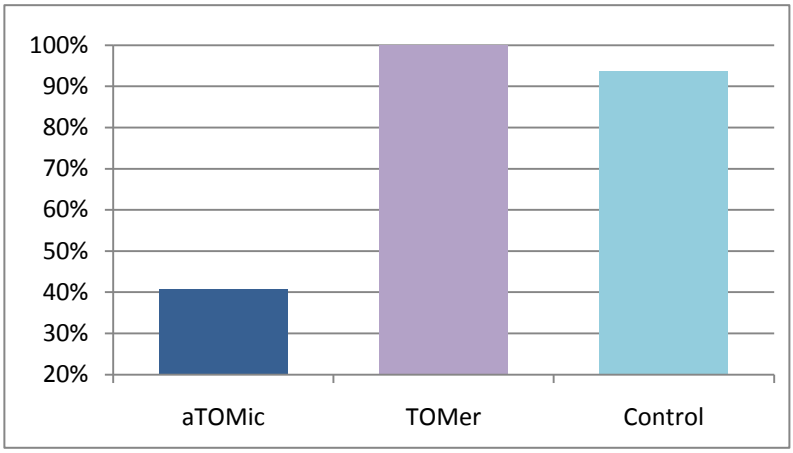

Figure 5. Percentage of correct responses to the target items

The results show that the aTOMics failed to grumble or show any disapproval reaction to the inappropriate request they were asked. Their average score (40.6% SD = 35.2%) was significantly lower that the TOMers (100%) and controls (93.8, SD = 17.7%) (U = 45, $p < .01$; U = 57.5, $p < .01$, respectively). There was no difference between the TOMers and controls (U = 21, $p = .37$).

In comparing the individual performance of the RBD patients we found that 6 out of 8 aTOMics performed significantly poorer than the control group ($p < .05$, using Crawford and Howell's, 1998 t-test). All the TOMer participants performed as well as the control group. For detailed information see Table 6.

These results cannot be explained by a general reluctance to demonstrate dissatisfaction, because the aTOMics did express dissatisfaction when asked to point to an item that was not in the picture (condition b) (average = 95.8%, SD = 11.8%). Only one participant pointed one time to some item in the picture although he was asked to point to an item that did not appear at all. (The TOMers and controls did not make any such mistake, 100% correct for both).

Summary

The results of this task show that the aTOMics lose some of their sensitivity to the appropriate context of the definite marker even when they are presented with isolated sentences detached from prior discourse, and the discourse conditions that require the definitness are the pictures in front of them.

This loss of sensitivity clearly demonstrates that linguistic abilities are affected by loss of TOM. Even when the context the aTOMics needed to consider was the experimental setting itself (and not an imagenery situation) and their interlocutor was the experimenter, they showed difficulty in identfying the appropriate context for using the definite marker.

4.1.4. Interim Summary and Discussion: Definiteness and aTOMia

We introduced three tasks to test the linguistic performance of RBD patients regarding the appropriate use of definiteness in discursive contexts.

We found that the two groups of RBD patients identified in chapter 3, as aTOMic and TOMer, scored significantly different in all three tasks. The aTOMics scored significantly lower than the TOMers and controls in all three, whereas the TOMers scored the same as the control group in all tasks and conditions. In all the tasks a test of individuall performance showed that most of the aTOMics and only a minority of the TOMers scored significantly different from the control group.

In the *production task* we found that, although the aTOMics received relatively high scores (over 83% correct), their scores were significantly lower than those of the two other groups and they showed a specific pattern of errors. Their difficulty was in producing the *definite reading* of the prepositional phrase. This finding can be explained by their disregard of the state of knowledge of their hearers. They preferred to use the indefinite reading as though the item they were referring to has not been mentioned previously and was not known to their hearers.

In the *judgment task* we found the aTOMics showed the same difficulty. They failed to judge if there was a need to add the definite marking to items that were introduced into context, and therefore were known to both interlocutors presented in the

passages. In this task we also saw that the same disregard to the state of knowledge of two interlocutours was evident in cases where they failed to add a description when needed. In the *infelicitous description contexts* the hearers in each situation needed a description in order to appropriately choose one of two similar items. The aTOMics failed in identifying this need. The TOMers and controls did not.

In the *grumble test* we found again that the aTOMics have difficulty in comprehending the appropriate use of the definite article. We asked the participants to point to some item, introduced with the definite article, in a picture. Because that item appeared twice, our instruction was supposed to elicit some kind of a grumbling reaction. The aTOMics failed to criticize our inappropriate instruction, while the TOMers and controls did.

In sum, the results of three tasks support one and the same conclusion: the decrease in aTOMic reasoning detected in the TOM battery affects the linguistic ability of aTOMic patients. They fail to appropriately use the definite article to mark items which are identifiable to their hearers and which are brand new or otherwise unidentifiable.

4.2. High or Low accessibility? The Use of Different Referring Expressions in Discourse

4.2.1. Production

We used two tasks to test the RBD patients' ability to use properly different kinds of referring expressions according to the accessibility of the different items in the mind of their addressees (see section 1.2.). In the first task we asked the participants to retell two short stories and we analyzed their use of high and low accessibility markers in the text they produced. The second task was an elicitation task. We presented pictures depicting three characters, two of them were similar and we asked a Wh- question which was designed to elicit a description of one of two similar figures. We tested whether the participants gave the relevant information that would enable the experimenter, as the addressee, to differentiate between the similar figures.

4.2.1.1. Retelling stories

Participants

The participants in the retelling task were 16 RBD patients, 9 aTOMics and the 7 TOMers. A group of 8 control participants also retold the stories. See Table 2 for detailed list of the patients who participated in this task.

Materials

We presented two stories (adapted from Pimkel, 2006), each story presented two characters and an interaction between them. The first, the 'Brothers story' was about two brothers who tried to help each other by giving, secretly, one to the other, a bigger portion of the wheat they harvested together. The second story, the 'Fishermen story' was about two fishermen who regularly brought unequal catch from their fishing trips. The one who caught fewer fish each time tried to improve his catch in different ways, but did not succeed.

Procedure

The task was described to the participants before the stories were read. They were told the experimenter will read them a short story and later they will be asked to retell the story to a recording for someone who has not yet heard it. For the RBD participants,

each story was read in a different session. At the end of the reading they were asked if they wanted to read it again by themselves. When they were ready, the page was put aside and the participants retold the story to the tape-recorder.

Results

Coding: First, we counted the number of all referring expressions (full NP+ description, proper names, demonstratives, pronouns, and null NPs).

Next we conducted two kinds of analyses: for the first, *'with corrections'*, we marked each referring expression that was inappropriate and then we corrected it to an appropriate marker and continued coding the text.

In the second analysis: *'without corrections'*, the inappropriate reference were marked and then we coded the following expressions while taking into account that the earlier reference assignment was unsuccessful or not clear.

The difference between these two methods was crucial to cases like the ones demonstrated below (7):

(7) *The fishermen returned to shore and then <u>he</u> told <u>him</u>.*

In this case the use of the high accessibility marker 'he' is inappropriate because there were two fishermen and we can't tell which of the two the pronoun refers to. The question is how to treat the second expression, 'him'. During on line listening, in natural conversation, there is no way of understanding who the narrator is referring to with this second expression if the earlier reference assignment of 'he' was not successful. Note, that in each story only two characters were presented so if the first referring expression was appropriate a listener could infer to whom the second expression is referring to according to the syntax of the sentence alone. Thus, the *'without correction'* analysis was conducted in order to simulate as closely as possible the real situation of a listener trying to comprehend a story told. When one reference is not clear it is harder to continue to build the mental picture of the situation.

The *'with correction'* analysis was conducted in order to try to balance between two research interests: one is the need to consider the real life situation in which conversations take place. The second is the need to describe precisely the patients' ability to use reference terms appropriately without "penalizing" them by counting

one inappropriate use as a continuing default. A second judge identified and categorized the referring expressions in 18 out of the 50 texts elicited. There was agreement on approximately 95% of the scores. Differences were resolved in discussion.

First we calculated the percent of referring expressions out of the total number of words used by each group. We found that, although the average number of words used by the different groups was different and varied, in all groups, approximately 20% were referring expression. These results are presented in Table 10.

Table 10. Average number and percentage of referring expressions used by aTOMics, TOMers and control participants (SD).

	Average number of words	Average number of referring expressions	Percent of referring expressions
aTOMic	211.8 (82.9)	42.9 (16.2)	20.4 (2.3)
TOMer	338.6 (101.4)	69 (19.8)	20.5 (2.2)
Control	321 (67.9)	60.3 (12.7)	19 (2.9)

The results of the two analyses for both stories are presented in Figure 6. The results show that the aTOMic group use less appropriate referring expressions than the other two groups tested.

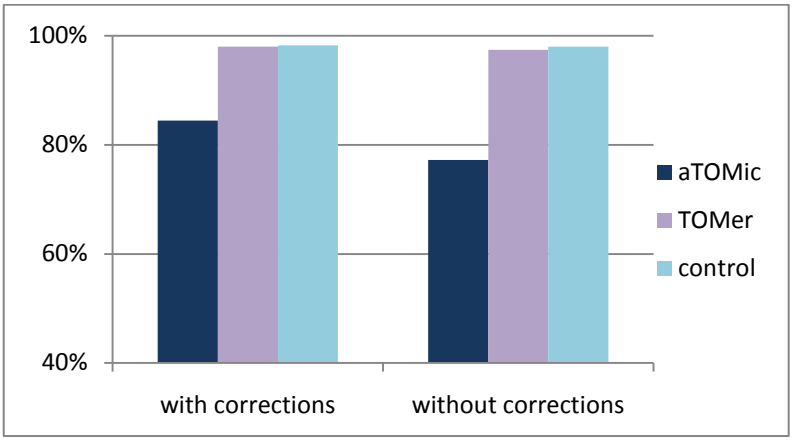

Figure 6. Percentage of correct referring expression according to the two analyses: *with corrections* and *without corrections*

Significant differences were found in both types of analysis. In the analysis *with correction*: the aTOMics (average = 84.5%, SD = 8.7%) scored significantly lower than the TOMers (average = 98%, SD = 1.3%) and controls (average = 98.2%, SD = 2.3) (U = 63, $p < .001$; U = 70, $p = .001$, respectively). The difference between the TOMers and controls was not significant (U = 29, $p = .3$).

In the *without corrections* analysis the same pattern of results was found. The aTOMics (average = 72.2%, SD = 14%) scored significantly lower than the TOMers (average = 97.4%, SD = 2.6%) and controls (average = 98%, SD = 2.4) (U = 61, $p = .001$; U = 61, $p = .001$, respectively). The difference between the TOMers and controls was not significant (U = 26.5, $p = .42$).

We also compared the individual performance of each participant to the performance of the control group. In the analysis *with corrections* each of the the aTOMic participants performed significantly lower than the control group ($p < .05$, using Crawford and Howell's, 1998 t-test). All the TOMer participants performed as well as the control group. In the analysis *without corrections* 7 out of the 9 aTOMics performed significantly lower than the control group ($p < .05$, using Crawford and Howell's, 1998 t-test), and again all the participants in the TOMer group performed as well as the control group. For detailed information see Table 6.

Another question we tried to answer was - In what way are aTOMics different from the other two groups in the use of referring expressions? Do they tend to use more high accessibility markers as if the information they have is common to them and their listeners? Do they assume their listeners share the same mental model of the situation they hold and therefore there is less need for lower accessibility markers? Or, do they use more low accessibility markers as if they assume their addressees are not attending and activating the different items in the context?

To answer this question we calculated the percent of low and high accessibility markers and compared the rate of each kind of referring expression used by the different groups.

We found that in the aTOMic group, 44.4% (SD = 10.2%) of the referring expressions used were *low accessibility markers* (those used to introduce new referents into discourse). In the TOMers group these expressions were 45.7% (SD = 5.2%) of the

total referring expressions they used, and for the controls it was 48.4% (SD = 11%) of the total referring expressions. The differences between the groups in the rate they used these expressions was not significant, between the aTOMics and TOMers, U = 35.5, p = .35; between the aTOMics and controls U = 38.5, p = .25; and between the TOMers and controls U = 28, p = .35. Accordingly the rate of high accessibility markers was also very similar between the groups (aTOMics: average = 55%, SD = 10.9%, TOMers: average 54.3%, SD = 5.2% and controls: average 51.6%, SD = 11%). No significant differences were found between the groups (between the aTOMics and TOMers U = 27.5, p = .35; between the aTOMics and controls U = 24, p = .23; and between the TOMers and controls, U = 21, p = .35).

Another important finding is the type of errors made by the participants. We calculated the number of errors each participant made (from the analysis without correction). The average number of referring expressions used by the aTOMics was 386, 77 (19.9%) were inappropriate. Most of the errors (63, 81.8% of the total errors) were cases in which the high accessibility marker was used instead of a lower, more informative, referring expression. The TOMers made only 12 (2.5%) errors out of a total 483 referring expressions. Eleven of these mistakes were of using inappropriately a high accessible term. The participants in the control group made 7 errors out of 422 referential expressions, 1.7%. All these errors were of the same kind again, overusing the high accessibility marker.

These findings show that aTOMics erred much more than the other two groups and that their common error was treating information as known to their addressees when in fact the use of a more informative, low accessibility marker would have been appropriate.

These results should be considered as preliminary because the texts used were very stringent, in using only two characters in each story. Maybe if the story presented one main character and/or additional characters were included, the pattern of using reference terms would have been different.

Summary

This task tested the participants' *production* of referring expressions by analyzing the texts they produced in a retelling story task. The results showed that the aTOMics

produced significantly more inappropriate referring expressions compared to TOMers and controls, in two kinds of analyses. Although the rate of the referring expressions produced by the three groups, out of the total number of words, was very similar.

We also tried to answer the following question - In what way the groups differ? Do aTOMics presume more mutual knowledge than is actually present? Do they tend to use more high accessibility markers or do they use more low accessibility markers, as if their listeners are not creating an active mental model of the story being told? The results are only preliminary but they show that the aTOMics use the same rate of low and high accessibility markers but they err more. They mostly overuse the high accessibility markers when lower, more informative markers should be used.

4.2.1.2. Production of appropriate differentiating descriptions

The aim of this task was to test whether the participants were able to choose a relevant feature which distinguishes efficiently between two similar characters while answering to the Wh question (*which character of the two?*), by which their hearer could identify the character they were referring to.

Participants

The participants in this task were 17 RBD patients, 10 were aTOMic and 7 were TOMer. A group of 7 control participants was also tested. See Table 2 for detailed list of the patients who participated in this task.

Materials

Twenty pictures were presented, each picture included three figures: two of the same type which differed in at least one feature (e.g., two giraffes, one smaller than the other) and a third figure of a different kind. In each picture, the first figure was performing an action on the second, and the second figure was performing the same action on the third figure, which was of the same type of the first one (see Fig. 7). We asked two Wh- questions about each picture, *which* subject (8) and *which* object (9).

(8) Eize girafa modedet et ha-yalda?
Which giraffe measures ACC the-girl?
Which giraffe is measuring the girl?

(9) Et Eize girafa ha-yalda modedet?
ACC which giraffe the-girl measures?
Which giraffe the girl is measuring?

Figure 7. An example of the items presented.

The important point was whether the participants would offer a feature that differentiates between the similar figures (e.g., their height or size), or some other feature, which they have in common (e.g., their color). If they chose a description that does not distinguish the two referents, or don't add a description at all, it would show that they are not considering which feature is pertinent for the addressee to zero in on the intended referent.

Procedure

We presented the task in two sessions. In the first, 20 pictures were presented. When presenting 10 of the pictures we asked *which- subject* questions, and to the other 10, we presented *which- object* questions. In the second session, the same 20 pictures were presented again and the alternate question was asked. The participants were asked to answer the question and not to point at the characters.

Results

Coding: if the participant mentioned a feature that differentiates between the two similar characters the answer was graded (1) correct, and if s/he mentioned a feature that does not differentiate or if he did not mention any feature it was graded (0) incorrect.

For example one of the aTOMic patients answered the question in (7) above with an ill defining feature:

(10) *Abraham*: which giraffe? The yellow giraffe with the brown spots.

The three groups scored high scores but the aTOMics erred more than the two other groups, showing they did not always consider which feature is relevant to successful retrieval of the correct referent. Figure 8 shows the average result of the three groups.

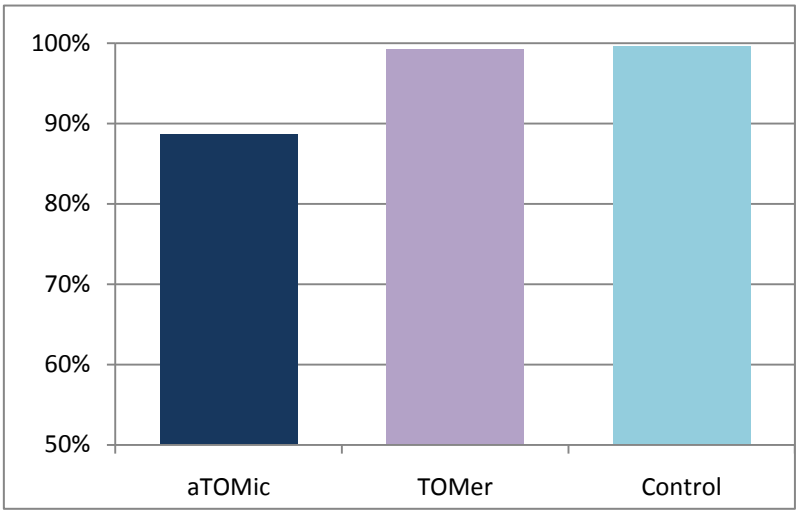

Figure 8. Percentage of correct responses of the three groups tested

The average score of the aTOMic group (average = 88.6%, SD = 9.8%) was significantly lower than the TOMers (average = 99.3%, SD = 1.9%) and controls (average = 99.6%, SD = 0.9%) (U = 62, p = .005, for both), there was no difference between the TOMers and control group (U = 25, p = .5). In comparing the individual performance of each RBD patient to the control group we found than 8 out of 10 aTOMics and 1 out of seven TOMers performed significantly poorer than the control

group ($p < .05$, using Crawford and Howell's, 1998 t-test). See Table 6 for detailed information.

Summary

In this task the aTOMics showed again insensitivity to the discourse situation, despite the fact that in this task they were not presented with an interaction between two characters, rather, they were active participants. The task required them to provide the relevant feature to answer the Wh-question.

Although they succeeded in most trials and their average score was above 85% correct, they still erred significantly more than TOMers and controls. Showing again the same insensitivity as speakers to provide the information that fits the accessibility of the item they are referring to in the minds of their addressees.

4.2.2. Metalinguistic tasks: Choosing the appropriate referent or reference term

In the tests we present below we asked the participants to judge which of two reference terms is more appropriate, or which of two characters is the intended antecedent of various reference terms.

In some of the tasks the options appeared after a preceding context was presented and in others the choices were presented within one sentence. In some of the tasks the participants were asked to justify their choice, in others we asked only for their preference.

In each task we presented a contrast between two options. We contrasted between referring expressions that signal high accessibility to those that signal low accessibility and intermediate accessibility rates. We also presented contrasts within the low accessibility markers between more or less informative phrases. The range of tasks and the variety of contrasts we tested were designed to present a detailed picture of the RBD patients' ability to use referential terms in discourse.

4.2.2.1. Choosing between antecedents for high accessibility terms *'hu/hi'* ('him/her') and intermediate accessibility term *'ze/ zot'* ('that masculine/that feminine')

Participants

The participants in this test were 21 RBD patients, 14 aTOMics and 7 TOMers. A group of 8 control participants was also tested. See Table 2 for the list of the patients who participated in this task.

Materials

In this task we tested the participants' ability to differentiate between two kinds of referential markers as hearers. We presented items composed of two sentences. In the first, two characters were introduced (see example 10 below). In the second sentence a reference term was used and the participants were asked to choose which of the two characters mentioned in the first sentence the term refers to. The referring expressions were either high accessibility markers (see 11a below), or intermediate accessibility marker (see 11b below). A correct response would be if the participant chose the first mentioned person when the reference term was a high accessibility marker (e.g., *'hu/hi'*) and the second, when an intermediate reference term was presented (e.g., *ze/ zot*).

(11) Itay melamed et Yonatan geografia.

Itay teaches Jonathan geography…

a. …*Hu* hevi sefer xadash la-shiur.
… *He* brought a new book to class.

b. …*Ze* hevi sefer xadash la-shiur.
…*That* brought a new book to class.

In all the items the two characters presented in the first sentences were always of the same gender in order to preclude a grammatical cue. The task included 8 items and 8 fillers. The fillers and the items were composed in a similar way. In the first sentence

we presented two or more characters and in the second sentence we referred to one of them and asked to which character the speaker is referring? For example (12) below:

(12) Dana called Miriam and asked her to prepare for a field trip, sandwiches for both of them. Who will bring the sandwiches?

Procedure

The items were presented in a semi-randomized order. The experimenter read each item in normal intonation and pace, and reread them if necessary. The participants were encouraged to follow the written items while they were read.

Results

Coding: We asked which of two characters introduced in the first sentence the speaker is referring to in the second sentence? The appropriate response was to choose the first mentioned character if a high accessibility marker was used and to choose the second mentioned character if the intermediate marker was used. An appropriate answer was scored as correct (1) and if the other character was chosen or if the participant said it could mean both, their answer was coded as incorrect (0). The answer that it could mean both was given only twice (out of 451 responses in all three groups), once by an aTOMic patient and once by a TOMer patient.

Pertest: The task was presented in a written form to 11 normal controls (between the ages of 19-24). Items, on which less than 85% of all control participants scored correct, were discarded from the final analyses. Four items were discarded due to this criterion. In one item the target answer was the first NP, in 2 the target NP was the second NP, and one was a filler item.

The average score of the aTOMics (average = 84.1%, SD = 12%) was significantly lower than the TOMers' score (average = 95.2%, SD = 4.5%) and controls (average = 95.2%, SD = 4.5%) ($U = 81.5$, $p < .01$; $U = 79.5$, $p = .01$, respectively). No significant difference was found between the TOMers and control participants ($U = 24.5$, $p = .48$). The scores of the three groups in identifying each of the two referents and the filler items are presented in Figure 9.

