



Contents lists available at ScienceDirect

## Journal of Neurolinguistics

journal homepage: [www.elsevier.com/locate/jneuroling](http://www.elsevier.com/locate/jneuroling)

## Research paper

# Is Theory of Mind the basis for exhaustivity in wh-questions? Evidence from TOM impairment after right hemisphere damage

Noga Balaban<sup>a</sup>, Petra Schulz<sup>b</sup>, Naama Friedmann<sup>a,\*</sup><sup>a</sup> Language and Brain Lab, School of Education and Sagol School of Neuroscience Tel Aviv University, Tel Aviv, 69978, Israel<sup>b</sup> Goethe-Universität Frankfurt, 60629 Frankfurt am Main, Germany

## A B S T R A C T

How do we know when to provide an exhaustive answer to a wh-question, which mentions all the items that satisfy the property being asked about? We explored the nature of this exhaustivity requirement by investigating whether it is grammatical or based on assessing the information needs of the person asking the question. In Experiment 1 we tested 14 individuals after right hemisphere damage who had Theory of Mind (TOM) impairment (aTOMia), which compromised their ability to assess the information needs of their interlocutor, but whose grammatical abilities were normal. The rationale was that if they provided an exhaustive answer to certain question types, the exhaustivity requirement could not depend on considering the asker's intentions. We assessed their responses to single wh-questions (e.g., *Who is painting?*) and multiple wh-questions (e.g., *Who is painting what?*) in a question-about-a-picture task compared to 5 right-hemisphere damaged patients with good TOM, and to healthy controls. The individuals with aTOMia often failed to provide exhaustive responses to the single wh-questions, but consistently produced exhaustive responses to the multiple wh-questions. This finding suggests that the source of exhaustivity in single wh-questions is based on TOM, whereas exhaustivity in multiple wh-questions is grammatically-based. Following up on a further result from Experiment 1, Experiment 2 investigated the effect of the exhaustivity requirement on attention: 6 additional participants with left visuo-spatial neglect after right hemisphere damage were tested with a similar task to assess whether their answers omitted figures on the left. Both experiments – Experiment 1 with eight neglect patients and Experiment 2 with six, revealed that the exhaustivity requirement also triggered attention shift to the left in individuals with left neglect: in their responses to single wh-questions they omitted the left figures, but in paired wh-questions they scanned the entire visual field and responded with exhaustive answers. The results provide clear evidence from neurolinguistics that exhaustivity in single wh-questions requires considerations of the asker's intentions, whereas multiple wh-questions mark the exhaustivity requirement grammatically.

## 1. Introduction

Imagine a former classmate of yours is telling you that she went hiking with a few of your old classmates during spring break. She says that they came across a field of special wild herbs and they decided that each of them would pick a different kind of herb. If you ask her *Who went on that trip?* her answer will probably include an exhaustive list of all your former classmates who fulfil the property of participating in that field trip. In other words, she will provide an exhaustive answer that fits a 'mention-all' reading of the wh-question. If you ask a paired wh-question like *Who picked what?* she will provide an exhaustive list of the pairs of each person and the kind of herb s/he picked. Similarly, if you ask a triple wh-question like *Who picked what where?* she will respond with an exhaustive list of triples, naming each person and what they picked and where they picked it (e.g., Maya picked mint near the pond and Esther picked sage near the rock).

In other contexts, the asker does not expect an exhaustive answer to single wh-questions. For example, if that friend asks you *Who did you see during spring break?* you would most likely provide a 'mention-some' reading, listing relevant mutual acquaintances whom you assume your friend is referring to, and not everybody you saw during that period of time. Importantly, the mention-some reading

\* Corresponding author.