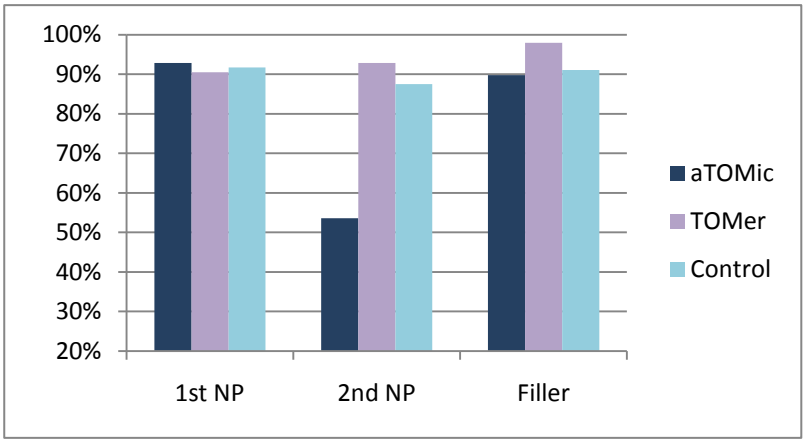

Figure 9. Percentage of correct responses of identifying 1st and 2nd NP mentioned in the sentence and filler items

The results show that the responses to the two NPs were different in the two groups; the aTOMics identified correctly the first NP mentioned in the sentences, as did the TOMer group. But in identifying the second mentioned NP, by the intermediate level accessibility term *ze*, the aTOMics scored significantly lower than the TOMers.

For the *first NP* mentioned the aTOMics average correct response was 92.9% (SD = 14.2%). Not significantly different from the TOMer (90.5%, SD = 16.3%) or the control group average score (95.2%, SD = 12.6%) (U = 45.5, p = .41; U = 52.5, p = .41, respectively). There was no significant difference between the TOMers and controls either (U = 28, p = .35).

In identifying the *second NP* there were significant differences. The aTOMics scores (53.6%, SD = 41.4%) were significantly lower than the TOMers' (92.9%, SD = 18.9%) and controls' (100%) (U = 75.5, $p < .05$; U = 80.5, p = .01, respectively). There were no significant differences between the TOMers and controls (U = 28, p = .35).

We also tested the individual performance on identifying the *second NP*. We found that 9 out of 14 aTOMics and 1 out of the 7 TOMers performed significantly poorer than the control group ($p < .05$, using Crawford and Howell's, 1998 t-test). For detailed information see Table 6.

The aTOMics score on the filler items (average = 89.8%, SD = 11.8%) was not different from the TOMers (average = 98%, SD = 5.4%) or the controls (average = 93.9%, SD = 7.6%) (U = 68, p = .08; U = 57, p = .28). There was no significant difference between the TOMers and control in this comparison too (U = 17.5, p = .2).

Summary

The participants in this task were asked to choose which of two antecedents the speaker is referring to with a certain referring expression. One term indicated a high degree of mental accessibility for the referent in the listener's mind. A speaker was expected to use this kind of term when referring to the first NP mentioned in the preceding sentence. The other was an intermediate accessibility marker; a speaker was expected to use this kind of term when referring to the second NP mentioned in the preceding sentence. By using a reference term which was not a high accessibility term (e.g., not a pronoun) the speaker signals to her hearer not to access the most highly accessible referent, but rather access a less accessible one. We found that the aTOMics were not sensitive to this difference, and did not take the change from the frequent pronoun (*hu*) to the less frequent one (*ze*) to signal a change in reference.

The results showed that the *total average score* of the aTOMic group was significantly lower than that of the TOMers and controls. But, there were differences in the way the aTOMics were able to identify the appropriate referent depending on the type of referring expression presented. All three groups correctly understood that when a high accessibility marker was used (i.e., a pronoun), the speaker w referring to the first NP in the sentence. But, a significant difference was found between the two RBD groups and between the aTOMics and controls in understanding the use of the *intermediate accessibility* marker[14]. The aTOMics missed the cue of using an irregular pronoun which indicates a change of reference, and said that the intermediate accessibility marker refers to the first NP, and not to the second NP, significantly more often than the other two groups.

An explanation for these results might be that when using pronouns the participants are activating a cognitive mechanism for identifying the NP which is first mentioned

[14] It should be noted that the intermediate accessibility marker (e.g., '*ze*', '*zot*') is much less common in spoken Hebrew than the high accessibility marker, and it is usually considered to signal a high register.

and is the agent in the sentence. The aTOMics performed very wellon this task but, identifying who had been the antecedent of the intermediate reference term was harder for them. We suggest that the source of this difficulty is TOM-related. The participants might be relying on TOM for understanding that the reason the speaker is using this term is to stress that she is referring to some other antecedent in the sentence, and not the most salient one. The results show that in this condition the aTOMics tended to choose the more prominent, first mentioned antecedent instead of the less salient second mentioned NP, as though they were not able to infer the reason for using the more uncommon marker.

In sum, the results of this task show that the aTOMics are less sensitive to the difference between the two kinds of referring expressions we presented. These results demonstrate again the relation between TOM and the specific linguistic ability of using referential terms.

4.2.2.2. Choosing between high accessibility markers (*hu/hi*) and low accessibility markers (recurrence of proper names)

Participants

The participants in this test were 20 RBD patients, 13 aTOMics and 7 TOMers. A group of 9 control participants was also tested. See Table 2 for detailed list of the patients who participated in this task.

Materials

This task tested whether participants appropriately used high accessibility markers (pronoun) and low accessibility markers (a proper name) according to the context they appear in. In each item two sentences were presented, in one the second occurrence of the same referent was marked by the same proper name, and in the other the first mention was of a proper name and the second occurrence was a pronoun. For example:

(13) *Zvi* nasa le-tiul ve-*Zvi* cilem harbe tmunot
Zvi went to-a-trip and-*Zvi* photographed many pictures
Zvi went on a trip and *Zvi* took many pictures.

(14) *Zvi* nasa le-tiul ve-*hu* cilem harbe tmunot
Zvi went to-a-trip and-*he* photographed many pictures
Zvi went on a trip and *he* took many pictures.

We asked the participants: *which sentence sounds more felicitous, sounds better, more natural in conversation*? The correct response, according to Ariel's accessibility theory (1990; 2001), would be to choose the sentence with the pronoun. This task questions the participants' meta-linguistic abilities, it demands probing of one's intuitions about their knowledge of their language, in other words, make a judgment about one's own linguistic knowledge.

The test included 10 test items and 5 fillers. The test items included 5 coordinated sentences and 5 embedded sentences. In the filler items the participants were asked to choose between semantically felicitous or infelicitous sentences (see 15 below) or between grammatical and ungrammatical sentences (see 16 below).

(15) a. I love to drive the car while my father is driving.

b. *I love to ride the car while my father is driving.*

(16) a. *Giora thinks Yifat likes him.*

b. Giora thinks Yifat likes himself.

Procedure

The test items were presented in a semi-randomized order after two practice items. The experimenter read each item in normal intonation and pace, and re-read them if necessary. The participants were encouraged to follow the items while the experimenter read them.

Results

Coding: An answer was coded correct (1) if the participant chose the sentence with the pronoun, and incorrect (0) if s/he chose the other option or if s/he said there was

no difference between the sentences. There were 12 responses of 'no difference' out of all responses given (207 in the aTOMic group, 461 in all groups together), all of them given by a-TOMic patients. The percent of correct answers in the three groups are shown in Figure 10.

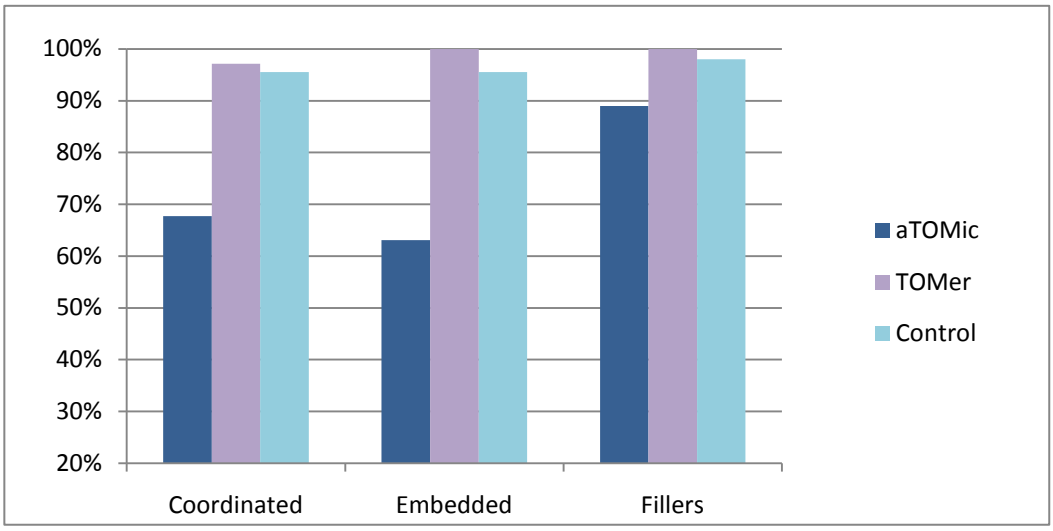

Figure 10. Percentage of correct answers of choosing the appropriate high accessibility marker

The results show that the aTOMics tended to prefer, more than the other groups, the sentences that repeated the same low accessibility marker (proper name) rather than the sentences that correctly used a high accessibility marker in the second appearance of the same reference (pronoun).

In the *coordinated sentences*, the aTOMic group scored (average = 67.7%, SD = 15.4%) significantly poorer than the TOMers (average = 97.1%, SD = 7.6%), and significantly poorer than the controls (average = 95.6%, SD = 8.8%) ($U = 85$, $p < .01$; $U = 107.5$, $p < .001$, respectively). There was no significant difference between the TOMer group and the control group ($U = 29$, $p = .42$). We also compared the individual performance of each aTOMic participant to the performance of the control group. We found that 8 out of 13 aTOMics performed significantly lower than the control group ($p < .05$, using Crawford and Howell's, 1998 t-test). See Table 6 for detailed information.

In the *embedded sentences* we found the same pattern of results: The aTOMic group scored (average = 63.1%, SD = 24.3%) significantly poorer than the TOMers (average

= 100%) and the controls (average = 95.6%, SD = 8.8%) (U = 87.5, p = .005; U = 106.5, p < .001, respectively). No difference was found between the TOMers and controls (U = 24.5, p = .25). In this condition 7 out of 13 aTOMics performed significantly lower than the control group (p < .05, using Crawford and Howell's, 1998 t-test). See Table 6 for detailed information.

In the *filler items* the aTOMics scored an average of 89.2% (SD = 15.5%), the TOMers scored 100% correct. The difference between the two groups was not significant (U = 63, p = .09). Controls scored an average of 97.8% (SD = 6.7). The differences between the aTOMics' and controls', and between the TOMers' and controls' were not significant either (U = 75.5, p = .13; U = 28, p = .37, respectively).

Summary

This task aimed to test the appropriate use of high accessibility markers. We presented two sentences and asked the participants to choose which of two sentences is more felicitous, the one in which the same referent is referred to twice by a proper name or the one in which it is referred to in the first mention by a proper name, and in the next appearance it is referred to by a pronoun (a high accessibility marker). The correct choice for all the items was the sentence that included a pronoun instead of a second mention of a proper name.

In this task too, significant differences were found between the aTOMics and the other two groups, the TOMers and the controls. The aTOMics preferred sentences in which two proper names appeared more than the other groups. This result of this task shows that the aTOMics prefer "not to trust" the mutual knowledge gathered during conversation.

Usually, a speaker uses a pronoun, a mark that only includes the gender and number features of the referent, if she takes into account that the referent is highly accessible to her addressee. She does not need to add more than that to ensure proper identification of the referent. The aTOMic patients, when taking the role of speakers, it seems, are not accurate enough in assessing the accessibility of the referent in the minds of their addressees and therefore, sometimes, choose to repeat, unnecessarily, the proper name of the referent. But note that in this task the only error possible is the one of repeating the proper name. It could be that the aTOMics have a general

difficulty in implementing their linguistic knowledge about all kinds of reference terms. In this task this difficulty could only be expressed in overusing the low accessibility term.

In the next task we presented two opposing situations, one which called for more informative reference term and the other which called for the use of a less informative reference term.

4.2.2.3. Choosing between two kinds of low accessibility markers: Less or more informative

Participants

The participants in this task were 17 RBD patients, 10 were aTOMic and 7 were TOMer. A group of 9 control participants was also tested. See Table 2 for detailed information about patients' participation.

Materials

The purpose of this task was to test whether the participants were sensitive to the common knowledge created during conversation by interlocutors. We presented two kinds of contexts in which two people were engaged in a conversation and asked participants to choose the appropriate way to continue the conversation. One option was a phrase containing more information (e.g., a proper name and a description) the other contained less information (e.g., a proper name alone). In one kind of context the characters were well-acquainted with each other and in the other they were not. The participants had to choose whether the information in the option they choose will be necessary for the figures described in the context, or will this information be superfluous.

The test included 16 items. For 6 of the items, the context introduced characters who were familiar with each other (see 17 below), so the appropriate continuation was the less informative option. For 6 other items (see 18 bellow), where the context introduced two figures that were not acquainted with each other, the more informative option was the correct continuation. Four other contexts served as fillers.

(17) Yifat and Nurit are sisters and they fight with each other quite often. Today Yifat annoyed her sister. Nurit went to their mother and said:

a. Mom, Yifat is annoying me.

b. Mom, Yifat *my sister* is annoying me.

(18) Efrayim is a garage owner and he is introducing the working procedures and people working in his garage to a new employee. How will he introduce the accountant?

a. Shmuel Weisman *is the garage's accountant*; he comes in every Wednesday and works until 4 in the afternoon.

b. Shmuel Weisman comes in every Wednesday and works until 4 in the afternoon.

Procedure

The items were presented in a semi-randomized order after one practice item. Participants were asked to choose one option and to justify their choice. The items were read by the experimenter and repeated if necessary.

Results

Coding: the answers were coded as correct (1) if the participant chose the appropriate answer or (0) if not. The *justifications* were coded as correct (1) if they included some mention of the common knowledge or relations between the characters presented, and (0) if an irrelevant justification was given.

The final score was calculated by multiplying the answer score with the justification score. This way, only correct answers with appropriate justification were given a final score (1). If participants either chose the inappropriate answer or did not justify it appropriately their final score was (0).

Below are examples of justifications given in answer to items (17) presented above. Dafna chose the option in which the sister is mentioned only by her name and justified it appropriately (19). Jacob on the other hand chose the option in which the superfluous description is added and justified his choice with an inappropriate justification (20). In reply to item (18) presented above both Sigalit and Dror chose

the appropriate, more informative option, but Sigalit justified her choice appropriately (21) and Dror did not (22).

(19) *Dafna*: *she (=the mother) knows she's her sister!*

(20) *Jacob* : *Because it's her sister.*

(21) *Sigalit*: *It's a new employee, he doesn't know who Shmeuel is and what's his job.*

(22) *Dror*: *He's emphasizing he works until 4.*

Pretest: the task was presented in a written form to a group of 10 healthy adults (between the ages of 20-24) in order to test if our choices of appropriate answers were not biased. The average total score was 99.6% (SD = 1.3%). Before summing the results we checked whether all items received at least 85% correct responses by the collected group of controls (9 who answered the written version in the pretest and 9 in an oral presentation that participated in the test). We found one filler item that did not meet this criterion. It was discarded from further analyses.

The responses of the three groups to the two kinds of items are presented in Figure 11. The 'more informative' contexts are the ones where an appropriate answer would be the more informative option, namely an NP and its description. The less informative contexts are the ones where an NP alone satisfied the needs of the interlocutors.

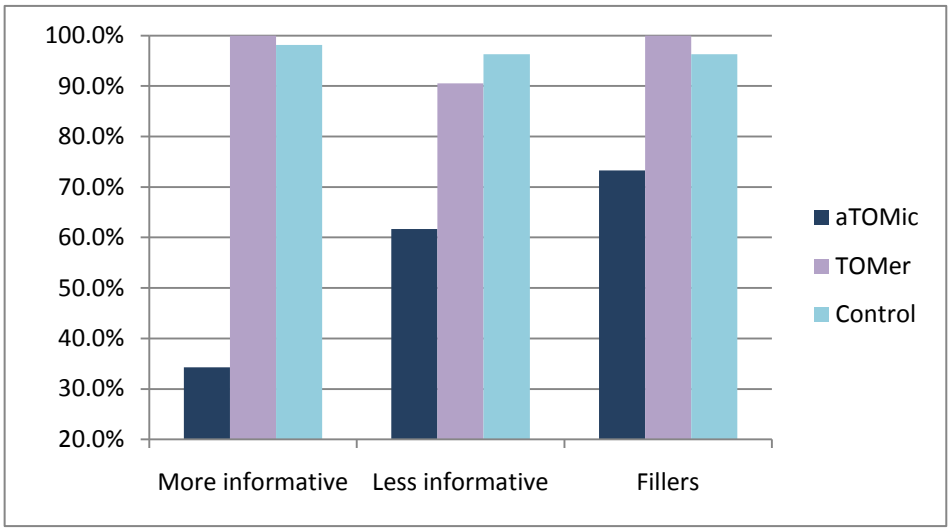

Figure 11. Percentage of correct responses to two kinds of contexts

The results show that the aTOMics performed worse than the other two groups in both kinds of contexts. They tended to choose the option which contained less information in the more informative contexts and the more informative options in contexts where this information was superfluous. In both cases the aTOMics demonstrated a breach of the *relevance maxim* and inappropriate use of reference terms according to Ariel's accessibility theory.

The aTOMics average *total score* (48.1%, SD = 22.6%) was significantly lower than the TOMers (average = 95.2%, SD = 9.4%) and the controls (average = 97.2%, SD = 5.9%) (U = 68, p = .001; U = 88.5, p < .001, respectively). No significant difference was found between the TOMers and controls (U = 34, p = .41). As Figure 11 indicates the same pattern of results was found for the two contexts presented.

In choosing appropriately the *more informative* options, the aTOMics score (average = 34.3%, SD = 24.7%) was significantly lower than the TOMers (average = 100%) and control's (average = 98.1%, SD = 5.6%) (U = 70, p < .001; U = 89.5, p < .001, respectively). No significant difference was found between the TOMers and controls (U = 28, p = .38). We also tested the RBD patients' individual performance as compared to the control group. We found that each of the aTOMic participants and none of the TOMers performed significantly poorer than the control group (p < .05, using Crawford and Howell's, 1998 t-test). For detailed information see Table 6.

In choosing appropriately the *less informative* options, the aTOMics score (average = 61.7%, SD = 27.3%) was significantly lower than the TOMer's (average = 90.5%, SD = 18.9%) and control's (average = 96.3%, SD = 7.3%) (U = 58.5, p = .01; U = 81.5, p = .001, respectively). No significant difference was found between the TOMers and controls (U = 34.5, p = .4). In comparing the individual performance of the RBD patients we found that 6 out of the 10 aTOMics and 1 out of 7 TOMers performed significantly poorer than the control group (p < .05, using Crawford and Howell's, 1998 t-test). For detailed information see Table 6.

We also compared the aTOMic's performance on the two test conditions using Wilcoxson test. We found that their performance on the *less informative* condition, when a less informative referring expression is the correct option, was significantly better than their performance in choosing appropriately the more informative option (T = 5, p = .01).

On the *filler items*, the aTOMics score was higher (average = 73.3%, SD = 30.6%) but still significantly lower than the TOMer's (average = 100%) and the control's (average = 96.3%, SD = 7.3%) (U = 56, $p < .05$; U = 67.5, $p < .05$, respectively). Again, no significant difference was found between the TOMers and controls (U = 28, $p = .37$).

We also compared the difference between the aTOMics' average score on the test items (average = 48%, SD = 22.7%) and their score on the filler items (average = 73.3%, SD = 30.6%). We found they performed significantly better on the filler items (T = 1, $p < .01$). This difference showed that although the filler items were harder for the aTOMics than for the two other groups, their performance on these items was still significantly better than on the test items.

Summary

In this task we presented participants a choice between two options to end a short interaction. The task was intended to test if the participants were able to fit a more or less informative ending depending on the context. To accomplish this task, the participants had to consider the relations between the characters presented and to evaluate the amount of information needed to continue the conversation. Crucially, they would also need to implement their linguistic knowledge about the appropriate reference term that would express the appropriate amount of information according to the interlocutors' information needs.

The results show that the aTOMics had difficulty in this task. In both contexts they did not translate the relations between the characters presented to choosing the appropriate reference term that contained the amount of information needed in the conversation.

Most of the errors the aTOMics made were in preferring the less informative option in contexts where more information was needed. By doing so they tended to breach Grice's maxim of relevance. Their responses also did not match Ariel's proposal regarding the appropriate use of accessibility markers. The aTOMics also erred in contexts where the less informative choice was appropriate. In both cases it seems that the lack of TOM causes the patients to dismiss the characters' mutual acquaintance as a key for choosing the appropriate answer.

In this task the participants were presented with a metalinguistic assignment, which *phrase is more appropriate in the context*. The results show that the aTOMics suffer lack in this knowledge, which was evident in their inability to attend to the relevant information in the context. For example, in the justifications for choosing the inappropriate answer to item (19) presented above, the participants quoted below in (23) and (24) focus on the irrelevant detail of the accountant's working hours rather than on the fact that he is being introduced to a new worker:

(23) Dror: *Because that (= the other option) is not appropriate, here he stresses that it is until 4.*

(24) Arye: *Because he explains to him that he arrives every Wednesday, so he won't hope to see him all week, he only works Wednesdays.*

It is important to note again that the TOMers answered the same way as did the control group, showing that the distinction in TOM is more productive than separating the participants on the basis of the occurrence of brain damage.

4.2.2.4. Choosing between NPs: A short NP or an informative description?

Participants

The participants in this task were 18 RBD patients, 11 were aTOMic and 7 TOMer. A group of 9 control participants was also tested. See Table 2 for a detailed list of the patients who participated in this task.

Materials

The aim of this task was to test whether the participants were sensitive to the relation between the interlocutors in a certain situation. We presented two kinds of interactions and asked the participants to choose between two possible endings. In half of the situations the characters presented in the items had a considerable amount of common knowledge (e.g., mother and father talking about their little boy's toys), and in the other half there was a small amount of mutual knowledge (e.g., a sales person and a customer who wants to buy a backpack). The participants were asked to choose between an answer that contained a small amount of information (short

answer) and an answer that contained more information (long answer). The short answers were appropriate to situations where the characters had considerable common knowledge and the long answers when there was little.

The examples below present the two possibilities. In (25) the short answer is appropriate because there is enough mutual information between the characters and the other answer includes too much information for that situation. In (26) the opposite is true; more information is needed so the short answer is not appropriate.

(25) Noam is three years old and he loves his teddy bear. He has only one teddy bear and he takes it with him everywhere he goes. When Noam's parents were packing for a weekend trip, Noam's mother reminded his father:

a. We must not forget to take Noam's teddy bear in the morning.

b. We must not forget to take the brown teddy bear with the rounded ears that's in Noam's bed.

(26) Ayelet wants to go on a trip to the Far East. The camping store has a large variety of bags, all of them on the top shelves. Ayalet can't reach the one she wants. How will she ask the sales women?

a. Please show me the bag from the shelf.

b. Please show me the big blue bag on the top shelf.

Procedure

The items were presented in a semi-randomized order after one practice item. Participants were asked to choose one option and to justify their choice. The items were read by the experimenter and read again if asked. The answers were recorded and later transcribed.

Results

Coding: The answers were coded as correct (1) if the participant chose the appropriate answer or (0) if not. A justification was coded as correct (1) if it included some mention of the common knowledge or relations between the characters presented, and (0) if it was irrelevant.

The final score was calculated by multiplying the answer score with the justification score. This way, only correct answers with appropriate justification were given a final correct score (1). If participants either choose the inappropriate answer or did not justify it appropriately their final score was (0).

Below are examples of appropriate and inappropriate justifications. In (27) two inappropriate justifications which were given in reply to (25) above and in (28) an appropriate justification. In (29) an inappropriate justification to (26) above and in (30) an appropriate one.

(27) Gila: answered (b: inappropriate long answer) and justified: *Because she is telling him which bear exactly, .. and he has a lot of teddy bears.*

[15]Ex: *No, it says the boy only has one and he takes it with him everywhere*.

G: *I still say (b) ..the one that in bed with him, So he won't be confused*.

And

Sachar: chose (b: inappropriate long answer) and justified: *Because it's more specific and detailed ..he'll know exactly which one to take and won't be confused.*

(28) Ahuva: chose (a: appropriate short answer) and justified: *It's more appropriate because he has only one.*

(29) Daniel: chose (a: inappropriate short answer) and justified: *She wants a backpack so she asks for one.*

(30) Sharon: chose (b: appropriate long answer) and justified: *She says exactly the one she wants, without wearing out the sales women.*

Pretest: The task was presented in a written form to a group of 8 healthy adults (between the ages of 20-24) in order to test if our choices of appropriate answers were not biased. The average total score was 94.4% (SD = 10.3%). Before summing the results we checked whether any item received less than 85% correct responses in the responses of the control participants (8 who answered the written version in the

[15] The experimenter interventions were very rare, only in cases where it seemed the participant misunderstood the fact in the story.

pretest and 9 who answered to the oral presentation. We found one test item that did not meet this criterion. It was discarded from further analyses.

The correct responses to the two kinds of contexts, those that were designed to elicit short answers and those that designed to elicit long answers are presented in Figure 12.

Results show that the aTOMics scored lower than the TOMers and controls on both items that required a shorter answer and in the answers that required the more detailed one. They tended to err more in choosing a detailed long answer although it was not needed, because the characters presented had a high rate of mutual knowledge. The *total score* of the aTOMics (average = 60.6%, SD = 21.9%) was significantly lower than the TOMers (average = 92.1%, SD = 5.4%) and control's score (average = 96.1%, SD = 5.8%) (U = 72, p = .001; U = 95, p < .001, respectively). There were no significant differences between the TOMers and control's total scores (U = 41, p = .17).

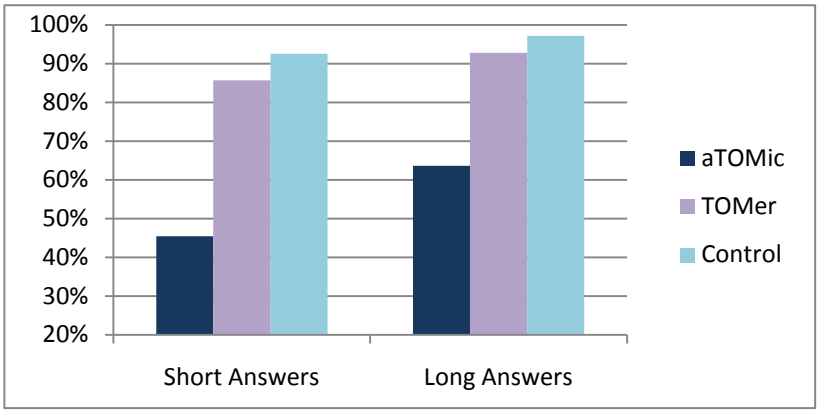

Figure 12. Percentage of correct responses to short and long answers given by the three groups

As Figure 12 indicates the same pattern of results was found for both the appropriately *short answers* and the appropriately *long answers*.

On the *short answers* condition the aTOMics' scores (average = 45.5%, SD = 37.3%) were significantly lower than the TOMers (average = 85.7%, SD = 17.8%) and

controls' score (average = 92.6%, SD = 14.7%) (U = 62, $p < 0.05$; U = 82.5, $p < 0.01$, respectively). There were no significant differences between the TOMers and controls' scores (U = 38, $p = .26$).

On the *long answers* condition the aTOMics' scores (average = 63.6%, SD = 25.9%) were also significantly lower than the TOMers' (average = 92.9%, SD = 12.2%) and control's score (average = 97.2%, SD = 8.3%) (U = 64, $p = .01$; U = 95, $p < .005$, respectively). There were no significant differences between the TOMers and controls' scores (U = 32.5, $p = .32$).

No significant difference was found in the aTOMics' performance in the two test conditions (T = 13, $p = .08$).

We also tested the individual performance on the two conditions. The comparison of each of the RBD patients to the control group showed that in the *short answer* condition 8 of the 11 aTOMics performed significantly poorer than the control group ($p < .05$, using Crawford and Howell's 1998 t-test). The TOMers performed no differently from the control group. In the *long answer* condition 6 out of the 11 aTOMics and 2 out of the 7 TOMers performed significantly poorer than the control group ($p < .05$, using Crawford and Howell's, 1998 t-test). For detailed information see Table 6.

Summary

The results of this task showed that the aTOMics had difficulty in choosing an answer according to the relations and mutual knowledge held by the characters presented in the context. They made both errors of inappropriately choosing the short answer when a more detailed answer was needed, and of choosing a detailed answer when a shorter one was sufficient.

In this task the participants were presented again with a metalinguistic assignment, choosing which of two answers is more appropriate in the context. The results demonstrated again that the aTOMics suffer lacked in this knowledge. Their choices breached Grice's maxim of relevance and showed the connection between the lack

they suffer in TOM and their linguistic abilities. In this task too, the TOMers scored no different then the control subjects.

4.2.3. Interim Summary and Discussion: Reference and aTOMia

The range of tasks presented in this section examined the way RBD aTOMic patients use referential terms in comparison to RDB TOMer patients and controls. The first two tasks tested the participants' abilities in *production.* In one task we asked the participants to retell stories and in the other we asked them to describe one of two similar figures. In both tasks the aTOMics performed significantly poorer than the TOMers and controls. The main conclusion from these tasks is that as speakers, patients who suffer aTOMia produce referring expressions with less sensitivity to their addressee compared to the other groups. In the retelling task they produced less referring expressions overall and made more errors of using the inappropriate reference term that did not allow their addressees to reach the referent they intended. Most of their mistakes were in choosing a high accessibility reference term when a lower accessibility marker should have been used. High accessibility terms (pronouns) should be used when the referent is highly accessible to the addressee, and there is no other referent that can compete with it. The aTOMics tended to use pronouns in contexts where a unique identification of the intended referent was not possible because the referent was not accessible enough.

In the second task we tested the aTOMics' ability to describe to their addressees (in this case the experimenter) one of two similar referents in a way that will allow her to identify the referent they intended her to pick out. In this case too, the aTOMics problem was of assuming too much mutual knowledge. They mistakenly assumed that their addressees already know their intention, or they simply ignored the differential accessibility with which mental representations are entertained by addressees, which should dictate the way they tailor their use of referential terms to fit their listeners state of knowledge.

The four other tasks tested the participant's metalinguistic knowledge about the proper way to use referential terms. In these tasks we asked the participants to make a choice (and in two cases to justify it) between two ways of referring, or to choose to which of

the two characters the speaker was referring to. In these four tasks the aTOMics again performed significantly poorer than the TOMers and controls.

In the first task we saw that the aTOMics preferred to name the most accessible referent, namely, the first mentioned character. They were not attentive to the change of accessibility term that could have directed them to choose the less available referent, namely, the second mentioned character.

In the second task we found an opposite tendency to the one found in the production tasks. The aTOMics preferred to excessively use a low accessibility term (a proper name) within the same sentence, as though there was no mutual knowledge about the sentence topic. We did not ask to justify the decision in this task (because all the items were of the same kind, therefore the reason would have to be repeated for all the items) but a few participants added spontaneous justifications. Below is an item and a reply given by one of the aTOMic patients:

(31) a. Gil phoned his sister Rina and *she* quickly answered the phone.

b. Gil phoned his sister Rina and *Rina* quickly answered the phone.

Dror: *The second.. here* (pointing to a) *it's not clear who, its written 'she'.*

We can learn from this reply that the participant did not acknowledge that the character was just introduced by her proper name. So in this case the aTOMia led to dismiss of mutual knowledge that would have led to a choice of a higher accessibility marker.

In the next two tasks we presented contexts in which the characters interacting were either closely acquainted or not, and asked the participants to complete the interaction with one of two options. The appropriate answer would be to choose the more informative reference term and the most detailed noun phrase if the partners in the conversation lacked mutual knowledge. In these two tasks we also asked the participants to justify their choices.

In both tasks the aTOMics performed significantly poorer than the other two groups but there were also differences between the results of the two tasks.

In choosing the appropriate reference term the aTOMics erred significantly more in choosing the less informative term when a more informative one was the correct choice. These results are similar to the ones found in the production tasks. The participants were not sensitive to lack of mutual knowledge and did not chose the phrase that fits this situation. They also erred in choosing a more informative answer when the characters had enough mutual knowledge, but significantly less.

In the task of choosing a more or a less informative NP, the aTOMics erred in both directions, showing a general disability to fit the level of informative content of their answer to the relations between the characters that were presented in the context.

In sum, we can conclude that we found a meaningful connection between aTOMia and the ability to properly use and understand referring expressions. The RBD patients who had difficulty in the TOM battery also performed significantly poorer than RBD patients who succeeded in the aTOMic battery in the different tasks that tested this ability; to tailor the linguistic message according to state of knowledge of two interlocutors.

In sum, the results show that the aTOMic patients use referential terms without considering the mutual knowledge two characters hold. This tendency is found both in production tests when they need to monitor the accessibility of different items in the mind of their hearers, and in cases where they need to judge the appropriateness of a character's production. The TOMers performance in all these tasks was not different from that of the controls.

4.3. Mental State Verbs

This section describes the way RBD patients use and comprehend mental state verbs (MSV). This topic is an interesting intersection point between TOM and language because these verbs are the linguistic tool for describing and reasoning about inner mental states. The appropriate use of MSV relies on the ability to consider semantic, lexical and syntactic characteristics of these verbs. We focused on three lexical-semantic characteristics: (a) Factivity (b) Certainty (c) Time frame.

4.3.1. Factivity

The purpose of this task was to test if RBD aTOMic patients distinguish between potentially factive (i.e., p-factive) verbs and non-factive verbs. The target sentences we tested were sentences which included a negated mental verb as the matrix verb. We asked the participants to judge the truth validity of their complement. Because these verbs were p-factives, their negation should not affect the truth validity judgment of their complement.

Participants

The participants in this task were 17 RBD individuals: 10 aTOMics and 7 TOMers, and 6 control subjects. See Table 2 for detailed list of the patients who participated in this task.

Materials

The purpose of this task was to test if participants distinguish between potentially factive (i.e., p-factive) verbs and non-factive verbs. We presented a truth verification task of four kinds of sentence complements; complements of p-factive predications in affirmative and negative and of non-factive predications, affirmative and negative. The point we tested was whether the participants understood that only the complements in the p-factive predications are presupposition. If they understood so, their judgment about the truth/ falsity of the complement should not change when the p-factive matrix verb is negated. For example:

(32) Pinchas knew the door was closed. Was the door closed? (yes)

(33) Pinchas didn't know the door was closed. Was the door closed? (yes)

In contrast, the complements in the non-factive predication are not presupposition, they are implications and therefore the truth verification judgment of these complements is changed when the implicative matrix verb is negated. For example:

(34) Shlomi wanted three of his friends to come over and play with him yesterday. Did those friends come over? (probably yes/ maybe).

(35) Shlomi didn't want three of his friends to come over and play with him yesterday. Did those friends come over? (probably not).

The list of predications was chosen after a pretest that was presented to 20 normal participants (between the ages of 15-64). We presented 16 MSV[16], half were p-factive and half non-factive. They were presented in 32 mental predications. In half the matrix verb appeared in affirmative (see example 32, 34 above) in the other half, the matrix verb was negated (see example 33, 35 above). In addition 5 more predications were presented in which the verb in the complement was negated (see example 36 below). The participants were asked to answer a truth verification question for each predication:

(36) Pinchas knew the door wasn't closed. Was the door closed? (no)

If more than 3 participants answered in contrary to the expected answer, the predication was omitted from the list. Ten predications were omitted due to this criterion, three affirmative, 5 negated and 2 in which the verb in the complement was negated.

The final test list included the predication that passed the pretest criterion and 2 extra verbs. The task included 15 verbs, 7 were potentially factive (i.e., p-factive) and 8 non-factive. All seven p-factive verbs appeared twice. Once the matrix verb was affirmative and once it was negated. Three of the verbs appeared three times: in the third, the verb in the complement was negated.

[16] We included sensory verbs like *hear* and *see* in order to vary the list as much as possible.

Of the 8 non-factive verbs, 6 appeared twice. Once the matrix verb was affirmative and once it was negated. One of these verbs appeared three times; a negation of the complementation was also added. Two other verbs appeared only once, in affirmative.

Three of the non-factive predications appeared with tensed complements, 5 appeared as infinitives. Examples of the tensed construction appear in (34-35) above, and of the infinitive construction in (37-38) below:

(37) Sagit was convinced she should give her secretary a bonus on her next salary. Will she give the bonus?(yes)

(38) Sagit was not convinced she should give her secretary a bonus on her next salary. Will she give the bonus?(maybe/ no)

The final task was presented in a written form to a group of 12 healthy adults (between the ages of 25-65) that also served to verify the answers.

Procedure

The final task included 32 sentences. After each sentence a truth verification of the complement was presented. The participants were asked if the complement was true/ false or maybe true. The sentences were presented in a semi-randomized order. Three practice trials were presented at the beginning.

Results

Coding: The answers were coded as correct (1) if they matched the expected answer according to the type of predication presented (and according to the normal controls in the pilot test) and incorrect if not (0).

In the final analysis we included only predication to which more than 85% of all healthy control participants (those who answered the task in the written form and the orally presented test and pretest). Nine predications were omitted according to this criterion. In 3 the matrix verb was a p-factive predicate that was negated, in 2 the matrix verb was non-factive that was negated, in 2 the verb in the complement was negated, and in 2 the matrix verb was a p-factive affirmative.

The total score of the aTOMic group (average = 74.3%, SD = 13.2%) was significantly lower than the TOMers (average =87.9%, SD = 6.4%), and control's

(average =89.4%, SD = 6.1%) (U = 59, $p < .001$; U = 53, $p < .01$, respectively). There was no significant difference between the TOMers and controls' total average score (U = 24.5, $p = .33$).

The sentences we wanted to focus on were the negated predications, both p-factive (see (33) above) and non-factive (see (35, 38) above). There were 4 negated p-factives and 4 negated non-factives that were included in the final analyses. The correct responses of the three groups to these 8 items are presented in Figure 13 below.

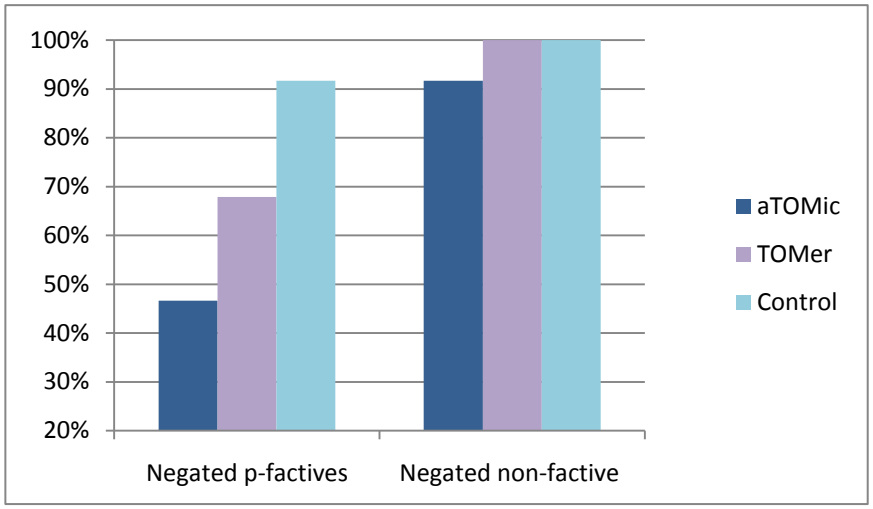

Fig 13. Percentage of correct responses to negated p-factives and negated non-factives

The results show that the aTOMics encountered difficulty in verifying the truth conditions of the complements of predications with negated p-factives matrix verbs. They did not show the same difficulty in verifying the truth of complements that were embedded to a negated non-factive matrix verb. The aTOMics' average score in negated p-factives was significantly lower than their score on non-factive negated predications (T = 0 , $p < .01$).

The average score in the *negated p-factive* condition of the aTOMic group (average = 46.7%, SD = 43.1%) was lower but not significantly different from that of the TOMers (average = 67.9%, SD = 34.5%), (U = 45, $p = .17$). But it was significantly lower than the average score of the control group (average = 91.7%, SD = 12.9%), (U = 47, $p < .05$). No significant difference was found between the TOMers and controls (U = 29, $p = .14$).

In comparing the individual performance of the aTOMics to the control group we found that 6 out of 10 aTOMics and 4 of the 7 TOMers also performed significantly worse than the control group ($p < .05$, using Crawford and Howell's 1998 t-test). For detailed information see Table 6.

The average scores in the '*negated non-factive*' condition were very high in all groups, and there were no significant differences among the three groups. The average score of the aTOMics was 91.7% (SD = 18%). The average score of both the TOMer group and control group was 100%. The difference between the aTOMics and TOMers was not significant (U = 42, $p = .26$). The difference between the aTOMics and controls was also non-significant (U = 36, $p = .27$). In comparing the individual performance of the aTOMics to the control group we found that only 2 out of 10 aTOMics performed significantly poorer than the control group ($p < .05$, using Crawford and Howell's, 1998 t-test).

The average score in the *affirmative* condition of the aTOMic group (average = 84.3%, SD = 11.7%) was significantly lower than that of the TOMers (average = 97.6%, SD = 4.1%) and controls (average = 95.8%, SD = 4.6%) (U = 61, $p < .01$; U = 51, $p = .01$, respectively). No significant difference was found between the TOMers and controls (U = 16.5, $p = .28$).

The average scores in the *negated complement* condition were very high in all groups, there were no significant differences between the three groups. The average score of the aTOMics was 95% (SD = 15.8%). The average score of the TOMer group was 92.9% (SD = 18.9%) and the control group scored 100%. The groups did not differ on this condition (aTOMics and TOMers, U = 33.5, $p = .46$; aTOMics and controls, U = 33, $p = .4$; TOMers and controls, U = 24, $p = .36$.

Summary

In this task we asked participants a truth verification question about affirmative and negative sentences in which the matrix verb was either p-factive or non-factive. The results of the pretest and the high number of items that were excluded from the final analyses show that healthy controls do not answer as expected according to the theoretical analyses of these constructions to many of the predications we presented. For example, they tend to answer 'no' to the verification question (39) below.

(39) Dan did not apologize that he came late to class. Did he come late to class?

One explanation might be that they are extending the scope of negation and including the complement in it. Another explanation might be that they were considering different situations in which one can apologize for being late. As if they were asked: *Does the fact that Dan apologized necessarily means that he came in late*?

The first explanation arises from some of the participants talk while answering the task. For example, one of the control subjects while answering (correctly) the question above (40) revealed something about her thought process.

(40) *Lilia: Didn't apologize, so he wasn't there, no, he was there.*

The same spontaneous reply came from an aTOMic patients to a number of the p-factive negated items for example:

(41) *Arye: If he didn't know that the door was closed then the door was not closed.*

The second possibility, e.g., that they were considering possible situations, can be backed up by the finding that most incorrect replies were 'maybe' and not 'no', the opposed answer to the correct one.

When we isolated the sentences to which more than 85% of the healthy participants answered as expected, after the pretests, we found that there were significant differences between the negated p-factives and negated non factives. The aTOMic group answered the negated non-factives the same as the controls. But, in answering the negated p-factives they erred significantly more than controls.

From these findings we conclude that the aTOMics has a specific difficulty in answering the negated p-factive predication. In these cases there is a need to entertain a contradiction: the truth condition of the complement stays the same while those of the complex sentence switch, For example:

(42) Ofer *found out* that there was chocolate in the drawer. Was there chocolate in the drawer? (yes)

(43) Ofer *did not find out* that there was chocolate in the drawer. Was there chocolate in the drawer? (yes)

These sentences require the participants to entertain two seemingly conflicting states of affairs. While the chocolate is in the drawer, the character presented does not hold that knowledge. It seems that representing these two opposing states of events is difficult for the aTOMic group.

The TOMers had less difficulty than the aTOMics in this task. Their scores were intermediate, not significantly different from controls, but not significantly different from TOMics either. The controls had no difficulty in this task, their answers were above 90% correct in both tasks.

4.3.2. Factivity in context

The aim of this task was to further explore the way aTOMics understood the p-factive lexical property. In this task we presented contexts and asked the participants to judge the felicity of different sentences that appeared after it.

Participants

The participants in this test were 19 RBD participants, 13 aTOMic, 6 TOMer, and 9 control participants. See Table 2 for detailed list of the patients who participated in this task.

Materials

We presented 4 situations in which people are saying sentences with mental verbs while they are engaged in some activity (e.g., looking for keys). We asked the participants to judge if the sentences were felicitous or not, and if they judged them as infelicitous, to offer a correction. The target sentences included a p-factive matrix verb complemented by an embedded phrase that was followed by a contradicting phrase using the conjunction 'but'. For example: One of the situations described a man sitting in his apartment trying to figure out where sounds of whining are coming from. He's saying:

(44) #I *know* that it's the neighbors' son, but it's not him.

Because a p-factive verb is the matrix verb, a contradiction of the complement creates an infelicitous sentence.

We also presented non-factive predication with the same kind of contradiction:

(45) I *think* it's a baby crying but there are no babies in our building.

In this case the contradiction does not cause the whole sentence to be infelicitous.

The test included a total of 32 items, 8 items after each of 4 contexts. In each of the 8 items, three were p-factives sentences that needed a correction (i.e., target sentences) and three, non-factives that did not need a correction. There were also 2 filler sentences in each situation that did not contain a mental verb. One required a correction and one did not.

Procedure

The contexts were presented in a semi-randomized order. One situation with two sentences was presented as a practice trial, in one of the sentences a correction was needed and the other did not.

Results

Coding: the target sentences were scored correct (1) only if the participants judged them as inappropriate, and gave an appropriate correction. If a target sentence was judged as incorrect but the correction offered was inappropriate, the answer was considered as incorrect (0). The filler items that needed correction were coded the same way. The answers to the mental predications and filler sentences that were felicitous were graded as correct (1) or incorrect (0) according to the answer to the yes/ no question.