E-mail address: [naamafr@tauex.tau.ac.il](mailto:naamafr@tauex.tau.ac.il) (N. Friedmann).<https://doi.org/10.1016/j.jneuroling.2019.100853>

Received 28 October 2018; Received in revised form 3 June 2019; Accepted 19 June 2019

Available online 17 July 2019

0911-6044/ © 2019 Elsevier Ltd. All rights reserved.

is not available in multiple (paired and triple) wh-questions, for which exhaustive readings, listing all pairs or triples, are obligatory (see e.g., Krifka, 2001).<sup>1</sup>

The main goal of this paper is to characterize the nature of the exhaustivity requirement in single and in multiple wh-questions, and specifically, to find out whether Theory of Mind (TOM) is involved in exhaustivity in wh-questions. To explore this issue, we tested how individuals with TOM impairment and intact grammatical abilities interpret single and multiple wh-questions.

### 1.1. Exhaustivity and its nature

Exhaustivity is a central aspect of the semantics of single wh-questions like *Who is reading a book?* and of paired wh-questions like *Who is eating what?* Under an exhaustive reading, the answer exhausts the relevant set of elements satisfying the predicate, an exhaustive list of individuals for single wh-questions, and an exhaustive list of pairs in paired wh-questions. In principle, the nature of the exhaustivity requirement for single and multiple wh-questions could be uniform or could differ according to question type. We consider three possibilities: one is that the source of the exhaustivity requirement in single and multiple wh-questions is identical, and both are guided by the intentions of the speaker. The second is that both are guided by purely grammatical principles. The third allows for different sources for the exhaustivity requirement in single and multiple wh-questions and suggests that single wh-questions are guided by the intentions of the speaker, whereas exhaustivity in multiple wh-questions is grammatically determined.

In the following we sketch these possibilities in turn. We consider first the possibility that exhaustivity in single and multiple wh-questions are both guided by the perspective of the speaker. Under this assumption, the source of exhaustivity is non-grammatical; both question types share the requirement to take into account the interlocutor's perspective and to decide on the optimal answer given the specific context of uttering the question and the **interlocutors' intention**. For single wh-questions, this account has been spelled out by theories such as van Rooij's (2003, 2004), who suggested that single wh-questions have a semantically underspecified meaning: the answer to the wh-question provides the smallest set that gives optimal relevance in the situation. Other proposals (e.g., Nelken & Shan, 2004) claim that single wh-questions possess two different semantic meanings, mention-some and mention-all. These proposals have to incorporate somehow that the person answering has to resolve the decision problem of which meaning to choose (van Rooij, 2003, and Nelken & Shan, 2004, referred to this as the "the asker's decision problem"). Although this line of reasoning has not been explicitly suggested for multiple wh-questions, in our view nothing speaks in principle against the possibility that multiple wh-questions have a semantically underspecified meaning as well, and also require consideration of the asker's intention to determine the reading. In this case multiple wh-questions would ask for the smallest list of pairs or triples that gives optimal relevance; what counts as optimally relevant would be different dependent on the specific context. (Note, however that this would be at odds with the fact that such questions always require exhaustive answers, namely, one would consider the asker's intention but the optimal list would always be the maximal, exhaustive one). According to this non-grammatical approach, exhaustivity in both question types is guided by Theory of Mind (henceforth also referred to as TOM): a person answering a single or a multiple wh-question needs to take into account the specific context in which the question is asked, including the speaker's intention, and then may or may not answer with an exhaustive list (see Section 1.2 for a definition of TOM).