There were 4 incidents where a participant judged a target sentence as inappropriate but did not give a felicitous correction (out of 840 responses of all subjects to all items). Three were in the aTOMic group and one in the control group.

Pretest: we presented the task to 8 healthy adults (between the ages 19-24). If an item received less than 85% correct responses in this group, it was discarded. Two items didn't meet this criterion, one was a filler item (average score 65%) and one a mental predication (average score 50%). The average score on the other items ranged between 88% - 100% correct.

The aTOMic group's *total score* in this task (average = 81.3%, SD = 11.6%) was significantly lower than the total score of the TOMer group (average = 97.4%, SD = 3.1%) and the control group (average = 98.6%, SD = 1.6%) (U = 75.5, $p < .001$; U = 115, $p < .001$, respectively). There was no significant difference between the TOMers and the control group (U = 32.5, $p = .59$). The responses to the mental predications are presented in Figure 14.

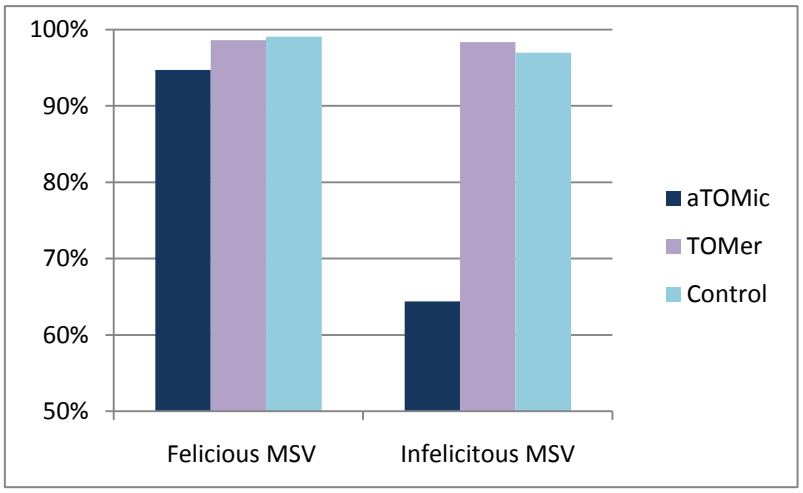

Figure 14: Percentage of correct responses to mental predications

Different patterns of results appear in response to the two kinds of mental predication presented. More than 90% of the mental predications that were felicitous were confirmed as correct by all three groups. But the identification of infelicitous MSV predication was considerably more difficult for the aTOMic group.

In the *felicitous condition* the aTOMics average score was 94.7% (SD = 10.1%), it was not significantly different from the TOMer average score (98.6%, SD = 3.4%) (U = 37.5, $p = .34$), and the controls scored 99.1% (SD = 2.8%) (U = 58.5, $p = .26$). No significant difference was found between the TOMers and controls either (U = 28.5, $p = .45$).

Crucially, the groups differed in their responses to the *infelicitous* sentences. The aTOMics' average score 64.5% (SD = 27.1%) was significantly lower than that of the average score of the TOMers 98.3% (SD = 4.1%) (U = 31.5, $p < .01$) and controls 97% (SD = 4.5%) (U = 91.5, $p < .001$). Again the difference between the TOMers and controls' scores was not significant (U = 24, $p = .39$).

In comparing the individual performance of each of the RBD patients to the control group we found that 9 out of 11 aTOMics but none of the TOMers performed significantly worse than the control group ($p < .05$, using Crawford and Howell's, 1998 t-test). For detailed information see Table 6.

The aTOMics' average score on the filler items was 89.6% (SD = 9.2%), not significantly different than that of the TOMers 97.6% (SD = 5.8%) (U= 49, $p = .06$). But, the aTOMics scored significantly lower than controls, who performed at ceiling on these filler items (average = 100%, U = 81, $p = .01$). There was no significant difference between the TOMers and controls on these items (U = 31.5, $p = .32$).

Summary

This task further tests the aTOMics ability to comprehend p-factive MSV. In this task they judged sentences, which included an embedded complement that was contradicted, as felicitous or infelicitous. When the sentences included a p-factive verb the complement is a presupposition and therefore its contradiction causes the sentences to be infelicitous. For example (46):

(46) *He knew that it was the neighbor's son but it wasn't. (him making the noise)*

The aTOMics scored significantly lower in judging these sentences as infelicitous than the TOMers and controls, showing that they see p-factives as non-factive that can be negated without affecting the sentence presuppositions.

These results demonstrate again that the aTOMics difficulty appears when there is a need to hold or consider two opposing viewpoints regarding the same situation. In this task they needed to consider some content of thought (which appeared in the complement), and its contradiction.

The TOMers, on the other hand. had no such difficulty and they were able to judge these sentences as infelicitous as well as controls.

4.3.3. Certainty

In this task we tested the participants' sensitivity to the different levels of certainty the different MSV verbs represent.

Participants

The participants in this test were 20 RBD patients, 13 were aTOMic and 7 TOMers. A control group of 12 participants was also tested. See Table 2 for detailed list of the patients who participated in this task.

Materials and Procedure

In order to test if the RBD patients were sensitive to the difference in the degree of certainty different mental verbs portray, we presented them with two-choice questions. The question described situations in which a high or low level of certainty was evoked and the participants were asked to choose between two sentences to continue the interaction. Only one of the two options signals the level of certainty that suits the situation. For example (47):

(47) Who would you prefer to buy a used car from?

a. From someone who *thinks* the car is in good condition.
b. From someone who *knows* the car is in good condition.

The correct option was (b) because '*know*' signals a higher degree of certainty then '*knows*'.

The task contained 12 target situations, and 6 filler items. There were two situations where the expected answer was the one with low certainty.

The task was presented in a semi randomized order after one practice trial.

Results

Coding: correct answers were scored as (1) incorrect answers were scored (0).

Two items were discarded from the final analysis because less than 85% of the participants in the control group answered them correctly.

The scores were very high in all three groups. But, the aTOMics' *total average* score (86.4%, SD = 8.8%) was significantly lower than that of the TOMers (94.5%, SD = 2.4%) (U = 74.5, p = .01) and controls (97.9%, SD = 3.1%) (U = 138, p = .001). The controls also scored significantly better then the TOMer patients (U = 66, p = 0.02).

In the analysis of the *mental items*: there was no significant difference between the average score of the aTOMics (average = 86% SD = 11.3%) and that of the TOMers (average = 92.9%, SD = 4.9%) (U = 60.5, p = .13). But the aTOMics scored significantly lower than the controls (average = 97.5%, SD = 4.5%) (U = p < .01). The TOMers also scored significantly less than the controls (U = 61.5, p = .05).

In the *filler items* there was no difference between the aTOMics (average = 87.2%, SD = 12.1%) and TOMers (average = 97.1%, SD = 7.6%) (U = 65, p = .07). The aTOMics scored significantly less than the controls (average = 98.6%, SD = 4.8%) (U = 120.5, p = .01). No significant difference was found between the TOMers and control group (U = 45, p = .41).

Summary

In this task we contrasted two statements and the participants had to choose which signals a higher level of certainty. We found that the aTOMics performance was similar to that of the TOMers in the mental predications, but both groups scored significantly lower than the control group. The overall high scores of the aTOMic group on the mental predications show that they do have an ability to compare between the levels of certainty of two MSVs when they are presented separately, as two different options and not as a contradicting condition within one option.

In this task the TOMers' score was significantly lower than that of the control's, on the mental items, but it is important to note that it was above 90% correct.

4.3.4. Matching semantics to tense of complement

This task intended to test the participants' sensitivity to the relations between the mental verb's meaning and the time inflection of their complement. As far as we know this lexical property has not been experimentally tested before.

Participants

The participants in this test were 21 RBD patients, 14 aTOMics and 7 TOMer. A group of 7 control participants were also tested. See Table 2 for detailed list of the patients who participated in this task.

Materials and Procedure

In this task we tested the participants' sensitivity to the time frame implied by various mental state verbs. Some mental verbs represent anticipation about the future (e.g., hope, wish-for). These verbs complements appear in future tense (e.g., I hope that I will see you in the meeting). Others relate to some event in the past (e.g., regret, sorry). The complements that appear after these verbs are in the past tense (e.g., I'm sorry that I didn't come to the meeting[17]).

We presented 10 mental predicates. Six semantically relate to the past (*was sorry, found out, was happy, noticed, discovered, regret*) and 4 semantically relate to the future (*planned, asked for, promised, predict*ed). Each verb was presented twice as the matrix verb of complex sentences. In one, the verb in the complement was tensed according to the matrix verb and the result was a felicitous sentence. In the other, the verb in the complement was tensed opposed to the semantic characteristics of the matrix verb and the predication created was infelicitous.

For example: (48) is appropriately tensed to past tense and (50) appropriately tensed to future. Sentences (49) and (51) are the infelicitous minimal pair of each.

[17] Sentences like *I'm sorry I won't be coming tomorrow* can be paraphrased as: *I'm sorry but I understood* that I will not be able to come tomorrow (P. Schultz, 2003).

(48) Ayala *gilta* se-hi *dibra* yoter miday.
Ayala *found-out* that-she *talked* too much.

(49) #Ayala *gilta* se-hi *tedaber* yoter miday.
Ayala *found-out* that-she *will-talk* too much.

(50) #Dani *bikesh* she-aba *kone* lo mamtak.
#Dani *asked* that-father *is buying* him candy.

(51) Dani *bikesh* she-aba *yikne* lo mamtak.
Dani *asked* that-father *will-buy* him candy.

In addition there were 8 filler items. All together the task included 28 sentences.

The sentences were presented in a semi randomized order after one practice trial. The participants were asked to judge if the sentences were felicitous or not and if not, to offer a correction.

Results

Coding: the answers to the infelicitous sentences were graded correct (1) only if the participants judged them as inappropriate, and gave an appropriate correction. If a sentence was judged as incorrect but the correction offered was inappropriate, the answer was considered as incorrect (0). The answers to the felicitous sentences were graded as either correct (1) or incorrect (0) according to the yes/no judgment answer.

There were three incidents where the subject failed to give an appropriate correction after judging the sentence as infelicitous (out of 868 total answers given). All three were in the aTOMic group.

The scores were very high in all three groups in both conditions. But, the *total score* of the aTOMic group (average = 86.8%, SD = 9.1%) was significantly lower than that of the TOMers total score (average = 97.4%, SD = 3.4%) ($U = 82.5$, $p < .01$) and the controls (average = 97.9%, SD = 3.5%) ($U = 125$, $p < .001$). There was no significant difference between the TOMers' and controls' scores ($U = 43.5$, $p = .21$).

Figure 15 presents the three groups' scores on the felicitous and infelicitous mental predications:

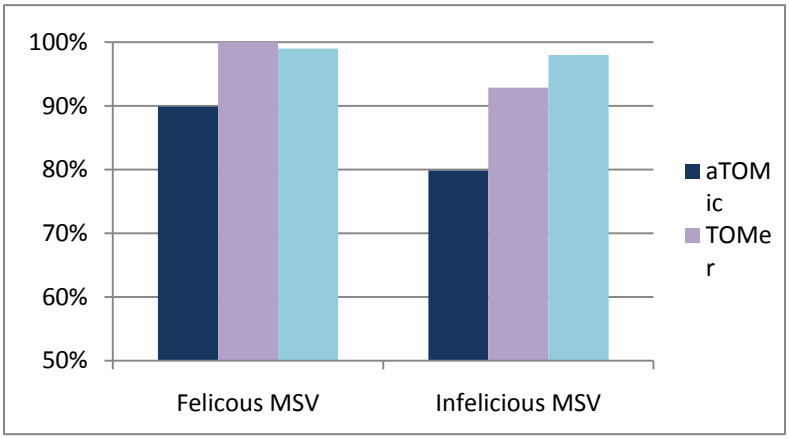

Figure 15. Percentage of correct responses to mental predications

These results show that the responses to the felicitous sentences with mental state verbs, when the tense of the complement fit the semantic of the matrix verbs were high for all groups. But identifying the infelicitous predications was harder for the aTOMics than for the other two groups.

In the *felicious condition*, the average score of the aTOMic group was 89.9% (SD = 15.2%), the TOMers' (average = 100%), and the controls' (average = 99%, SD= 3.2%). No significant difference was found between the aTOMics' and TOMers' scores (U = 70, p = .06), and between the aTOMics and controls (U = 95.5, p = .07) or the TOMers and control (U = 31.5, p = .39)

In the *infelicitous condition*, the aTOMic group scored (average = 79.8%, SD = 15.2%) significantly less than the TOMers (average = 92.9%, SD = 9.5%) and controls (average = 98%, SD = 4.2%), (U = 75, p < .05; U = 120, p = .001, respectively). There was no difference between the TOMers' and controls' average scores (U = 45, p = .17).

The aTOMics scored significantly better on the felicitous compared to the infelicitous predications (T = 4.5, p = .01). We also tested the individual performance of the RBD patients compared with the control group. We found that 11 out of 14 aTOMics and 3 out of 7 TOMers performed significantly poorer than the controls. This is due to almost perfect scores in the control group.

In the *filler items*, the average score of the aTOMic group was 92% (SD = 10.5%), the TOMers' average was 100%, and the controls' average was 96.3% (SD = 6%). No significant differences were found between the aTOMics and TOMers (U = 70,

p = .06), between the aTOMics and control (U = 83.5, p = .22), or between the TOMers and control (U = 24.5, p = .16).

Summary

In this task we tested an attribute of MSV that has not yet been tested, as far as we know. We tested whether the participants were sensitive to the time frame the MSVs imply. We contrasted between sentences that were felicitous, the verb in the complement was tensed according to the matrix verb, or infelicitous, when the complement was tensed in opposition to the complex verb.

In this task the aTOMics scored very high showing again that their ability to understand the semantic aspect of the verb was not compromised. Still, their scores were significantly lower than those of the TOMers and the controls. It might be that lexical information that is implied in these verbs is less available to these patients.

4.3.5. Interim Summary and Discussion: Mental verbs and aTOMia

We presented 4 different tasks to test different aspects of MSV understanding to the aTOMic group the TOMer group and a group of control participants.

The understanding of factivity was tested in two tasks; in the first we asked participants a truth verification question about affirmative and negative sentences in which the matrix verb was either factive or non-factive.

During the process of composing the first task we found that a large number of items had to be discarded and replaced due to the fact that a large percent of the control participants that took part in two pretests and the final task answered in an unexpected way. This finding highlights the need to consider the division between p-factives and non-factives (Kiparsky & Kiparsky, 1971), or at least the presumption that readers comprehend them as such.

Still, in the few items that the control group answers were compatible with the theoretical argument, the aTOMics did not reach the same high scores as TOMers and controls. We also found that the aTOMic group answered better on the negated non-factives items than on the negated p-factives. They answered to the negated non-

factives as well as the controls. But, in answering to the negated p-factives they erred significantly more than controls. We conclude from these results that aTOMic patients have trouble when they need to consider a conflict between the state of affairs described in the complement and the negated matrix verb.

The second task presented to the participants strengthened the results of the first study. We found that the p-factive predications in this task which included contradictions to the complement of the sentence using 'but', were not corrected by the aTOMic patients as much as they were corrected by the TOMers and controls. This task too shows that the aTOMics have difficulty when they are confronted with a need to consider two opposing viewpoints regarding the same situation. One aspect they should have considered in this task is a statement about the content of some thought, which appears in the complement, and the second is its contradiction.

The other two tasks presented tested two different lexical-semantic aspects of MSV understanding. In one we contrasted two statements and the participants had to choose which signals a higher level of certainty. In this task, the aTOMics answered as well as the TOMer group, but both RBD groups scored less than the control group. The overall high scores of the all RBD patients on the mental predications point to the conclusion that their ability to compare between the two MSVs when presented separately is not compromised due to loss of TOM.

The last task tested whether the subjects were sensitive to the tense the MSVs are implying. We contrasted between sentences that were felicitous, the verb in the complement was tensed according to the matrix verb, or infelicitous, when the complement was tensed in opposition to the complex verb.

In this task the aTOMics scored very high showing again that their ability to understand the semantic aspect of the verb was less compromised. Still, their scores were significantly lower than those of the TOMers and the controls. The reason for this difference might be that the aTOMics, because they lack sensitivity to the point of view of the speaker, are willing to accept sentences that include two points of views regarding the occurrence of the reported event, one that is represented by the mental matrix verb and the other by the embedded verb.

In sum, the results show that aTOMic patients have difficulty in understanding certain aspects of MSV. While they understand their semantics and can infer about different certainty levels they denote, when they need to entertain a contradiction between two conflicting situations regarding the mental realm, they have difficulty inferring and understanding the consequence of the situation.

4.4. Syntactic abilities

In this chapter we will present a number of tests that were used to assess the syntactic abilities of the right brain damaged (RBD) participants.

The first domain we focused on in the assessment of the syntactic abilities of the RBD participants was the comprehension and production of sentences derived by Wh-movement. The first structures we tested were relative clauses. Relative clauses proved to be sensitive to syntactic impairments in various populations. Comprehension and production of relative clauses were found to be impaired in agrammatic aphasia (Friedmann, 1998; Friedmann & Shapiro, 2003; Grodzinsky, Piñango, Zurif, & Drai, 1999), children with syntactic SLI (Friedmann & Novogrodsky, 2004; Novogrodsky & Friedmann, 2006), orally-trained children with hearing impairment (Friedmann & Szterman, 2006), and children with Down Syndrome (Thordardottir, Chapman, & Wagner, 2002), as well as normal children in the first stages of language acquisition (Friedmann, Belletti, & Rizzi, 2009). We also tested the RBD patients' ability to comprehend two types of Wh-questions, *which subject* and *which object. Which object* questions, like object relative clauses, are derived by wh-movement of the object across the subject. The comprehension of these structures was found to be difficult for agrammatic aphasia patients (Friedmann & Shapiro, 2003), for children with S-SLI (Friedmann and Novogrodsky (in press), for hearing impaired children who had difficulty in *object relative* comprehension (Friedmann & Szterman, 2006; Haddad-Hanna & Friedmann, 2009; Nave, Szterman, & Friedmann, 2009). This task was also important because it served as a direct measure of the RBD patients' syntactic and discoursive abilities. The details of this measurement will be explained in section 4.4.3.

Another domain we tested was the participants' ability with respect to the syntactic constraints on pronouns (Chomsky's 1981 binding principles). This ability was considered important because we found that aTOMic patients had difficulties in using pronouns as referring devices in discourse. We wanted to test whether this difficulty is related to their syntactic abilities, and to tease apart syntactic and TOM sources for the difficulty in using pronouns. Binding comprehension is another domain that was found to be sensitive to syntactic impairments in aphasic patients (Ruigendijk, Vasić & Avrutin, 2006; Vasić, 2006). Typically developing Hebrew-speaking children also

showed late development of this ability, which seemed to be mastered only around the age of 6 (Ruigendijk, Friedmann, Novogrodsky, & Balaban, 2010).

A further syntactic ability we tested was the ability to produce complex sentences that include sentential complements embedded to a mental state verb. This ability was considered important for our research because some researchers claimed that TOM ability, specifically the acquisition of false belief understanding, is dependent upon the prior acquisition of this syntactic ability (J. de Villiers, 2007; J. de Villiers & Pyres, 2002; de Villiers & de Villiers, 2000). We expected that if RBD patients suffered syntactic difficulty it would surface in one of these tests and structures.

4.4.1. Comprehension of relative clauses

Comprehension of relative clauses was assessed using *a binary sentence-picture matching task* (Bambi ZTI, Friedmann, 1998; Friedmann & Novogrodsky, 2002).

Participants

The participants in this test were 18 RBD patients, 11 aTOMic patients, and 7 TOMer patients. See Table 2 for detailed list of the patients who participated in this task.

Materials and Procedure

The task comprised of 60 semantically reversible sentences and 20 picture pairs, which were shown three times. The sentences were presented in a random order: 20 simple SVO sentences (e.g., example 52), 20 subject relative clauses (e.g., 53), and 20 object relative clauses (e.g., 54). Because Hebrew verbs agree with the subject in gender, number and person, all sentences included characters of the same gender and number, in order to preclude an agreement cue on the verb.

(52) ha-yeled mesaben et ha-pinguin.
The-boy soaps ACC the-penguin
'The boy is soaping the penguin'.

(53) tare li et ha-yeled she-mesaben et ha-pinguin.
Show me ACC the-boy that-soaps ACC the-penguin
'Show me the boy that is soaping the penguin'.

(54) tare li et ha-pinguin she-ha-yeled mesaben.
Show me ACC the-penguin that-the-boy soaps
'Show me the penguin that the boy is soaping'.

While the participant heard a sentence, she was shown two pictures (e.g., Fig. 16) and was asked to point to the one that matched the sentence she heard. In one picture the roles matched the sentence; in the other the roles were reversed.

Figure 16. An example of a picture pair presented in the relative clause-picture matching task.

Results and discussion

The scores of the RBD patients in this task are presented in Table 11. The results in show that all RBD patients performed well in this task. The average score of the average between the two kinds of relative clauses was 96% (SD = 4.5%).

Table 11. Percentage of correct responses on relative clause comprehension

	Simple	Subject Relative	Object Relative
Tzipora	100	100	95
Dafna	100	95	90
Abraham	95	85	85
Sason	90	95	100
Arye	95	95	90
Dror	95	100	90
Jacob	100	100	95
Daniel	95	95	85
Sachar	100	100	100
Simon	100	100	95
Gila	100	100	80
Yaron	95	100	95
Tzvi	100	100	100
Sigalit	100	100	100
Moshe	100	100	100
Sharon	100	100	100
Ahuva	95	100	100
Ayal	100	100	100

*The results of the a-TOMic patients appear on shaded background.

All the participants performed above chance level on all sentence types (using a binominal test, $p < .05$). These results show the RBD patients do not show any difficulties comprehending relative clauses. These results are in sharp contrast to patients who suffer damage to their left frontal hemisphere and are diagnosed with agrammatic aphasia, who comprehend object relatives only at chance level (Friedmann & Shapiro, 2003; Grodzinsky et al., 1999). Poor performance on these structures is also found for children with S-SLI at the end of their elementary school years, while normal 6 years old perform above chance level (Friedmann & Novogrodsky, 2004). Orally-trained children with hearing impairment also perform significantly lower on this task compared to a control group of younger children

(Friedmann & Szterman, 2006). So, whereas the comprehension of object relatives is a very clear and sensitive clinical marker for syntactic impairment in various populations, the individuals with right brain damage who participated in the current study showed normal performance in this structure, indicating preserved syntactic abilities.

4.4.2. Production of Relative Clause

Relative clause production was tested using a preference elicitation task (BAFLA ADIF, Friedmann, 1998; Novogrodsky & Friedmann, 2006).

Participants

The participants in this test were 14 RBD patients, 7 aTOMic patients, and 7 TOMer patients. Seven control patients were also tested. See Table 2 for detailed list of the patients who participated in this task.

Materials and Procedure

The participants heard a sentence and were asked to answer a preference question that required the production of a relative clause. Half of the questions elicited a *subject relative clause* (55) and the other half, an *object relative clause* (56). The sentences and question were presented in either feminine or masculine inflection, according to the participant's gender.

(55) There are two wo/men, one wo/man is reading a newspaper, the other wo/man is reading a book. Which wo/man would you prefer to be? Please start your answer with "The wo/man.."

(56) There are two wo/men, the doctor is examining one wo/man and a nurse is examining the other wo/man. Which wo/man would you prefer to be? Please start your answer with "The wo/man..".

The task included 12 sentences, 6 were *subject relative clause,* and 6 *object relative clause*. The sentences were presented in a semi-random order.

Results and discussion

The results of the RBD patients and the average results of the control group are presented in Table 12.

Table 12. Percentage of correct responses on relative clause production

	Subject Relative	Object Relative	Average
Tzipora	100	100	100
Dafna	100	100	100
Abraham	100	40	70
Arye	100	50	63
Dror	67	60	67
Jacob	100	33	75
Gila	83	80	82
Yaron	100	80	90
Tzvi	100	100	100
Sigalit	100	100	100
Moshe	100	60	80
Sharon	100	100	100
Ahuva	100	100	100
Ayal	100	100	100
Controls (n=7)	98	100	99

*The results of the a-TOMic patients appear on shaded background.

The total average score of the RBD group was 86.8% (SD = 15.2%). The controls total average score was 98.8% (SD = 3.1%). The RBD's scores on subject relative clause was very high (average = 96.4%, SD = 9.6). Their production of object relative clause (which is the condition that other populations who suffer syntactic difficulties show low performances at), was lower (average = 78.8%, SD = 25.2%). There were participants in both the aTOMic and the TOMer groups who had difficulties in this task. However, crucially, 7 of the participants, 5 TOMers and 2 aTOMics produced 100% of the target sentences correct. Two other participants, one TOMer and one

aTOMic, reached the high score of 80% correct. The aTOMic and TOMers together produced 160 responses in this task, only two were ungrammatical, the others were inappropriate for different reasons. In contrast, patients who suffer left hemisphere damage who were diagnosed as aphasic had severe difficulty even in producing subject relative clause in a repetition task, producing only 33% correct repetitions and an average of 22% correct responses in an elicitation task (Friedmann, 2006).

Because of the crucial finding that almost no ungrammatical relative clause was produced, we suggest that some of the difficulty the aTOMics experienced was due to the complexity of the task, and possibly to its pragmatic demands. The participants were asked to imagine which character of the two presented to them by the experimenter they prefer to be. This task might be difficult not because the syntactic demand but because these patient might have difficulty in imagining herself or himself as someone else in different social situations, as the items in the task demand. Therefore, in hindsight, other tests in which the participants can demonstrate their ability to produce relative clauses by describing figures in pictures (Friedmann & Szterman, 2006; Novogrodsky & Friedmann, 2006) would have been more appropriate for this group in the current study.

4.4.3. Comprehension of wh-questions

The ability of the RBD patients to comprehend Wh questions was assessed using a task in which the participants had to choose the right figure from 3 figures in the picture.

Participants

The participants in this task were 7 RBD patients, 5 of them aTOMic and 2 TOMers. See Table 2 for detailed list of the patients who participated in this task.

Materials and Procedure

Twenty pictures were presented, each picture included three figures: two of the same type which differed in at least one feature (e.g., a blue and a purple elephant; a blond and a red-head girl) and a third figure of a different kind. In the picture, the first figure was performing an action on the second, and the second figure was performing the

same action on the third figure, which was of the same type of the first one (Figure 17). Each picture was presented twice, each time with a different type of question: 20 *which subject* questions (see 57 below) and 20 *which object* questions (see 58 below).

(57) eyzo yalda melatefet et ha-ima?

Which girl caresses ACC the-mother?

Which girl caresses the mother?

(58) et eyzo yalda ha-ima melatefet?

ACC which girl the-mother caresses?

Which girl does the mother caress?

Figure 17. An example for a picture used in the question comprehension task

The experimenter asked a question while the participant was looking at a picture. The participant was then requested to point at the figure that answered the question.

Results and discussion

The scores of the RBD patients who participated in this task are presented in Table 13.

The total average result was 98.7% (SD = 3.5%). All the patients, except one aTOMic patient, performed this task perfectly.

Table 13. Percentage of correct responses on *subject* and *object* wh-questions comprehension

	Subject	Object
Tzipora	100	100
Dror	94	88
Daniel	100	100
Sachar	100	100
Gila	100	100
Sharon	100	100
Ayal	100	100

*The results of the a-TOMic patients appear in shaded background.

These results stand in direct contrast to the task producing appropriate descriptions (section 4.2.1.2.) in which the same materials (pictures and questions) were used, but the task was different. In the task described here the participants were requested to point at the appropriate figure, in order to test the syntactic comprehension ability. In the discoursive task the assignment was to describe the character. That assignment tested the ability to choose the appropriate description considering the context the figures appeared at and the aTOMics had difficulty completing that task.

4.4.4. Comprehending Binding Principles

Chomsky's (1981) Binding theory presents two principles that structurally regulate the relations between pronominal elements and their antecedents within the same sentence. The first principle determines the constraint on reflexive interpretation and the second determines the constraints on pronouns interpretation:

A: A reflexive is bound in its governing category.

B: A pronoun is free in its governing category.

An element is 'bound' if its antecedent c-commands it and is co-indexed with it. 'Free' means not bound. The 'governing category' is the smallest maximal projection (IP or NP) which contains the pronoun/reflexive, its governor and an accessible subject. In other words, these principles state that reflexives must have a local antecedent on which they depend referentially, and that pronouns must be locally free. The pronouns' antecedents are found outside the clause they appear in. For example, in (59) the reflexive is locally bound and in (60) the pronoun is locally free (because the antecedent of *him* is outside the clause the pronoun appears):

(59) The king said the doctor$_i$ pinched himself$_i$.

(60) The king$_i$ said the doctor pinched him$_i$.

Patients who suffer damage in their left hemisphere and are diagnosed with agrammatic aphasia (Broca's aphasia) have been found to be impaired in reference assignment. Studies indicate that their interpretation of pronouns is impaired while their reflexive interpretation was unimpaired. Thus, when subjects err they allow a pronoun to co-refer with a local antecedent, which is the closest to it (Ruigendijk, Vasić & Avrutin, 2006; Vasić, 2006). In other words, they are missing condition B of Chomsky's binding principles.

In order to test the possibility that RBD participants also have difficulty in reference assignment within a sentence we presented a task that included reflexives and pronouns in three kinds of sentence structures (coordination, subject relative clause, sentential complement). The relative clause sentences are interesting because in this construction the pronoun's antecedent is the closest NP, while the reflexive's antecedent is the farthest (61). In the other conditions the distance between the reflexive and its antecedent is the smallest (62) (63).

(61) *The penguin* that washed the boy soaped him/ *himself* .

(62) The boy and the penguin met and then the *penguin* washed him/ *himself* .

(63) The boy said that the *penguin* washed him/ *himself*.

This task, evaluating the syntactic ability of the RBD patients to distinguish between the use of pronouns and reflexives, according to the binding principles is used here to explore the connection between syntactic and discoursive abilities. As described before (4.1.3.), we found that the aTOMic patients had difficulties in using referring expressions within discourse. In this task we test whether these difficulties stem from syntactic difficulties.

Participants

The participants in this test were 20 RBD patients, 13 aTOMic patients, and 7 TOMer patients. See Table 2 for detailed list of the patients who participated in this task.

Materials

The comprehension of pronominal expressions was tested using a sentence-picture matching task. The test included 72 sentences, half included a pronoun and half - a reflexive. The test included 24 coordinated sentences (see 64 below), 24 were sentences with sentential complement (see 65 below), and 24 were relative clause sentences (see 66 below). The sentences were presented in a semi-randomized order. Examples of the three kinds of sentences are presented below:

(64) Ha-yalda ve-ha-jirafa nifgeshu ve-az ha-yalda raxca *et acma* / *ota*

'The girl and the giraffe met and then the girl washed herself / her'

(65) Ha-yeled mesaper she-ha-penguin mesaben *et acmo/ oto*.

'The boy said that the penguin is soaping himself/ him'

(66) Ha-penguin she-raxac et ha-yeled siben et *acmo/ oto*

'The penguin that washed the boy soaped himself/ him'

The two NPs in each sentence always had the same gender in order to preclude gender agreement cue for the choice of the antecedent. The sentences were presented together with one page that included two pictures, one matched the sentence and the other presented an incorrect antecedent as the object, as shown in Figure 18.

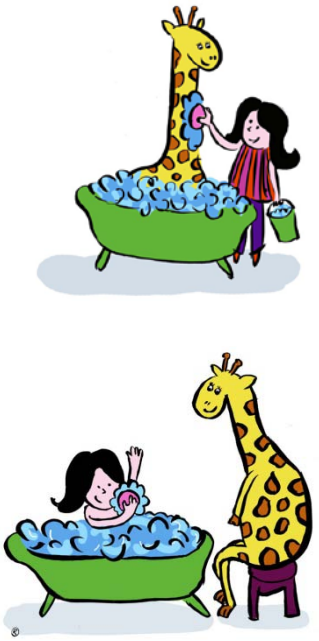

Figure 18. Example of a picture pair in the binding comprehension task

Procedure

The sentences were orally presented in a neutral intonation and normal rate. The sentences were presented without time limitation, and were repeated as many times as requested or needed. The participants were asked to choose the picture that matched the sentence they heard. If a participant changed her response, only the final answer was included in the calculation. The items were presented in a semi random order after two practice trials.

Results and discussion

Coding: correct responses were graded a score of (1) and incorrect responses a score of (0). Note that RBD patients often suffer neglect, a neurological deficiency which causes them to disregard their left visual field (see Table 1 for detailed information about the participants). Therefore, while coding the answers we disregarded the mistakes that were due to this problem. In other words, if a patient gave an incorrect answer when the appropriate answer appeared on the leftmost side of the page, that answer was discarded. There were 5 pictures (out of 18) in which one figure appeared in the left side of the page. There were 18 mistakes of this kind, which were made by

5 patients from the aTOMic group that had neglect (out of 888 answers given by the aTOMics all together).

The average score of both groups for both kinds of referring expressions are presented below in Table 14.

Table 14. Percentage of correct responses to pronouns and reflexives by the aTOMic and TOMer groups

	Pronouns	Reflexives
aTOMic	96	98
TOMer	100	100

The total average score of the aTOMic group was 96.7% (SD = 3.6%). The total average score of the TOMer group was 99.6% (SD = 0.7%).

The results in Table 14 show that the aTOMic group reached high scores in comprehending both kinds of referring expression, all the participants performed above chance level in both kinds of expressions (using a binominal test, $p < .05$). These scores show that the brain damage these patients suffer does not create a syntactic difficulty in comprehending the appropriate use of pronouns reflexives.

If we consider again the results of chapter 2 regarding the use of pronouns as referring devices in discourse we can conclude that the difficulty found there is not related to a syntactic disabilities. Whereas both groups of RBD patients performed over 95% correct in the syntactic tasks presented, the participants in the aTOMic group were unable to use pronouns appropriately only in the discoursive context and tasks.

4.4.5. TOM and sentential complements

One aspect of the relations between TOM and linguistic abilities that received particular attention is the connection between acquisition of complement construction and mastering false belief tasks. Several findings led researchers (J. de Villiers, 2004, 2007; J. de Villiers & Pyers, 2002; Hale & Tager-Flusberg, 1999) to suggest there is a

unidirectional process of development, namely, only once children are able to represent embedded proposition they become capable of representing false beliefs.

Although this discussion concerned the process of acquisition, the results of the current study can be relevant because they shed light on the dependency between the use of complement construction and false belief understanding. More precisely, whether adults who suffer loss to their TOM ability also lose the ability to create and understand such sentences.

During the fourth year of life children master important abilities both in language and in cognitive development. The ones under focus here are the ability to master sentence forms involving mental state verbs and their sentential complements and the ability to master first order false belief tasks. J. de Villiers and Pyers (2002) conducted a longitude study that followed 28 children from age 3;1 and tested them in four sessions along one year. The children were presented different version of first order false belief tasks (e.g., unseen displacement and unexpected content) and a number of linguistic measures (some derived from spontaneous speech and others from structured tasks). One of the tasks was a test of memory for complements. The children were shown pictures that illustrated short stories about a character making a mistake, telling a lie or holding a false belief. Half the items described acts of communication (e.g., say, tell) and half acts of thinking (e.g., think, believe). The children were asked to report the content of the character's belief or statement (i.e., what did he say/ think). A second linguistic task tested children's understanding of medial answers to wh-questions. They were presented short stories about an event that occurred at a certain time (e.g., a girl tore her dress in the afternoon) and a later reporting of the event (e.g., the girl telling about it to her mother that evening) the children were asked: *When did the girl say what she ripped?* The researchers interpreted a wrong answer (e.g., that afternoon) as indicating that the children failed to subcategorize the complement under the verb. As expected the overall success rate in all tasks increased during the year of study. The researchers also found significant correlations between memory for complements (for both communication and mental verbs) and all false belief tasks but not between the performance on the wh-question task and false belief tasks. In their analysis the researchers distinguished between children who passed the memory for complementation test and those that passed the false belief tasks and found that 5 out of the 6 children that failed syntax (performed

less than 10 out of 12 correct) were able to pass false belief (performed more than 5 out of 6 correct) and 13 out of the 23 children that passed the syntax task, failed false belief. This difference was significant. This showed that the children acquired the ability to retain complement structure before they could pass false belief tasks. The researchers also conducted a regression analysis to test which linguistic feature (e.g., length of MLU, using complements in spontaneous speech, answering wh-questions and retaining the information in the complements) was the best predictor of future success on false belief tasks. They found that the best predictor was the success in the memory for complements test. But the reverse was not found. Success on false belief task did not predict language ability as measured in a later round of testing.

The same pattern of results was found in testing deaf children whose language acquisition and false belief understanding was delayed considerably relative to their peers (de Villiers & de Villiers, 2000). In this population too, the best predictor for success on false belief understanding was the success in producing sentential complements of mental state verbs. These findings led the researchers to conclude that the acquisition of complement structure is a perquisite for understanding false belief. They conclude:

> *"The point here is that a child who fails to retain the appropriate syntactic representation for a complement construction will not have it available as a form of mediating representation for false-belief-understanding. If false-belief-understanding depends on such a representation, the child will then fail false-belief tasks. Our data above confirm this prediction…..we wish to argue that mastery of the linguistic structures bridges the two strands of development, the grammatical developments that encode events, and the understanding of others' behavior and the construction of a theory about it. The structural achievement of complements in language provides a representation for encoding false beliefs: with this representation the child can build the explicit theory on which such reasoning depends*." (*J. de Villiers & Pyres, 2002, pp. 1057-1060).*

The researchers also state that they do not claim all understanding of TOM relies on this syntactic acquisition, but they do claim that it is a prerequisite for achieving false belief understanding (J. de Villiers, 2007; J. de Villiers &Pyres, 2002).

Another study that demonstrated the importance of complement structure acquisition to false belief understanding was conducted by Hale and Tager-Flusberg (2003). Their study was a training study in which preschoolers who failed false belief tasks and understanding sentential complements were trained on both these tasks. The false belief task they were tested and trained on was one of unseen displacement. In the sentential complement task the children were shown pictures of a character doing something and saying she is doing something else. The children were asked what the character said she was doing. The children were randomly assigned to one of three training groups. They were either trained on false belief tasks, sentential complements or on understanding relative clause, as a control condition. The children attended two training sessions within a week, in each, 4 trails of each of the tasks were presented. During the false belief training the children were asked the anticipation question about where the character will look for the item that was moved and they were given either corrections or positive feedback. In the sentential complement training the children were told and shown a story of a boy performing some action on one figure and saying he performed it on a different one. They were asked either *What did he say?* Or *Who did he say he __?* Here too corrections were offered. Three to five days after the last training session all the children were administered a post-test which included a TOM task of displacement similar to the one used in the practice sessions and two other false belief tasks (change of content and appearance reality test). They were also administered a similar test of sentential complements. The results showed that the group trained on false belief tasks improved dramatically on these tasks only, but the group that was trained in understanding sentential complements improved dramatically both in their syntactic ability and in their performance on false belief tasks. The control group that was trained on relative clause sentences did not improve on TOM post tests. The group that trained on relative clause sentences improved on those only. These results strongly support the claim put forward by J. de Villiers and Pyers (2002) that the acquisition of complements is strongly related to the ability to represent false belief situations. Note that it also showed that children who were trained on a certain kind of false belief are able to use their acquired knowledge and represent similar situations although their improvement on understanding sentential complements has not improved significantly. So the training on false belief seems to allow a bypass for the need to first acquire the syntactic knowledge of complements.

One important note relevant to both studies above was highlighted by Tomasello and Lohmann (2003) who noticed that in the task of sentential complementation comprehension the children were shown and practiced with situation of deception or mistakes. The character was doing one thing but reporting she was doing something else, or admitting a mistake she has made. This can explain (at least partly) the improvement in performance in false belief tasks after practicing sentence complements, an explanation that does not rely on syntactic acquisition.

The researchers suggested that a proper control condition to this option will be one in which children were presented with a deceptive scenario but without any linguistic description at all. Such a control condition is necessary to separate between deceptive experience and language and thus to determine whether the experience itself, with no linguistic mediation at all does not create an understanding of false belief. The researchers conducted a training study to test other possibilities of connection between the ability to master false belief understanding and language acquisition. One important connection is the learning of mental state terms and concepts (Astington & Jenkins, 1999; Olsen, 1998). The second is discoursive experience children acquire while interacting with other people. Understanding beliefs are mental states that can be false and contradictory or alternative to others' beliefs is dependent on experience in different discursive situations. In these interactions children experience misunderstandings and clarification about people’s differing perspectives or understandings of situations (Harris 1996; Siegal, 1999).

So, in order to broaden the possibilities for understanding what kind of training can aid understanding of false belief Lohman and Tomaello (2003) presented different conditions of training: one was of full discourse – they showed the children deceptive objects and talked about them using mental terms while using complement structures. The second was discourse only: the deception was explained without mental terms and complement structures. The third was a situation of no language: a deception was shown and the only words used were 'Look – see this" - no further elaboration was given. In the fourth condition the sentence complements were highlighted and there was no mention of the deceptive aspect of the situation. The experimenter raised questions about the information in the complements (e.g., what do you think, what does the puppet think). The children were engaged in giving complements to mental

state verbs. After the training sessions the children were post-tested on three different false belief tasks and sentence complement understanding. The results showed that the full discourse training group outperformed each of the other groups on the post test task. They also found that training of sentential complements which include mental state predicates as matrix verbs was sufficient by itself to facilitate children's understanding of false belief. The researchers conclude:

"The current findings are thus the strongest evidence (at least using a training methodology) that linguistic experience is a strong facilitator, perhaps even necessary condition, in the development of children's false belief understanding… availability of sentential complement syntax as a representational format both seem to make independently important contributions to the ontogenetic process" (p. 1139)

So in both studies a strong connection between understanding sentence complements and the acquisition of false belief understanding was evident.

The data from the current study cannot directly contribute to the discussion of acquisition of false belief understanding. But we were interested to study whether the aTOMic patients, who had an acquired impairment, produced less sentential complements of the kind discussed above; whether the strong correlation between false belief understanding and sentential complement understanding is evident in any way in this population.

To do this we analyzed the sentences produced in the justification given by all the participants in the a-TOM battery (see Appendix). We compared the rate the aTOMics used the complement structure (sentential complements to mental verbs) to the rate of use by the other two groups, the TOMers and controls.

Participants

The participants in this test were 26 RBD patients, 17 aTOMic patients, and 8 TOMer patients. The control group included 14 adults. See Table 2 for detailed list of the patients who participated in this task.

Materials and Procedure

The aTOMic battery (Chapter 3) included 8 kinds of tasks that test different aspects of TOM. In seven out of the eight categories the participants were asked, in addition to a yes/no question, for a justification of their answers. We used these justifications to evaluate the syntactic abilities of the participants with respect to sentential embedding. We counted how many sentential complement (SC) embedded to a mental verb occurred in the justifications for each participant. We considered in this analysis only the justification that were coded as *appropriate justifications* (see chapter 3) because only when a participant comprehended the aTOMic aspect of the item presented we can expect the use of this syntactic construction. To arrive at the percent of sentential complements used we calculated the number of answers with SC out of the total number of answers that were coded as appropriate justifications given by each participant to each of the items.

Results and discussion

The results are presented in Figure 19.

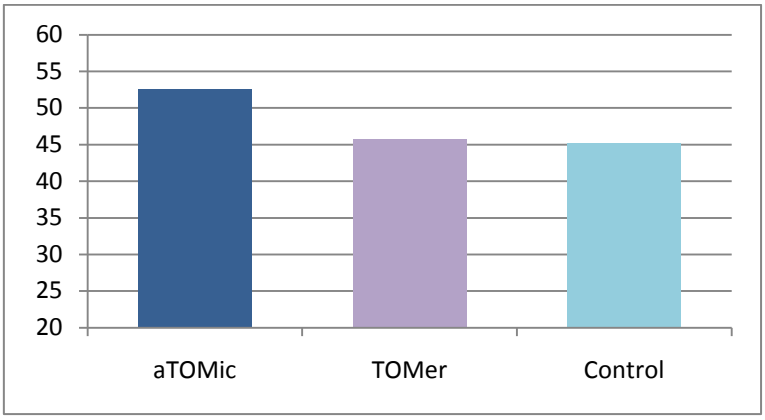

Figure19. Percent of sentential complements used in the justifications given in reply to the aTOMic battery

The results showed that the aTOMic patients produced on average (52.5%, SD = 21%), the TOMers (average = 45.7%, SD = 20.6%) and controls (average = 45.1%, SD = 13.3%) produced less SC but the differences between the three groups were not significant. Between the aTOMics and TOMers (U = 47.5, p = .22), between the aTOMics and control (U = 86, p = .2) and between TOMers and controls (U = 56.5, p = .5). The results of this study showed that there was no significant difference

between the loss of TOM and a tendency to create less SC. There were no ungrammatical sentences or cases of omission of the complementizer.

So, although significant differences were found between the ability of the different patients to answer and justify the answers to the TOM tasks appropriately, no significant differences were found in the syntactic composition of their answers.

These results point again to a dissociation between syntactic abilities which are not compromised due to the brain damage these patients suffered, and a decreased ability to put this syntactic knowledge to appropriate use during conversation.

4.4.6. Interim Summary and Discussion: aTOMia and Syntactic Abilities

In this chapter we reviewed various syntactic tasks the RBD completed in order to assess their syntactic abilities. We found that both groups of RBD patients, aTOMics and TOMers, did not suffer syntactic disabilities.