Another possibility is that exhaustivity, both in single and in multiple wh-questions, is **grammatical** in nature. Under this assumption, the exhaustivity requirement is exerted by means of the structure of the question and its underlying syntactic/semantic representation. The central assumption of such an approach is that the exhaustive reading is derived via a universal quantifier in the semantic representation of the fronted wh-word. This universal quantifier dictates that all elements fulfilling the property in the question are listed (i.e., they are exhausting the question domain). Consider for example the question *Who is reading what?*. The semantic representation of the fronted wh-word *who* contains a universal quantifier that is interpreted as 'list all individuals who ...'. This 'grammatical' approach has been explicitly proposed for multiple wh-questions (see Bošković, 1998; Hagstrom, 1998) and for embedded wh-questions (e.g., Groenendijk & Stokhof, 1984) (see also the overview in Onea & Zimmermann, in press). Extending this grammatical account to single wh-questions, exhaustivity in single wh-questions would also involve a universal quantifier. Accordingly, both multiple and single wh-questions would necessarily include universal quantification, which obligatorily triggers exhaustive answers. However, a purely grammatical account for single wh-questions can already be ruled out, because single wh-questions also allow for non-exhaustive mention-some answers (e.g., Hamblin, 1973). This indicates that single wh-questions do not always contain a universal quantifier.

Let us now consider the third possibility: that the exhaustivity requirements in single and in multiple wh-questions are different in nature. The two question types differ in that multiple wh-questions, unlike single wh-questions, include two (or more) wh-words that stand in a **mapping** relation. Mapping is a function that takes every domain member satisfying the predicate in the question and maps it to the set of objects of this predicate. Consider the question *Who is eating what?*: here the domain of eaters is mapped onto the domain of food items. In this case mapping would pair each eater with what she is eating, so that an answer to a paired question would be an exhaustive paired-list (for the notion of mapping, see Krifka, 2001; for mapping in language acquisition research, see Schulz & Roeper, 2011; Schulz, 2015; Villiers, Ning, Liu, Zhang, & Jiang, 2018). It might be that this core difference between single and multiple wh-questions, both in terms of mapping and the manifestation of two (or more) wh-words, is the basis for the different exhaustivity requirements in the two question types.

<sup>1</sup> Multiple wh-questions presuppose a more-than-one-answer context in most languages, including Hebrew, the language tested in the current study. This presupposition seems to be absent only with echo-questions and with so-called quiz-questions such as *Who killed John F. Kennedy when?* (e.g., Dayal, 1996; Krifka, 2001). These were not part of our experimental design.

Given the three different approaches delineated above, neurolinguistic evidence regarding exhaustivity in wh-questions is instrumental, as it can inform us which of these accounts is more consistent with psychological reality. The aim of the current empirical investigation was to decide between these three approaches by studying a special population: adults who following brain damage show impairment in TOM and are hence impaired in their ability to evaluate the interlocutor's conversational needs, but whose grammatical abilities are unimpaired. If exhaustivity in both single and multiple wh-questions relies solely on TOM, these individuals are expected to fail in answering both types of questions, i.e., to not provide exhaustive answers in contexts in which unimpaired adults would do so. If exhaustivity relies solely on grammar, these individuals are expected to succeed on both wh-question types. If exhaustivity relies on TOM in single wh-questions and on syntax/semantics in multiple wh-questions, these individuals are predicted to fail to exhaust appropriately when answering single questions, which only rely on TOM, but should consistently exhaust when answering multiple wh-questions. The general rationale is this: if an individual is impaired in TOM but still provides appropriately exhaustive answers to a certain type of question, this will indicate that providing exhaustive answers in this type of question is not dependent upon TOM ability. Alternatively, if an individual (whose sole deficit is in TOM) shows difficulty in providing exhaustive answers in contexts where these are required, this will indicate that exhaustivity relies on TOM, i.e., on the need to fulfill the informational requirements in the discourse.

### 1.2. Theory of Mind and Theory of Mind impairment

Theory of Mind (TOM) is broadly defined as the ability to stand in someone else's shoes and see the world through their eyes. This cognitive ability allows one to understand that others may hold various mental understandings, beliefs, desires, emotions, and intentions, which may differ from their own, and which motivate people's behavior. In order to engage in meaningful discourse, conversation partners need to "read the mind" of the interlocutor and to tailor their utterances according to the interlocutor's understanding and informational needs (Brownell & Stringfellow, 1999; Martin & McDonald, 2003; Sperber & Wilson, 2004; for alternative approaches that ascribe a smaller role for TOM in conversational success see e.g., Pickering & Garrod, 2004). In these respects, TOM is a cognitive basis that is required for the language abilities involved in discourse pragmatics. Experimental TOM tasks require participants to consider the information to which others have access, and to take into account expectations and intentions of others (Apperly, 2011; Happé, 1994; Happé, Brownell, & Winner, 1999; Myers, 2001; Surian & Siegal, 2001).