In relative clause comprehension the RBD patients achieved almost perfect scores, well above chance level. This task has been proven as a sensitive tool for detecting syntactic difficulties in other population. The RDBs' high scores stand in contrast to the low ability of aphasic patients and SLI children who comprehend object relatives at chance level (Friedmann & Shapiro, 2003; Grodzinsky et al., 1999; Friedmann & Novogrodsky, 2004), and children with hearing impairment who also have difficulty in this task (Friedmann & Szterman, 2006). The RBD patient's ability to produce relative clause sentences was also better than that of aphasic patients (Friedmann, 2006).

The results of wh-question comprehension task are especially important to the comparison between the RBD patients' preserved linguistic abilities and their deficit in TOM related linguistic abilities. In this chapter we described the syntactic test of the ability to comprehend *subject* and *object which* questions. We used a picture selection task in which the participants pointed to one of three figures according to a question they heard. All but one RBD patient received the perfect score of 100% correct in this task. This result stands in contrast to the results of the task presented in section (4.2.1.2.) in which the same pictures and questions were presented but the

participants were asked to describe rather than point to the appropriate figure. In the description task the participants need to consider the other figures in the picture and choose a feature that describes and differentiates between two similar figures. In other words, they had to consider, as speakers, which information will allow their addressees to know which of the two characters they choose. In this task the aTOMics performed significantly poorer than TOMers and control participants (for details see section 4.2.1.2). Thus, they have the syntactic ability that allows them to understand wh-questions, but they lack the language tools to communicate which is the exact figure they thought was the correct answer to the question.

The third domain we tested was the RBD patients' ability with respect to the syntactic constraints on pronouns (Chomsky's 1981 binding principles). We found an important dissociation between their preserved ability to apply the restriction of the (syntactic) binding principles and their inability to use pronouns appropriately as reference terms in discourse (see section 4.2.1. and 4.2.2.). Their performance on the binding principles task was above 95% in correctly comprehending pronouns and reflexives. This result is much higher than that of aphasic patients in a similar task (Ruigendijk, Vasić & Avrutin, 2006; Vasić, 2006) and typically developing Hebrew-speaking children around the age of 6 (Ruigendijk et al., 2010).

The fourth syntactic ability we tested was the ability to produce complex sentences that include sentential complements embedded to a mental state verb. We found that between 23.5% to 67.6% of the justifications the RBD patients produced in answer to the different items in the aTOMic battery (chapter 3) included sentential complements. We also found that the rate the a-TOMics used these constructions was not different from the rate the TOMers and the control group used them. From these results we learn that this syntactic ability is not compromised due to the brain damage and the cognitive disability that follows it.

In sum, we found that different key syntactic abilities are preserved in the group of RBD patients we tested. Both syntactic abilities that are known to be difficult for different population suffering syntactic loss or delay (i.e., relative clause comprehension and production) and syntactic abilities specifically related to the TOM ability (binding principles and sentential complements).

5. General Summary and Discussion

The aim of this study was to describe the effect of acquired damage to Theory of Mind on language. In order to achieve this aim we conducted the study in two phases. First, we identified a group of patients that were right brain damage (RBD) and suffered a substantial loss to their TOM. Second we assessed the linguistic abilities of these participants in a long list of language tasks, and compared their performance with the performance of a group of RBD patients who did not show decrease in their TOM, and to healthy controls. The results of each of the tested domains were discussed in the interim discussions (sections 3.4; 4.1.4.; 4.2.3.; 4.3.5.) In this chapter we will highlight the main conclusions and the relations between them.

Our main finding was that damage to TOM has a specific effect on linguistic abilities. The patients who had difficulties in the TOM tasks demonstrated difficulty only in linguistic tasks that rely on the need to consider the point of view of others. At the same time, their syntactic abilities as well as other linguistic abilities that do not rely on considering the state of knowledge of their interlocutor remain intact.

Significant findings were obtained in each of these two phases of the research; The TOM assessment and the language assessment.

5.1. Right brain damage and TOM

The first important finding is that the group of RBD patients who participated in this study was not a homogeneous group with respect to their TOM abilities. Note that the patients that took part in the research were a random group of right brain damaged patients; they were not included in the research according to their clinicians' assessment regarding any pragmatic or TOM difficulties.

We created a battery of tasks (i.e., aTOM battery) in reference to earlier studies that tested TOM in various age groups and populations. The battery included a variety of items, assessing 8 different aspects of TOM. We asked the classic false belief question about a misplaced item, second order false belief questions, questions about *faux pas* situations, and questions about understanding situations of misunderstanding of intentions, and of surprise. The various items offered the participants a chance to

reason about various social situations (e.g., surprises, white lies, misunderstanding, and social embracement). Some of the tasks had been previously presented to RBD patients in different studies using various methodologies (e.g., knowledge gaps: Happe et al., 1999; Tompkins et al., 2008) others were not yet presented, as far as we know (e.g., understanding teaching: Ziv & Frye, 2004). The broad variety of items enabled us to collect a detailed evaluation of the various aspects of TOM ability for each patient.

Two subgroups clearly emerged within the group of participants with right hemisphere damage. One group showed significant difficulty in almost all the categories. The other group showed normal performance in all items and categories (see chapter3). Accordingly, we divided the group of RBD patients. In the first group all the participants performed significantly poorer than the healthy control group. Each of the participants' average score in this group was less than 70% correct (i.e., aTOMic group). The other group did not differ significantly from the healthy control group, and their average scores on the whole aTOM battery were above 80% correct (i.e., TOMer group).

Our ability to identify a decrease in TOM only in some of the right brain damaged patients, is a highly important finding to the study of TOM after RBD. It demonstrates that characterization of participants according to the hemisphere in which they were impaired is too crude. The fact that a patient suffers RH damage can not inform us about the condition of his TOM. Therefore there is a need for a direct TOM assessment. The TOM status of each patient with right hemisphere damage should thus be assessed using a battery like the one used in the current study.

Previous studies showed correlation between damage to the right hemisphere and discoursive or communicative functions (Adenzato & Bucciarelli, 2008; Brownell et al., 1997; Brownell & Stringfellow, 1999; Davis et al., 1997; Johns, 2008; Mar, 2004; Marini et al., 2005; Weylman et al., 1989), others found correlation between the location of the brain damage and specific aspects of TOM abilities (Amodio & Frith, 2006; Berthoz et al., 2002; Bird et al., 2004; Calarge et al., 2003; Frith & Frith, 2003; Saxe & Powell, 2006; Saxe & Wexler, 2005; Siegal & Varley, 2002). The methodology used in the current study allowed the findings about the cognitive function to serve as the basis for the demonstration of other cognitive functions (e.g.,

linguistic abilities) that might be TOM-related but were not necessarily in direct correlation with the general location of the brain damage.

The distinction between aTOMics and TOMers can also explain diverse reports regarding the relations between RBD and loss of TOM. Some researchers found that damage to the right hemisphere led to such a loss. They showed that RBD patients failed to reason and attend to situations which require a capacity to attribute thoughts, beliefs and intentions in order to understand them properly (Brownell et al., 2000; Happé et al., 1999; Winner et al., 1998). But other studies argued against such relations and showed that with careful design of items and the use of tasks that do not burden the patients' linguistic abilities (e.g., sentence verification task instead of justification elicitation) the RBD patients do not show a decrease in TOM (Tompkins et al., 2008; Siegal et al., 1996).

The finding from the current study, that patients who suffered damage to the right hemisphere were not a uniform group, could explain these conflicting reports (Joanette & Ansaldo, 1999; Myers, 2001; Slessor, Phillips, & Bull, 2007). The most important point is that the loss of TOM might be graded and not all patients that show a decrease in TOM are completely unable to reason about the mental realm. Therefore it is reasonable to expect that some tasks that test TOM will be easier than others, although they are both designed to capture the same ability.

The identification of the distinction between RBD patients who were aTOMic or TOMer is also important from a clinical point of view. The aTOMic battery, designed and tested in the current research, can serve as tool for distinguishing between patients who require treatment for TOM difficulty, and other patients, who do not require such treatment (Blake et al., 2002; Penn, 1999; 2000; Perkins, 2005).

One sensitive tool that was found in the current study and can be useful for clinicians is the analysis of the various justification provided by the participants in the aTOM battery. These justifications were revealing; the aTOMics were found to concentrate much more on the physical aspects of the situation than on the mental aspects, whereas the TOMers tended to justify their choices not differently from controls. Such an analysis of a patient's justifications given in the aTOM battery can serve as a basis for initiating treatment.

5.2. aTOMia and Linguistic abilities

We tested the two groups of participants on tasks that probed various linguistic abilities. Two types of linguistic abilities were tested and compared. One type of linguistic abilities tested was linguistic abilities which we predicted that could be affected by TOM impairment. The other type of linguistic abilities are purely syntactic abilities, which have been shown to be impaired in populations with syntactic impairment, and that are therefore sensitive markers for syntactic impairments (Friedmann, 1998; Friedmann & Shapiro, 2003; Friedmann & Szterman, 2006; Friedmann & Novogrodsky, 2004; Grodzinsky et al., 1999; Novogrodsky & Friedmann, 2006). Some of these tests were also instrumental in distinguishing between two possible explanations for a deficit in questions, topic, and focus abilities, as these abilities are syntactically represented in the CP node in the syntactic tree. Tests of the syntactic abilities related to CP enabled us to decide whether it is the CP that is impaired or the pragmatic abilities. We found that the aTOMic patients had difficulty in the TOM-related linguistic tasks only. Crucially, their linguistic-pragmatic difficulty was not accompanied by other linguistic disabilities, and their syntactic abilities, even those that are sensitive to syntactic impairments after brain damage were completely within the normal range. The TOMers' performance also did not differ from that of the control group; they succeeded in both kinds of tasks. This finding demonstrated a clear dissociation between TOM-related linguistic abilities and other, syntactic abilities. When TOM is preserved in spite of the damage to the right hemisphere the linguistic abilities that are TOM related are also preserved.

The linguistic tasks offered more insights and deeper understanding about the nature of TOM impairment. First, we presented tasks designed to test the participants' ability to produce and comprehend definite articles as means of marking NPs' accessibility (Ariel, 1990; 2001). From these tasks we learned that aTOMics disregard the question of whether or not information has been accumulated between the interlocutors during the discourse when applying the definite marker in discourse. Their use of the definite system to mark whether an item was accessible or not was found to be impaired compared to the TOMers and controls. The aTOMics overused the definite articles and therefore improperly introduced new items into the discourse (as if they were

already accessible). They were also insensitive to the need to add description to items in order to allow their proper identification from the context.

In the tasks that tested if the participants were aware of the degree of accessibility of a referent to their addressee, we found again that the aTOMics tended to disregard information status when tailoring their utterances. They failed to differentiate between given or new information by choosing the appropriate referential terms for each level of accessibility. In some cases the overused the low accessibility markers without considering the addressees' ability to recognize to which of the items presented they are referring. In other tasks they overused the low accessibility markers as if no mutual knowledge of the characters presented has been accumulated during the discourse.

In the production tasks the aTOMic speakers tended to treat new information as if it were shared and known to their addressees, both in the task of retelling stories and in the task that asked to generate a proper differentiating description. For example, one of the aTOMic patients, while retelling a story about two fishermen, described the interaction between the characters in a way that made it very difficult to distinguish to which of the characters he is referring to with each pronoun:

Yigal: .. the one that caught less fish told him because you have a bigger net, so he said to him: give me the net, you take the net, so they switched nets. The next day, they sailed again and despite the fact they switched nets he came in with a larger catch. Later he said to him: No, it's because of the location in the sea, lets switch locations...

In the metalinguistic tasks we also found the opposite pattern, when the aTOMics preferred to treat given information as relatively new. They preferred sentences that used a proper name twice over sentences in which the second mention of the referent could be by a pronoun. This demonstrated their preference for overusing too low accessibility term. For example, when asked to choose which of the two sentences bellow sounds better, David, one of the aTOMic patients preferred the second sentence:

(a) Danny went to see *Margalit* and discuss with *her* the economical condition of the market.

(b) Danny went to see *Margalit* and discuss with *Margalit* the economical condition of the market.

And he added to explain his choice: *to discuss 'with her' is when you actually meet each other*.

In sum, we found that aTOMics tend to disregard the status of the information gathered during conversation while tailoring their utterances.

From the tasks that tested the patients' ability to use and comprehend mental state verbs we found dissociation between relatively good performances on the tasks that related to the semantic aspects of the verbs to the tasks that questioned the lexical property of factivity. Only when aTOMics had to consider in parallel two conflicting aspects of a discoursive situation their performances decreased significantly. From this finding we learned that in order to test the abilities of adults (i.e., participants that have already acquired full mastery of their language) there is a need to test more than semantic acquisition. We found that TOM affects certain lexical attributes, while not affecting the semantic aspects. In these tasks too we found that the TOMers performed almost as well as the control group.

5.3. aTOMia and syntactic abilities

Another interesting result of this study is the good syntactic performance demonstrated by all the RDB patients. We tested the RBD patients' syntactic ability using a variety of tasks.

We presented tasks that tested the participants' ability to use and comprehend relative clauses. We found that their performance on relative clause comprehension was almost 100% correct, and so was their performance on production. These results stand in direct contrast to agrammatic aphasic patients, who show, on the same tasks, great difficulty in both comprehension and production of relative clauses (Friedmann, 2006).

We also tested the RBD patient's comprehension of another structure that is a sensitive marker for syntactic deficits in agrammatic aphasia, wh-questions, and again we found almost perfect performance.

We also tested two syntactic abilities which are directly related to the abilities we found to be compromised in the aTOMic group. One was a test of the binding principles which constrain the distribution of pronouns and reflexives within sentences (Chomsky, 1981). This ability can been seen as complementary to the use of pronouns within discourse which was found to be difficult for aTOMic patients (see section 4.2.). Thus, the two aspects of pronouns use – the syntactic and the discoursive, TOM-related aspects, can be separately assessed. Whereas in the assessment of the discoursive aspects of pronouns we found that the participants with TOM impairment were severely impaired, these patients succeeded in the task assessing the syntactic aspects of pronoun use. All the participants had very high scores in the syntactic binding assessment, showing their intact syntactic ability which stands in contrast to their inability in the discursive tasks.

The other syntactic ability we tested that was related to TOM was the use of sentential complements. Some researchers studying children's acquisition of TOM and syntactic abilities (J. de Villiers, 2007; J. de Villiers & Pyres, 2002; Lohman & Tomasselo, 2003) showed a dependency between the acquisition of sentential complement clauses and the performance on false belief tasks. Our study showed that when TOM impairment is acquired (rather than developmental), the ability to use sentential complements is not affected. When aTOMic patients did give appropriate mental justifications while answering the aTOM battery (all patients gave some appropriate mental justifications), they used sentential complements to mental verbs at the same rate the TOMers and controls did.

These findings are important for the discussion about dissociation between TOM and other linguistic abilities (Ariel, 2008, forthcoming; Blake et al., 2002; Sperber & Wilson, 2002). We found that damage to TOM does not result in a general linguistic impairment. The linguistic difficulties that the aTOMic RBD patients show are all related to their TOM impairment.

This dissociation can be demonstrated by comparing results we gathered to a test that was administered in two ways, thus creating two tasks, one syntactic and one TOMic. The syntactic test was of comprehending wh-questions (section 4.4.3.). We presented pictures of three characters performing some act on one another (e.g., see Fig. 7 and 17) and asked a wh-question about the subject or the object. The participants were

asked to point at the appropriate figure in the picture. All RBD patients performed this task almost perfectly. In the TOMic version of this task (section 4.2.1.2.) we asked wh-questions about the same pictures, however, in this task, we asked the participants to give a description of the character rather than to point at it. This task demanded they produced a description that highlights some feature in the picture that differentiates one of two similar figures so their addressee (in this case the experimenter) could identify which character they intended her to pick out (see Fig. 7 and 17). In this task the aTOMics performed significantly poorer than the TOMers and controls.

In accordance to the framework we presented in the Introduction (section 1.1.1.) we conclude that the linguistic difficulties aTOMic patients face are all and only pragmatic difficulties (Sperber & Wilson, 2002; Wilson, 2005). We did not find any evidence for a general decrease in other linguistic abilities due to RBD.

In sum, this study demonstrates that linguistic abilities may sometimes be compromised by brain damage, although direct damage to the linguistic system of RBD patients was not evident. This damage is indirectly caused when linguistic performance builds on cognitive abilities, such as those relying on TOM. We look forward to implementing these findings in the clinical field in order to help the special population of TOM deficient RBD patients we identified and described in the current study.

References

Abbott, B. (2005). Definite and indefinite. In K. Brown (Ed.) *Encyclopedia of Language and Linguistics*, 2nd ed. (pp. 392-399) Oxford: Elsevier.

Adenzato, M., & Bucciarelli, M. (2008). Recognition of mistakes and deceits in communicative interaction. *Journal of Pragmatics, 40,* 608-629.

Aichhorn, M., Perner, J., Weiss, B., Kronbichler, M., Staffen, W., & Ladurner, G. (2008). Temporo- parietal junction activity in Theory of mind tasks: Falsness, beliefs or attention. *Journal of Cognitive Neuroscience, 21,* 1179-1192.

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience, 7,* 268-277.

Amsterlaw, J., Lagattuta, K. H., & Meltzoff, A. N. (2009). Young children's reasoning about the effects of emotional and physiological states on academic performances. *Child Development, 80,* 115-133.

Apperly. I. A., Sampson, D., & Humphreys, G. W. (2009). Studies of adults can inform accounts of theory of mind development. *Developmental psychology, 45,* 190-201.

Ariel, M. (1990). *Accessing Noun Phrase Antecedents*. London: Routledge.

Ariel, M. (2001). Accessibility theory: An overview. In: T. Sanders, J. Schilperoord & W. Spooren (Eds.), *Text Representation: Linguistics and Psycholinguistics Aspects* (pp. 29-87). Amsterdam: Benjamins.

Ariel, M. (2008). *Pragmatics and Grammar*. Cambridge: Cambridge University Press.

Ariel, M. (Forthcoming). *Defining pragmatics*. Cambridge: Cambridge University Press.

Avram, I., & Armon-Lotem, S. (2005). The autonomous contribution of syntax and pragmatic: the acquisition of the Hebrew definite article. In: A. M. Di Sciullo and R. Delmonte (Eds.), *Universal grammar and the external systems* (pp. 169-182). Amsterdam: John Benjamins.

Arnold, J. E., Benneto, L., & Diehl, J. J. (2009). Reference production in young speakers with and without autism: Effects of discourse status and processing constraints. *Cognition, 110,* 131-146.

Astington, J. W. (1998). Theory of mind, Humpty Dumpty, and the icebox. *Human Development, 41,* 30-39.

Astington, J. W., & Jenkins, J. M. (1995). Theory of mind development and social understanding. *Cognition and Emotion, 9,* 151-165.

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relationship between language and theory of mind development. *Developmental Psychology, 35,* 1311-1320.

Astington, J. W., Pelletier, J., & Homer, B. (2002). Theory of mind and epistemological development: The relation between children's second order false belief understanding and their ability to reason about evidence. *New Ideas in Psychology, 20,* 131-144.

Baltaxe, C. A. M. (1977). Pragmatic deficits in the language of autistic adolescents. *Journal of Pediatric Psychology, 2,* 176-180.

Baltaxe, C. A. M., & D'Angiola, N. (1992). Cohesion in the discourse interaction of Autistic, Specifically Language impaired and normal children. *Journal of Autism and Developmental Disorders, 22,*1-21.

Baron-Cohen, S., Cox, A., Baird, G., Swettenham, J., Nightingale, N., Morgan, K., Drew, A., & Charman, T. (1996). Psychological markers in the detection of autism in infancy in a large population. *British Journal of Psychiatry, 168,* 158-163.

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advances test of theory of mind: Evidence from very high functioning adults with autism or Asperger syndrome. *Journal of Child Psychology and Psychiatry,* 38, 813-822.

Baron- Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic-child have a theory of mind. *Cognition, 21,* 37-46.

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with

Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders, 29*, 407-418.

Baron-Cohen, S., Ring, H., Moriarty, J., Schmitz, B., Costa, D., & Ell, P. (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.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "reading the mind in the eyes" test revised version: Study with normal adults, and adults with Asperger syndrome or high functioning autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42,* 241-251.

Beeman, M. J., Bowden, E. M., & Gernsbacher, M. A. (2000). Right and left hemisphere cooperation for drawing predictive and coherence inferences during normal story comprehension. *Brain and Language, 71,* 310-336.

Berthoz, S., Armony, J. L., Blair, R. J. R., & Dolan, R. J. (2002). An fMRI study of intentional and unintentional (embarrassing) violations of social norms. *Brain, 125,* 1696-1708.

Bird, C. M., Castelli, F., Malik, O., Frith, U., & Husain, M. (2004). The impact of extensive medial frontal lobe damage on `Theory of Mind' and cognition. *Brain, 127,* 914-928.

Bishop, D. V. M. (1998). Development of the children's communication checklist (CCC): A method for assessing qualitative aspects of communicative impairment in children. *Journal of Child Psychology and Psychiatry, 39*, 879-891.

Blake, M. L. (2006). Clinical relevance of discourse characteristics after right hemisphere brain damage. *American Journal of Speech and language Pathology, 15*, 255-267.

Blake, M. L. (2007). Perspectives on treatment for communication deficits associated with right hemisphere brain damage. *American Journal of Speech and Language Pathology, 16,* 331-342.

Blake, M. L., Duffy, J. R., Myers, P. S., & Tompkins, C. A. (2002). Prevalence and patterns of right hemisphere cognitive/communicative deficits: Retrospective data from an inpatient rehabilitation unit. *Aphasiology, 16*, 537-547.

Booth, J. R., Hall, W. S., Robison, G. C., & Kim, S. Y. (1997). Acquisition of the mental state verb know by 2- to 5- year-old children. *Journal of Psycholinguistic Research, 26,* 581- 603.

Bowler, D. M. (1992). Theory of mind in Asperger syndrome. *Journal of Child Psychology and Psychiatry, 33,* 877-893.

Brownell, H., Griffin, R., Winner, E., Friedman, O., & Happé, F. (2000). Cerebral lateralization and theory of mind. In: S. Baron-Cohen, H. Tager-Flusberg and D. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (2nd edition). (pp. 306-333) Oxford: Oxford University Press.

Brownell, H., Pincus, D., Blum, A., Rehak, A., & Winner, E. (1997). The effects of right hemisphere brain damage on patients' use of terms of personal reference. *Brain and Language, 57,* 60-79.

Brownell, H., & Stringfellow, A. (1999). Making requests: Illustrations of how RBD can affect discourse production. *Brain and Language, 68,* 442-465.

Brunet, E., Sarfati, Y., Hardy-Baylé, M. C., & Decety, J. (2000). A PET investigation of the attribution of intentions with a nonverbal task. *NeuroImage, 11,* 157-166.