Acquired TOM impairment has mainly been studied in patients who suffered brain damage in the right frontal and temporal brain areas (Amodio & Frith, 2006; Balaban, Friedmann, & Ziv, 2016; Happé et al., 1999; Frith & Frith, 2003; 2006). These patients have been reported to show difficulties in discourse understanding (Blake, 2006, 2007; Johns, Tooley, & Traxler, 2008), in inferring the beliefs others' convey (Apperly, Samson, Chiavarino, & Humphreys, 2004), and in monitoring intentions (Danckert, Stöttinger, Quehl, & Anderson, 2012). Their contribution to conversation is described as tangential and egocentric, irrelevant, off topic or incorrect, digressive, and centered on isolated details (Blake, 2006; Johns et al., 2008). Other studies showed that patients with right hemisphere damage have difficulties in articulating requests (Brownell & Stringfellow, 1999), appreciating jokes (Cheang & Pell, 2006), and attending to implicit details in stories and integrating them to narratives (Bartels-Tobin & Hinckley, 2005; Frestl, Walther, Guthke, & von Cramon, 2005; Marini, Carlomagno, Caltagirone, & Nocentini, 2005). Many of the studies mentioned above studied the discourse abilities of patients after right hemisphere brain damage at the group level, analyzing all right-hemisphere damaged patients together. Other studies employed a different approach and used a battery of TOM tasks (aTOMia battery) to identify two subtypes within the group of patients with right hemisphere damage (Balaban, Friedmann, & Ziv, 2016) as well as within the group of patients after traumatic brain injury (Balaban, Biran, & Sacher, 2019). The aTOMia battery allowed identifying individuals with impaired TOM abilities (patients with aTOMia) and individuals who, despite their brain damage, are not impaired (non-aTOMic). Testing specific linguistic abilities, it was found that patients with aTOMia also showed a specific impairment in their use of referring expressions. For example, they used a full name instead of a pronoun after a person had already been introduced to the discourse, and pronouns when a person was introduced in the discourse for the first time (Balaban, Belletti, Friedmann, & Rizzi, 2016; Balaban, Friedmann, & Ariel, 2016; similar findings were also reported in Chantraine, Joannette, & Ska, 1998; Sherratt & Bryan, 2012). Their syntax, on the other hand, can be intact (Balaban & Friedmann, 2014; Balaban, Friedmann, & Ariel, 2016; Brady, Armstrong, & Mackenzie, 2006; Marini et al., 2005). This asymmetry between good performance on grammatical tasks and low performance on TOM-related tasks was shown in a study testing comprehension of subject and object wh-questions in patients with TOM impairment (Balaban & Friedmann, 2014). The test was exactly the same in the grammatical and TOM-related conditions, but the form of response differed: the participants were asked wh-questions about pictures that included three figures (e.g., *Who is washing the boy?*, when presented with a picture of a blue elephant washing a boy who is washing a purple elephant, see Friedmann & Szterman, 2011; Friedmann & Novogrodsky, 2011). In the grammatical, non-TOM condition, when requested to respond by pointing to the appropriate figure in a picture, the patients with aTOMia performed flawlessly, showing normal syntactic abilities regarding wh-question comprehension. In the TOM condition, the same patients were asked the same wh-questions, but this time they were requested to respond verbally, i.e., to describe the correct figure in a way that would allow the hearer to distinguish between two similar figures, e.g., between a purple and a green elephant. In this condition, the patients failed to provide the relevant answer, because of their inability to attend to the other person's informational needs. For example, they would answer "the elephant", when there were two elephants in the pictures that differed in their color.