Calarge, C., Anderson, N. C., & O'Leary, D. (2003). Visualizing how one brain understands another: A PET study of theory of mind. *American journal of Psychiatry, 160,* 1954-1964.

Campbell, A. L., Brooks, P., & Tomasello, M. (2000) Factors affecting young children's use of pronouns as referring expressions. *Journal of Speech, Language and Hearing Research, 43,* 1337-1349.

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development, 63,* 1-174.

Carston, R. (2002). *Thoughts and utterances*: *The pragmatics of explicit communication*. Oxford: Blackwell.

Cheang, H., & Pell M. (2006). A study of humor and communicative intention following right hemisphere stroke. *Clinical Linguistics Phonetics, 20,* 447-462.

Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris Publications.

Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Koshi, B. Webber, and I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10-63). Cambridge: Cambridge University Press.

Colle, L., Baron-Cohen, S., & Hill, J. (2007). Do children with Autism have a theory of mind? A non verbal test of Autism vs. specific language impairment. *Journal of Autism Developmental Disorder, 37,* 716-723.

Coulson, S., & Williams, R. F. (2005). Hemispheric asymmetries and joke comprehension. *Neuropsychologia, 43,* 128-141.

Crawford, J. R., & Garthwaite, P. H. (2002). Investigation of the single case in neuropsychology: Confidence limits on the abnormality of test scores and test score differences. *Neuropsychologia, 40,* 1196-1208.

Crawford, J. R., & Howell, D. C. (1998). Regression equations in clinical neuropsychology: An evaluation of statistical methods for comparing predicted and observed scores. *Journal of Clinical and Experimental Neuropsychology, 20,* 755-762.

Cutting, A., L., & Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: Individual differences and interrelations. *Child Development, 70,* 853-865.

Davis, G. A., O'neil- Pirozzi, T. M., & Coon, M. (1997). Referential cohesion and logical coherence of narration after right hemisphere stroke. *Brain and Language, 56,* 183-210.

Dennis, M., Lazenby, A. L., & Lockyer, L. (2001). Inferential language in high-function children with autism. *Journal of Autism and Developmental Disorders, 31,* 47- 54.

de Villiers, J. G. (2004). Getting complements on your mental state (verbs). In Van Kampen, J., & Baauw, S. (Eds.), *Proceedings of 2003 GALA Conference*, LOT, Utrecht, pp. 13–26.

de Villiers, J. G. (2007). The interface of language and theory of mind. *Lingua, 117,* 1858-1878.

de Villiers, J. G., & de Villiers, P. A. (2000). Linguistic determinism and the understanding of false beliefs. In P. Mitchell, & K. Riggs (Eds.), *Children's Reasoning and the Mind* (pp. 189-226). Psychology Press, Hove, UK.

de Villiers, J. G., & Pyres, J. (2002). Complements to cognition: A longitudinal study of the relationship between complex syntax and false belief understanding. *Cognitive Development, 17,* 1037-1060.

Eisele, J. A., Lust, B., & Aram, D. M. (1989). Presupposition and implication of truth: Linguistic deficits following early brain damage. *Brain and Language, 61,* 376-394.

Ferstl, E. C., Walther, K., Guthke, T., & von Carmon, D. Y. (2005). Assessment of story comprehension deficits after brain damage. *Journal of Clinical and Experimental Neuropsychology, 27,* 367-384.

Field, M. (1997). The role of factive predicates in the indexicalization of stance: A discourse perspective. *Journal of Pragmatics, 27,* 799-814.

Fine, J., Bartolucci, G., Szatmari, P., & Ginsberg, G. (1994). Cohesive Discourse in Pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24,* 315-329.

Fletcher, P. C., Happé, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S. J., & Frith, C. D. (1995). Other minds in the brain: A functional imaging study of "theory of mind" in story comprehension. *Cognition, 57,* 109-128.

Fodor, J. A. (1992). A theory of the child's theory of mind. *Cognition, 44,* 283-296.

Fournier, N. M., Calverley, K. L., Wagner, J. P., Poock, J. L., & Crossley, M. (2008). Impaired social cognition 30 years after hemispherectomy for intractable epilepsy: The importance of the right hemisphere in complex social functioning. *Epilepsy and Behavior, 12,* 460-471.

Friedmann, N. (1998). *Functional categories in agrammatic production: A cross linguistic study.* Doctoral dissertation, Tel Aviv University.

Friedmann, N. (2006). Speech production in Broca's agrammatic aphasia: Syntactic tree pruning. In Y. Grodzinsky & K. Amunts (Eds.), *Broca's Region.* Oxford University Press.

Friedmann, N., Belletti, A., & Rizzi, L. (2009). Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua, 119,* 67-88.

Friedmann, N., & Gvion, A. (2003). Sentence comprehension and working memory limitations in aphasia: A dissociation between semantic and phonological encoding. *Brain and Language, 86,* 23-39.

Friedmann, N., & Novogrodsky, R. (2004). The acquisition of relative clause comprehension in Hebrew: A study of SLI and normal development. *Journal of Child Language, 31,* 661-681.

Friedmann, N., & Shapiro, L. P. (2003). Agrammatic comprehension of simple active sentences with moved constituents: Hebrew OSV and OVS structures. *Journal of Speech Language and Hearing Research, 46,* 288-297.

Friedmann, N., & Szterman, R. (2006). Syntactic movement in orally trained children hearing impairment. *Journal of Deaf Studies and Deaf Education, 11,* 56-75.

Frith, U. (2001). Mind blindness and the brain in autism. *Neuron, 32,* 969-979.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society B: Biological Sciences, 358,* 459-473.

Frith, U., & Frith, C. D. (2006). The neural basis of mentalizing. *Neuron, 50,* 531-534.

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: an fMRI study of `theory of mind' in verbal and nonverbal tasks. *Neuropsychologia, 38,* 11-21.

Gallagher, H. L., Jack, A. I., Roepstorff, A., & Frith, C. D. (2002). Imaging the intentional stance in a competitive game. *NeuroImage, 16,* 814-821.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: III. Speech acts* (pp. 41–58). New York: Academic Press.

Grice, P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Grodzinsky, Y., Piñango, M. M., Zurif, E., & Drai, D. (1999). The critical role of group studies in neuropsychology: Comprehension regularities in Broca's aphasia. *Brain and Language*, *67*, 134-147.

Gundel, J. K., Ntelitheos, D., & Kowalsky, M. (2007). Children's use of referring expressions: Some implications for theory of mind. *ZAS Papers in Linguistics, 48,* 1-21.

Gvion, A., & Friedmann, N. (2008). FriGvi: Friedmann Gvion battery for assessment of phonological working memory. *Language and Brain, 7,* 161-180. (in Hebrew)

Haddad-Hanna, M., & Friedmann, N. (2009). The comprehension of sentences with and without syntactic movement in Arabic-speaking with hearing impairment. *Language and Brain*, *9*, 79-104. (in Arabic)

Hale, C. M., & Tager-Flusberg, H. (2003). The influence of language on theory of mind: A training study. *Developmental Science, 6,* 346-359.

Hale, C. M., & Tager-Flusberg, H. (2005). Social communication in children with autism: The relationship between theory of mind and discourse development. *Autism, 9,* 157-178.

Happé, F. (1993). Communicative competence and theory of mind in Autism: A test of relevance theory. *Cognition, 48,* 101-119.

Happé, F. (1994). An advanced theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders, 24*, 129-154.

Happé, F. (1995). *Autism*. Cambridge, MA: Harvard University Press.

Happé, F., Brownell, H., & Winner, E. (1999) Acquired 'theory of mind' impairments following stroke. *Cognition, 70,* 211-240.

Happé, F., & Loth, E. (2002). Theory of mind and tracking speakers' intentions. *Mind and Language, 17,* 24-36.

Happé, F., Winner, E., & Brownell, H. (1998). The getting of wisdom: Theory of mind in old age. *Developmental psychology, 34,* 358-362.

Harris, P. (1996). Desires, beliefs, and language. In P. Caarruthers & P. K. Smith (Eds.), *Theories of Theories of Mind* (pp. 200-220). Cambridge University press.

Hawkins, J. A. (1978). *Definiteness and indefiniteness: A study of reference and grammaticality prediction.* London: Croom Helm.

Hickmann, M., Kail, M., & Roland, F. (1995). Cohesive anaphoric relations in French children's narratives as a function of mutual knowledge. *First Language, 15,* 277-300.

Hough, M. S. (1990). Narrative comprehension in adults with right and left hemisphere brain-damage: Theme organization. *Brain and Language, 38,* 253-277.

Hollebrandse, B., Hobbs, K., de Villiers, J., & Roeper, T. (2007). Second order embedding and second order false belief. *Language Acquisition and Development: Proceedings of GALA 2007,* 270-280.

Hughes, C., Adlam, A., Happé, F., Jackson, J., Taylor, A., & Caspi, A. (2000). Good test-retest reliability for standard and advanced false belief tasks across a wide range of abilities. *Journal of Child Psychology and Psychiatry, 41,* 483-490.

Joanette, Y., & Ansaldo, A. I. (1999). Clinical note: Acquired Pragmatic Impairments and Aphasia. *Brain and Language, 68,* 529-534.

Johns, C. L., Tooley, K. M., Traxler, M. J. (2008). Discourse impairments following right hemisphere brain damage: A critical review. *Language and Linguistic Compass, 2,* 1038-1062.

Jollife, T., & Baron-Cohen, S. (1999). The strange stories test: a replication with high-functioning adults with autism or Asperger syndrome. *Journal of Autism and Developmental Disorders, 29,* 395-406.

Kail, M., & Hickmann, M. (1992). French children's ability to introduce referents in narratives as a function of mutual knowledge. *First Language, 12,* 73-94.

Karmiloff-Smith, A. (1979). *A functional approach to child language: A study of determiners and reference*. Cambridge: Cambridge University Press.

Kasher, A. (1991). On the pragmatic modules: A lecture. *Journal of Pragmatics, 16,* 381-397.

Kiparsky, P., & Kiparsky, C. (1971). Fact. In D. Steinberg & L. Jacobovits (Eds.) *Semantics* (pp. 345-369). Cambridge, MA: University Press.

Lee, A., Hobson, R. P., & Chiat, S. (1994). I, You, Me, and autism: An experimental study. *Journal of Autism and Developmental Disorders, 24,* 155-176.

Leslie, A. M. (1987). Pretense and representation: The origins of theory of mind. *Psychological Review, 94,* 412-426.

Leslie, A. M. (2000). How to acquire a representational theory of mind. In D. Sperber (Ed.), *Metarepresentations: A Multidisciplinary Perspective.* (pp. 197-223). Oxford: Oxford University Press.

Lohmann, H., & Tomasello, M. (2003). The role of language in development of false belief understanding: A training study. *Child Development, 74,* 1130-1144.

Lord, C., Cook, E. H., Leventhal, B. L., & Amaral, D. G. (2000). Autism Spectrum Disorders. *Neuron, 28,* 355-363.

Losh, M., & Capps, L. (2003). Narrative ability in high functioning children with Autism or Asperger's syndrome. *Journal of Autism and Developmental Disorders, 33,* 239-251.

Loukusa, S., & Molianen, I. (2009). Pragmatic inference abilities in individuals with Asperger syndrome or high functioning autism. A review. *Research in Autism Spectrum Disorder, 3,* 890-904.

Malle, B. F. (2005). Folk theory of mind: Conceptual foundations of human social cognition. In R. Hassin, J. Uleman, & J. Bargh (Eds.), *The new unconscious*. New York: Oxford University Press.

Mar, R. A. (2004). The neuropsychology of narrative: Story comprehension, story production and their interrelation. *Neuropsychologia, 42,* 1414-1434.

Marini, A., Carlomagno, S., Caltagirone, C., & Nocentini, U. (2005). The role played by the right hemisphere in the organization of complex textual structures. *Brain and Language, 93,* 46-54.

Martin, I., & McDonald, S. (2003). Weak coherence, no theory of mind, or executive dysfunction? Solving the puzzle of pragmatic language disorders. *Brain and Language, 85,* 451- 466.

Mason, R. A., Williams, D. L., Kana, R. K., Minshew, N., & Just, M. A. (2008). Theory of Mind disruption and recruitment of the right hemisphere during narrative comprehension in autism. *Neuropsychologia, 46,* 269-280.

Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2006). The effect of perceptual availability and prior discourse on young children's use of referring expressions. *Applied Psycholinguistics, 27,* 403-422.

Mayer, M. (1969). *Frog, Where are you?* New York: Dial Book for Young Readers.

McCabe, K., Houser, D., Ryan, L., Smith, V., & Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences of the United States of America, 98,* 11832-11835.

McDonald, S. (2000). Exploring the cognitive basis of right hemisphere pragmatic language disorder. *Brain and Language, 75,* 82-107.

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18 month old children. *Developmental Psychology, 31,* 838-850.

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

Montgomery, D. (1997). Wittgenstein's private language argument and children's understanding of the mind. *Developmental Review, 17,* 291-320.

Montgomery, D. (2002). Mental verbs and semantic development. *Journal of Cognition and Development, 3,* 357-384.

Moore, C., Brynt, D., & Furrow, D. (1989). Mental terms and development of certainty. *Child Development, 60,* 167-171.

Moore, C., & Furrow, D. (1991). The development of the language of belief: The expression of relative certainty. In D. Frye and C. Moore. *Children's theories of mind: Mental states and social understanding* (pp. 173-193). Hillsdale, NJ, England: Erlbaum.

Myers, P. S. (2001). Towards a definition of RHD syndrome. *Aphasiology, 15,* 913-918.

Myers, P. S. (2005). Profiles of communication deficits in patients with right cerebral hemisphere damage: Implications for diagnosis and treatment. *Aphasiology, 19,* 1147-1160. (Original work published in 1979).

Naigles, L. R. (2000). Manipulating the input: Studies in mental state acquisition. In B. Landau, J. Sabini, J. Jonides, & E. Newport (Eds.), *Perception, Cognition and Language: Essays in honor of Henry and Lila Gleitman*. Cambridge, Mass: MIT.

Nave, M., Szterman, R., & Friedmann, N. (2009). Comprehension and production of Wh-questions in Hebrew speaking children with hearing impairment: Additional evidence for syntactic difficulty. *Language and Brain, 9,* 1-29. (In Hebrew).

Nixon, S. M. (2005). Mental state verb production and sentential complements in four year old children. *First Language, 25,* 19-37.

Novogrodsky, R., & Friedmann, N. (2006). The production of relative clauses in syntactic SLI: A window to the nature of the impairment. *International Journal of Speech-Language Pathology, 8,* 364-375.

Olsen, D. R. (1998). On the origins of beliefs and other intentional states in children. In J. W. Astington, P. L. Harris, & D. Olsen (Eds.), *Developing Theories of Minds* (pp. 414- 426). Cambridge University Press.

Ozonoff, S., & Miller, J. (1996). An exploration of right- hemispheric contributions to the pragmatic impairments of autism. *Brain and Language, 52,* 411- 434.

Papafragou, A. (2002). Mindreading and verbal communication. *Mind and Language, 17,* 55-67.

Papafragou, A., Cassidy, K., & Gleitman, L. (2007). When we think about thinking: The acquisition of belief verbs. *Cognition, 105,* 125-165.

Pell, M. D. (1999). Fundamental frequency encoding of linguistic and emotional prosody by right hemisphere damaged speakers. *Brain and Language, 69,* 161-192.

Pell, M. D. (2006). Cerebral mechanisms for understanding emotional prosody in speech. *Brain and Language, 96,* 221-234.

Penn, C. (1999). Pragmatic assessment and therapy for persons with brain damage: What have clinicians gleaned in two decades? *Brain and Language, 68,* 535-552

Penn, C. (2000). Paying attention to conversation. *Brain and Language, 71,* 185-189.

Perkins, M. R. (2005). Pragmatic ability and disability as emergent phenomena. *Clinical Linguistics and Phonetics, 19,* 367-377.

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that…": Attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology, 39,* 437-471.

Pimkel, M. (2006). *Philosophical tales for children*. Tel Aviv: Ahuzat Bait Pub. (In Hebrew).

Polak, A., & Harris, P. L. (1999). Deception by young children following noncompliance. *Developmental Psychology, 35,* 561-568.

Prince, E. F. (1981). Towards a taxonomy of given-new information. In P. Cole (Ed.), *Radical Pragmatics* (pp. 223-256). New York: Academic Press.

Prince, E. F. (1992) 'The ZPG letter: Subjects, definiteness, and information status' In Mann, W. C., & Thompson, S. A. (Eds.), *Discourse description: Diverse linguistic analyses of a fund-raising text.* Amsterdam/Philadelphia: John Benjamins. pp. 295-326.

Rehak, A., Kaplan, J. A., & Gardner, H. (1992). Sensitivity to conversational deviance in right hemisphere damaged patients. *Brain and Language, 42,* 203-217.

Richard, M., Girouard, P. C., & Décarie, T. G. (1999). Personal pronouns and perspective taking in toddlers. *Journal of Child Language, 26,* 681-697.

Rajendran, G., & Mitchell, P. (2007). Cognitive theories of Autism. *Developmental Review, 27,* 224-260.

Roth, D., & Leslie, A. M. (1998). Solving belief problems: towards a task analysis. *Cognition, 66,* 1-31.

Ruigendijk, E., Friedmann, N., Novogrodsky, R., & Balaban, N. (in press). Symmetry in comprehension and production of pronouns: A comparison of German and Hebrew. *Lingua*.

Ruigendijk, E., Vasić, N., & Avrutin, S. (2006). Reference assignment: Using language breakdown to choose between theoretical approaches. *Brain and Language, 96,* 302-317.

Russell, B. (1905). On denoting. *Mind, 14,* 479-493.

Russell, T. A., Schmidt, U., Doherty, L., Young, V., & Tchanturia, K. (2009). Aspects of social cognition in anorexia nervosa: Affective and cognitive theory of mind. *Psychiatry Research, 168,* 181-185.

Rutherford, M. D., Baron-Cohen, S., & Wheelwright, S. (2002). Reading the mind in the voice: A study with normal adults and adults with Asperger syndrome and high functioning Autism. *Journal of Autism and Developmental Disorders, 32,* 189-194.

Sabbagh, M. A. (1999). Communicative intentions and language: Evidence from right hemisphere damage and autism. *Brain and Language, 70,* 29-69.

Sabbagh, M. A., & Flynn, J. (2006). Mid frontal EEG alpha asymmetries predict individual differences in one aspect of theory of mind: Mental state decoding. *Social Neuroscience, 1,* 299-308.

Sampson, D., Apperly, I. A., Chiavarino, C., & Humphreys, G. W. (2004). Frontal and temporo- parietal lobe contributions to theory of mind: Neuropsychological evidence from a false belief task with reduced language and executive demands. *Journal of Cognitive Neuroscience, 16,* 1773-1784.

Saxe, R., & Powell, L. J. (2006). It's the thought that counts: Specific brain regions for one component of theory of mind. *Psychological Science, 17,* 692-699.

Saxe, R., & Wexler, A. (2005). Making sense of another mind: The role of the right tempro-parietal junction. *Neuropsychologia, 43,* 1391-1399.

Schaeffer, J., & Hacohen, A. (2003). The dissociation between grammar and pragmatics: evidence from English SLI. In *Proceedings of the Israel Association for Theoretical Linguistics,* IATL 19, Ben-Gurion University of the Negev, Beer-Sheva.

Schaeffer, J., & Matthewson, L. (2005). Grammar and pragmatics in the acquisition of article systems. *Natural Language & Linguistic Theory, 23*, 53-101.

Scholl, B. J., & Leslie, A. M. (1999). Modularity, development and theory of mind. *Mind and Language, 14*, 131-153.

Schulz, P. (2003). *Factivity: Its nature and acquisition*. Max Niemeyer Verlag: Tubingen.

Schwanenflugel, P. J., Fabricius, W. V., & Noyes, C. R. (1996). Developing organization of mental verbs: Evidence for the development of a constructivist theory of mind in middle childhood. *Cognitive Development, 11,* 265-294.

Schwanenflugel, P. J., Henderson, R. L., & Fabricius, W. V. (1998). Developing organization of mental verbs and theory of mind in middle childhood: Evidence from extensions. *Developmental Psychology, 34,* 512-524.

Shamay-Tsoory, S. G., Tomer, R., & Aharon- Perez, J. (2005). The neuro-anatomical basis of understanding sarcasm and its relationship to social cognition. *Neuropsychology, 19,* 288-300.

Siegal, M. (1999). Language and thought: The fundamental significance of conversational awareness for cognitive development. *Developmental Science, 2,* 1-34.

Siegal, M., Carrington, J., & Radel, M. (1996). Theory of mind and pragmatic understanding following right hemisphere damage. *Brain and Language, 53,* 40-50.

Siegal, M., & Valrey, R. (2002). Neural systems involved in 'theory of mind'. *Neuroscience, 3,* 463-470.

Silliman, E. R., Diehl, S. F., Bahr, R. H., Hnath-Chisolm, T., Zenko, C. B., & Friedman, S. A. (2003). A new look at performance on Theory of Mind tasks by adolescents with autism spectrum disorder. *Language, Speech, and Hearing Services in Schools, 34,* 236-252.

Sinclair, M. (1995). Fitting pragmatics into the mind: Some issues in mentalistic pragmatics. *Journal of Pragmatics, 23,* 509-539.

Slessor, G., Phillips, L. H., & Bull, R. (2007). Exploring the specificity of age-related differences in Theory of Mind tasks. *Psychology and Aging, 22,* 639-643.

Slomkowski, C., & Dunn, J. (1996). Young children's understanding of other people's beliefs and feelings and their connected communication with friends. *Developmental Psychology, 32,* 442-447.

Smith, A. (2008). The empathy imbalance hypothesis of autism: A theoretical approach to cognitive and emotional empathy in Autistic development. *The Psychological Record, 59,* 273-294.

Spanoudis, G., Natsopoulos, D., & Panayiotou, G. (2007). Mental verbs and pragmatic language difficulties. *International Journal of Language and Communication Disorders, 42,* 487-504.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition*. Oxford: Blackwell. (Original work published in 1986)

Sperber, D., & Wilson, D. (2002). Pragmatics, modularity and mind-reading. *Mind and Language, 17,* 3-23.

Strauss, S., Ziv, M., & Stein, A. (2002). Teaching as a natural cognition and its relations to preschoolers developing theory of mind. *Cognitive Development, 104,* 1-15.

Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Developmental Psychology, 30,* 395-402.

Surian, L., Baron-Cohen, S., & Van der Lely, H. (1996). Are children with Autism deaf to Gricean Maxims? *Cognitive Neuropsychiatry, 1,* 55-71.

Surian, L., & Siegal, M. (2001). Sources of performance on theory of mind tasks in right hemisphere damaged patients. *Brain and Language, 78,* 224-232.

Tager-Flusberg, H. (1989). A psycholinguistic perspective on language development in the autistic child. In G. Dawson (Ed.), *Autism: Nature, diagnosis and treatment* (pp. 92– 115). New York: Guilford Press.

Tager-Flusberg, H. (2000). Language and understanding minds: Connections in autism. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from autism and developmental cognitive neuroscience.* Second Edition. Oxford: Oxford University Press.

Tager-Flusberg, H., & Sullivan, K. (1994). A second look at second order false belief attribution in autism. *Journal of Autism and Developmental Disorders, 24*, 577- 586.

Talwar, V., Gordon, H. M., & Lee, K. (2007). Lying in elementary school years: Verbal deception and its relation to second order belief understanding. *Developmental psychology, 43*, 804-810.

Thompson, S. A. (2002). "Object complements" and conversation: Towards a realistic account. *Studies in Language, 26*, 125-163.

Thordardottir, E. T., Chapman, R. S., & Wagner, L. (2002). Complex sentence production by adolescents with Down syndrome. *Applied Psycholinguistics, 23*, 163-183.

Tomasello, M., & Habrel, K. (2003). Understanding Attention: 12 and 18 month- olds know what is new for other persons. *Developmental Psychology, 39*, 906-912.

Tompkins, C. A., Fassbinder, W., Blake, M. L., Baumgaertner, A., & Jayaram, N. (2004). Inference generation during text comprehension by adults with right hemisphere brain damage: Activation failure versus multiple activation. *Journal of Speech, Language, and Hearing Research, 47*, 1380-1395.

Tompkins, C. A., Lehman-Blake, M. T., Baumgaertner, A., & Fassbinder, W. (2001). Mechanisms of discourse impairment after right hemisphere brain damage: Suppression in inferential ambiguity resolution. *Journal of Speech, Language, and Hearing Research, 44*, 400-415.

Tompkins, C. A., Scharp, V. L., Fassbinder, W., & Meigh, K. M. (2008). A different story on "Theory of mind" deficit in adults with right hemisphere brain damage. *Aphasiology, 22*, 42-61.

Uryase, D., Duffy, R. J., & Liles, B. Z. (1991). Analysis and description of narrative discourse in right-hemisphere-damaged adults: A comparison to neurologically normal and left hemisphere- damaged aphasic adults. In T. E. Prescott (Ed.), *Clinical aphasiology.* Austin, TX: Pro-Ed. Vol. 19, pp. 125–138.

Vasić, N. (2006). *Pronoun Comprehension in Agrammatic Aphasia*: *The Structure and Use of Linguistic Knowledge*. LOT: Netherlands Graduate School of Linguistics.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory of mind development: The truth about false belief. *Child Development, 72,* 655-684.

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

Weylman, S. T., Brownell, H., Roman, M., & Gardner, H. (1989). Appreciation of indirect requests by left and right brain damaged patients: The effects of verbal context and conventionality of wording. *Brain and Language, 36,* 580-591.

White, S., Hill, E., Happé, F., & Frith, U. (2009). Revisiting the strange stories: Revealing mentalizing impairments in autism. *Child Development, 80,* 1097-1117.

Wilson, D. (1999). Relevance and relevance theory. In R. Wilson & F. Keil (Eds.) *MIT Encyclopaedia of the Cognitive Sciences*., Cambridge MA: MIT Press, pp. 719-722.

Wilson, D. (2005). New directions for research on pragmatics and modularity. *Lingua, 115,* 1129-1146.

Wilson, D., & Matsui, T. (1998). Recent approaches to bridging: Truth, coherence, relevance. *UCL Working Paper in Linguistics, 10,* 1-28.

Wilson, D., & Sperber, D. (1993). Linguistic form and relevance. *Lingua, 90,* 1-25.

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

Winner, E., Brownell, H., Happé, F., & Blum, A. (1998) Distinguishing lies from jokes: Theory of mind deficits and discourse interpretation in right hemisphere brain- damaged patients. *Brain and Language, 62,* 89-106.

Ziatas, K., Durkin, K., & Pratt, C. (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,* 755-763.

Ziv, M., & Frye, D. (2004). Children's understanding of teaching: The role of knowledge and belief. *Cognitive Development, 19,* 457-477.

Ziv, M., Solomon, A., & Frye, D. (2008) Young children's recognition of the intentionality of teaching. *Child Development, 79,* 1237-1256

Zuckerman, S., Vasic, N., Ruigendijk, E., & Avrutin, S. (2002). Experimental evidence for the subject rule. *Proceedings IATL*. Utrecht University.

Appendix: Justifications

			Appropriate		Inappropriate	
		Expected Response	**Specific and Explicit**	**General and Implicit**	**Mental inappropriate**	**Physical or Irrelevant**
2nd order False Belief	A mother hides a puppy she intended to be a surprise for her son. The son finds out and tells a friend he will be getting a dog. Later the friend called and asked if he can come over to see the dog, before the present was revealed by the mother. We asked – what will the son answer his friend?	To refer to what the son thinks about his mother state of knowledge, her intention that it will be a surprise	No, he can't, because his mother is home. If they go to see the dog the mother will know that he knows. (Gila)	No, not yet, because the party didn't start yet and he didn't get his dog yet. (Yaron)	Yes, because he is so happy about getting the dog (Tzvi)	Yes, so they can get ready for the party (Dafna)
Knowledge gaps	A student hide a note in her pencil case, she planned to copy from it during a test. Before the test began, her teacher asked her to come over with her pencil case because she needed a pen. The student handed her the note. We asked: why did she do that?	To refer to the student's misunderstanding of her teacher's intentions	Te student was sure the teacher knew she wanted to copy (in Hebrew- it is almost as saying to cheat). (Sigalit)	She was afraid she will get caught (Moshe)	She had no choice the teacher found out she copied (Sara)	Because it was in the pencil case and the teacher saw it. (Abraham)

			Appropriate		Inappropriate	
		Expected Response	**Specific and Explicit**	**General and Implicit**	**Mental inappropriate**	**Physical or Irrelevant**
Instruction TPK	A sport teacher is teaching all his classes a new game. One student was missing but the teacher forgo that he was not with the class when he taught the game. We asked: Will he try to teach that student the next lesson?	To refer to the teacher false belief about his pupil's knowledge	He won't teach him because he thinks he already knows. (Ahuva)	The teacher will presume he didn't get it and leave him alone. (Ayal)	The boy missed class, he doesn't know the game so the teacher will teach him. (Tzipora)	Of course he'll teach how else will he be able to participate? (Jacob)
Instruction TPNK	At the end of the school year a teacher got a report about the children she'll be teaching next year. The report said that one of the children, Yonatan, can't read yet. During the summer Yonatan learned to read. We asked: Will the teacher try to teach Yonatan how to read?	To refer to the teacher false belief about her pupil's knowledge.	Yes, because she doesn't know that she learned. (Dror)	She'll try to test him, I know that every teacher checks. (Gila)	No, he(=the pupil) knows. (Dafna)	Yes, because the teacher is tested by her student's ability to read. (Dov)
Faux pas	A girl started drum lessons, her cousin comes over to her family's house for dinner and complains about the noise the awful students in the drum class are making. We asked – did the cousin know that his cousin started drum class, and how do you if he knew or not?	To refer to the cousins unawareness to the girl's feelings.	He didn't know. He wouldn't have said the students were no good if he knew his cousin was one of them. (Sigalit)	He didn't know. He would have been more intelligent and not say it to his cousin. (Moshe)	He knew. .. how could he not know his cousin was there, they live close by. (Yigal)	He knew, Because he told everyone, he was there and he told everyone about it. (Sachar)

			Appropriate		Inappropriate	
		Expected Response	**Specific and Explicit**	**General and Implicit**	**Mental inappropriate**	**Physical or Irrelevant**
Surprise	A man is planning to surprise his girlfriend with tickets to a trip abroad. He tells her they will be going to Jerusalem. His girlfriend finds the receipt and understands a surprise is planned. Nevertheless when they approach the airport she cries out – "Wow, I can't believe it, I'm so surprised, we're going abroad, that's wonderful!". We asked : (1) Is it true what the girlfriend said, 'was she really surprised?' (2) why do you think so?	To refer to the girlfriends state of knowledge and her motive for acting surprised.	*She didn't want him to know she knew about the tickets* (Avraham)	*So he won't feel bad.* (Sigalit)	*Because she was happy about it*.(=the trip) (David)	*I suppose (because) he didn't go to Jerusalem, he drove to the air-port.* (Sasson)
Empathy	Two students take a course test, Miri and Gilad. The results are posted on a bulletin board. Miri passed and Gilad failed. Only Miri saw the results, Gilad was at work and didn't see them. Gili called Gilad and said she was sorry, Gilad didn't know what she was talking about and Miri replied, Oh, never mind, forget it. We Asked : why did Miri say she was sorry?	To refer to her friends disappointment.	Because her friend didn't pass the test. (Sigalit)	Because she thought he didn't pass the test (Danny)	She saw the results, he didn't so she said she was sorry. (Sachar)	Because she saw the results and he didn't. (Simon)

			Appropriate		Inappropriate	
		Expected Response	**Specific and Explicit**	**General and Implicit**	**Mental inappropriate**	**Physical or Irrelevant**
2nd order Empathy	(1)When Miri made the call, did she think Gilad knew the results, or that he didn't know? (2) Why do you think so?	To refer to her presuppositions when approaching the friend.	(1)She thought he knew. (2) Otherwise she wouldn't have said she was sorry, she didn't want to hurt him. She would have waited for him to find out on his own. (Sharon)	(1)She thought he knew. (2)I know that from his answer. (Gila)	(1)Surely she knew that he didn't know. (2) She knew he was not in school; she gave him something to think about. (Yigal)	(1)No (2)Because he wasn't there that day (Sara)
Cartoon	A shark pulling out a stick	To refer to what the shark intended the man will think	A shark is pulling out a hand so the man will jump, so he'll think that someone is asking for help , and he will eat him up. (Yaron)		There's a man here that is frightened of the shark. (Dror)	A shark can't hold a stick like that. (Sasson)

אוניברסיטת תל-אביב

בית הספר לחינוך

ע"ש חיים וג'ואן קונסטנטינר

ההשלכות השפתיות של פגיעה נרכשת ב- Theory of Mind

חיבור לשם קבלת התואר "דוקטור לפילוסופיה"

מאת:

נגה בלבן

הוגש לסנאט של אוניברסיטת תל אביב

אפריל 2010

עבודה זו נעשתה בהדרכת

פרופ' נעמה פרידמן ופרופ' מירה אריאל

* המחקר מומן בעזרת מענק מחקר שניתן ע"י האגודה הישראלית לקרנות מחקר וחינוך (מענק דוקטורנטים מס. 28).

תודות

אני מודה למטופלים על הנכונות לתרום מזמנם ולשתף אותי בחלק מהחוויות שלהם. למדתי מכם שיעור חשוב על בריאות ומוגבלות ועל האופן בו אנשים מתמודדים עם ההבדל ביניהם.

תודה לד"ר רוזנטול, ד"ר אגרנובה, ד"ר גביעון וחנה קרפין שאיפשרו לי להכיר את המטופלים במחלקותיהן.

תודה לנבדקי הבקרה ששיתפו פעולה בהכנעה.

תודה לחברי מעבדת שפה ומח על החברות והתמיכה, אני שמחה להיות חלק מהקבוצה.

תודה למורותי,

למרגלית זיו, על שפתחת בפני את תחום המחקר הכי מעניין באקדמיה ועל כך שהארת את עיני לפשטות ולמורכבות שניתן לגלות בו, שלובים יד ביד.

למירה אריאל, על ההדרכה המקצועית והמרתקת במרחבי הפרגמטיקה והעזרה בגילוי האופן שבו תיאוריות מתעוררות לחיים.

לנעמה, על כך שהכל היה כמו שהיה, תודה גדולה.

היה מרגש להכיר את האהבה וההתלהבות שאת משקיעה בכל מה שאת עושה, היצרתיות, הפתיחות, הידענות, הדיוק, הסקרנות וההנאה השמחה, אני מקוה שאלו יהיו לי מקור להשראה.

מבחינתי מה שסיכמנו בעמודים אלו - זו רק ההתחלה.

תודה לחברים, לאלו שהסתכלו מרחוק ולא שאלו, לאלו שהתעניינו בסקרנות והקשבה ותודה מיוחדת למי שממש התקרב ונגע.

למשפחתי היקרה, תודה גדולה

לאמא ואבא שתומכים אוהבים ועוזרים לי בכל, אין מילים להודות, אני מקוה שבמעשים אצליח.

תודה אוהבת לילדי, מיה ונמרוד, על הסבלנות והעוד סבלנות.

והכי, לבן עמי, אהובי, בלעדיך לא היה כלום.

אלף- בית

ויסלבה שימבורסקה

לְעוֹלָם כְּבָר א׳ אֵדַע

מָה חָשַׁב עָלַי א׳.

אִם ב׳ לֹא סָלְחָה לִי עַד הַסּוֹף.

מַדּוּעַ ג׳ הֶעֱמִיד פָּנִים שֶׁהַכֹּל בְּסֵדֶר.

מָה הָיָה חֶלְקָהּ שֶׁל ד׳ בִּשְׁתִיקָתוֹ שֶׁל ה׳.

לְמַה קְוָה ו׳, אִם קְוָה.

לָמָּה ז׳ שָׁכְחָה עַל אַף שֶׁהֵיטִיבָה לָדַעַת.

מָה הָיָה לְח׳ לְהַסְתִּיר.

מָה רָצְתָה ט׳ לְהוֹסִיף.

אִם לְכָךְ שֶׁהָיִיתִי בַּסְּבִיבָה

הָיְתָה מַשְׁמָעוּת כִּלְשֶׁהִי

בְּעֵינֵי י׳, כ׳ וּשְׁאָר הָאָלֶף-בֵּית.

תרגם מפולנית רפי וייכרט

תקציר

תיאורית המיינד היא יכולת קוגניטיבית המאפשרת לאדם להבחין בין מצבים מנטליים שונים, להבחין בין אמונות וכוונות של אנשים שונים ולתפוס את הקשר בין האמונה, הכוונה ומצב הידע של אדם והתנהגותו בסיטואציות חברתיות שונות. היא מאפשרת לתפוס כי יתכן שלאדם אחר אמונה שונה וידע שונה ועל כן היא מהותית ובסיסית לכל אינטראקציה חברתית ותקשורתית. ליקוי ביכולת זו יוצר קושי בתחומים רגשיים והתנהגותיים. אחת מהפעילויות המרכזיות בחיי היום יום התלויה ביכולת זו היא היכולת לנהל שיח בעל משמעות, כלומר, להבין מסרים עקיפים או סמויים, הנחות שאינן מובעות במפורש ושפה לא ליטרלית.

קשיים שונים בתחום זה, של התנהלות בשיח, אותרו אצל מטופלים שסבלו מפגיעה מוחית בהמיספרה הימנית. נמצא כי מטופלים אלו מתקשים לעיתים להפיק טקסטים קוהרנטים, להבין ביטויים אירוניים ומטפוריים, להבין את העיקר (point) בסיפורים או את שורת המפתח בבדיחות ולהבין כוונות של בני שיחם. בשל מאפיינים אלו, מחקרים שונים במהלך עשרים השנים האחרונות קושרים בין פגיעה ב׳תיאורית המיינד׳ ופגיעה בהמיספרה הימנית. מחקרי הדמיה מוחית שבוצעו בקרב אוכלוסיה זו כמו גם בקרב אוכלוסיות בריאות מאששות את קיומו של הקשר בין תיאורית המיינד ותפקוד תקין של ההמיספרה הימנית.

על כן, בתחום השיח, פגיעה בהמיספרה הימנית עשויה להתבטא באי יכולת להבין שהידע של בן השיח הוא שונה משל הדוברת. בהתאם, ניתן לצפות שלפגיעה זו תהיה גם השפעה לשונית, בתחומים שבהם היכולת הלשונית תלויה בכך שהדוברת והנמען יקחו בחשבון את עמדת בני שיחם ביחס לאינפורמציה העולה בשיחה. שאלה זו, בדבר יכולותיהם הלשוניות של מי שנפגעו ביכולת הקוגניטיבית של תיאורית המיינד, לא נחקרה עד כה לעומק באופן זה.

על כן, במחקר זה בדקנו באיזה אופן משפיעה פגיעה בתיאורית המיינד על יכולות לשוניות ספציפיות של מטופלים שסבלו פגיעה בהמיספרה הימנית.

המחקר התנהל בשני שלבים. מטרת השלב הראשון היתה לזהות קבוצת מטופלים הסובלים מליקוי בתיאורית המיינד מתוך קבוצה של חולים שסבלו פגיעה בהמיספרה הימנית. כדי לזהות קבוצה זו פיתחנו סוללת מבדקים (סוללת אתומיה) הכוללת משימות המקיפות מגוון היבטים של ׳תיאורית המיינד׳. בחנו את הביצועים של 25 פגועי ההמיספרה הימנית בהשוואה לקבוצת בריאים הדומים להם. הממצא המרכזי של שלב זה היה שקבוצת פגועי המיספרה ימין היא קבוצה הטרוגנית מבחינת יכולותיה הקוגניטיביות בתחום תיאורית המיינד. מצאנו כי 17 מהמטופלים עם הפגיעה המוחית סובלים גם מפגיעה בתיאורית המיינד.

בשלב השני של המחקר בחנו יכולות לשוניות ספציפיות, הקשורות ליכולת לקחת בחשבון את עמדתו של בן השיח במהלך התפתחות השיחה, בשתי קבוצות המטופלים אלו שנמצאה אצלם פגיעה בתיאורית המיינד (להלן ליתום) ואלו שלמרות הפגיעה המוחית, לא סבלו מפגיעה כזו בהשוואה לקבוצת בריאים הדומים להם.

היכולות הלשוניות שבחנו הן:

<u>שימוש ביידוע להצגת ישויות חדשות לתוך השיחה</u>: משימה זו דורשת מהדוברת ניטור של הידע של הנמען במהלך השיחה ביחס לישויות השונות המוצגות בה והבחנה בין היישויות החדשות עבור הנמען לבין אלו המוכרות לו. הוספת הא הידיעה לפריט המוצג בשיחה מבטאת את ההנחה של הדוברת כי הפריט מוכר לנמען. למשל: משפטים כמו 'הגישה את העבודה', 'החזירה את התיק', 'נכנסה אל החדר', תקין שיאמרו במסגרת שיחה, רק אם ברור לדוברת כי הנמען יודע באיזה 'עבודה', 'תיק' או 'חדר' מדובר. יכולת זו נבחנה באמצעות שלושה ניסויים שונים שבדקו את יכולת ההפקה וההבנה של הנבדקים בתחום זה.

יכולת נוספת שנבחנה היא היכולת לעשות <u>שימוש תקין בביטויים מרפררים</u>, הן כדוברים והן כנמענים. בדקנו האם יש הבדל בין האופן שבו מטופלים שאבחנו כ'לתומים' לעומת כאלו שלא סבלו פגיעה כזו עושים שימוש במגוון אמצעי הרפרור העומדים לרשותם כדוברי השפה. האם השימוש באופני הרפרור המגוונים נעשה באופן שיהיה ברור לנמעניהם למי הכוונה בכל שלב בשיח? האם המטופלים, כנמענים, מבינים בשלבי השיחה השונים אל מי מתכוונת הדוברת? למשל, משפט הפותח תיאור של אינטראקציה בין שניים בכינוי גוף ('הוא פתח את דבריו בהבטחה...'), מבלי שהוצגו קודם הדמויות, אינו מאפשר לנמען לזהות את מי מהשניים תיארה הדוברת. יכולת זו נבדקה באמצעות 6 ניסויים שונים. שניים בחנו את יכולת ההפקה של הנבדקים וארבעה אחרים את ההבנה המטאלינגוויסטית שלהם לגבי סוגי אמצעי הרפרור השונים.

תחום לשוני נוסף שבדקנו הוא האופן בו המטופלים <u>משתמשים בפעלים מנטליים וכיצד הם מבינים אותם</u>. פעלים מנטליים, כמו *חשב*, *ידע* ו- *זכר* מכילים במשמעותם אספקטים שונים שמתייחסים למצבים מנטליים שונים ביחס למציאות בשיח. הכרת האופן בו הלתומים והתומרים עושים שימוש במאפיינים הלקסיקליים השונים של פעלים אלו חשובה להבנת הקשר בין תיאורית המיינד ויכולות לשוניות. יכולת זו נבחנה באמצעות 4 ניסויים שונים.

בנוסף, בחנו יכולות לשוניות נוספות של מטופלים אלו, כמו הבנה והפקה של משפטי זיקה וכן את ההבנה של כללי הכבילה ע"פ חומסקי באמצעות 4 ניסויים שונים. יכולות אלו אינן קשורות לתיאורית המיינד אולם הן מאפיינות אוכלוסיות מחקר אחרות הסובלות מליקוי בתיפקודם של מרכזי השפה במח. מטרת בדיקה זו היא לבחון האם פגיעה לשונית ממנה סובלים הלתומים היא פגיעה לשונית כוללת. כמו כן, בחנו את יכולתם לעשות שימוש במשפטים משועבדים, שלגביהם נטען בספרות, כי רכישתם היא תנאי הכרחי להבשלת 'תיאורית המיינד' אצל ילדים צעירים.

הממצאים במבדקים הלשוניים הראו כי קבוצת המטופלים הלתומים הצליחה באופן עקבי פחות טוב מקבוצת המטופלים שתיאורית המיינד שלהם נמצאה תקינה רק במטלות הלשוניות הקשורות לתיאורית המיינד. קבוצת הלתומים התקשתה בכל התחומים הלשוניים שנבדקו ואילו הקבוצה השניה הצליחה במבדקים במידה שווה לקבוצת הבקרה. לעומת זאת, במשימות הלשוניות שאינון תלויות בתיאורית המינד, כמו גם בהפקה של משפטים משועבדים הצליחו שתי קבוצות המטופלים בדומה לנבדקי הביקורת.

מן הממצאים אנו מסיקות כי קבוצת המטופלים שסבלה פגיעה בהמיספרה הימנית אינה קבוצה אחידה. לפיכך ישנו צורך לזהות אצל מי מהמטופלים ישנו ליקוי בתיאורית המינד ואצל מי מהם לא היתה השפעה על יכולת זו. אנו מקוות שהכלים שפותחו במהלך המחקר יוכלו לשמש ככלי אבחוני בקליניקה ובכך יקדמו את איכות הטיפול לה זוכה קבוצה יחודית זו של מטופלים.

מסקנה נוספת היא כי תיתכן פגיעה לשונית מוגדרת אצל מטופלים שסובלים מפגיעה בהמיספרה הימנית, על אף שלא סבלו מפגיעה ממוקדת במרכזי השפה במח.

לפיכך, אנו למדות כי פגיעה ביכולת הקוגניטיבית של תיאורית המינד גוררת איתה קשיים לשוניים ספציפיים. ליקויים אלו אינם מלווים בחוסר יכולת לשונית רחבה יותר, הם מופיעים רק ביכולת הלשוניות הקשורות לתחום תיאורית המינד.