In the current study we used this dissociation between normal syntactic ability and impaired TOM in adults with brain damage in the right hemisphere to examine exhaustivity in wh-questions. We asked whether TOM is needed to consistently provide exhaustive answers to single and multiple wh-questions. If patients with aTOMia show difficulty in providing exhaustive answers when these are required in a given context, this would indicate that this ability is dependent upon the consideration of others' point of view. On the

other hand, if these patients are able to provide appropriate exhaustive answers, this would indicate that exhaustivity does not rely on ToM ability, and specifically not on evaluating the speaker's intention. Importantly, if individuals with TOM impairment perform differently on the two types of wh-questions, it suggests that the exhaustivity requirement originates from different sources.

## 2. Experiment 1: exhaustivity and aTOMia in single and paired questions

### 2.1. Method

#### 2.1.1. Participants

A group of 19 right-hemisphere damaged patients took part in the study, 8 of them female and 11 male (see [Appendix A](#), Table A1 for a detailed description of the participants). Their ages ranged between 32 and 67 years ( $M = 53$  years;  $SD = 10$ ). Fourteen of the participants were native Hebrew speakers, and 5 others moved to Israel before they were 10 years old and have spoken Hebrew for more than 55 years. Fourteen of the patients sustained one incident of CVA (cardiovascular accident, i.e., stroke) in their right hemisphere, and two sustained recurrent strokes (YA and YI). One of the CVA patients (MI) also suffered vertobasilar meningioma. Two other patients were surgically treated for removal of a tumor, one (AS) in the right parieto-frontal-temporal area and the other (HA) in the right parietal lobe. Nine of the participants had left hemispatial neglect, a neurological phenomenon characterized by a difficulty to attend to the left side of the visual field. One participant (SA) had hemianopia, a loss of visual ability to half of the visual field. All participants were tested at least 2 months post their brain damage.

The aTOMia battery ([Balaban, Friedmann, & Ziv, 2016](#); [Balaban, Ziv, & Friedmann, 2010](#)), a comprehensive test of Theory of Mind, was administered to all participants. The battery includes items testing eight aspects of TOM, two items per category. Each category presents a type of social scenario that requires inference of intentions, beliefs, feelings, and expectations of another person, and identification of gaps in knowledge and point of view. The eight categories included in the battery are first order false belief, 2nd order false belief, understanding knowledge gaps, understanding scenarios in which teaching was initiated, understanding white lies, comprehending embarrassing social situations (faux pas), comprehending the need to hide emotions, and mental-image cartoons ([Baron-Cohen, O'Riordan, Stone, Jones, & Plaisted, 1999](#); [Happé, 1994](#); [Happé et al., 1999](#)) (see [Appendix B](#) for details of the aTOMia battery).

Performance on the TOM battery indicated that 14 of the 19 right-hemisphere damaged participants had a severe deficit of TOM (aTOMia), with scores ranging between 13% and 60% correct ( $M = 43\%$ ,  $SD = 13\%$ ; the average performance on this battery of healthy adults in ages similar to the participants in the current study is 98.7%,  $SD = 2.3$ ; none of the controls made more than a single error. [Balaban, Friedmann, & Ziv, 2016](#)). These 14 participants with aTOMia showed a deficit in evaluating intentions, beliefs, and differences in points of view. The other 5 participants did not show aTOMia, with scores of above 80% correct on the aTOMia battery; two of them scored 100% correct, and the three others scored 89%, 88%, and 81% correct.<sup>2</sup>

Furthermore, we assessed the syntactic abilities of the participants using three tasks that examined their comprehension of the thematic roles of the arguments in relative clauses and wh-questions, their interpretation of pronouns, and their production of embedded sentences. These domains were chosen because they are directly related to the grammar of wh-questions, and are known to be clinical markers for grammatical impairments in agrammatic aphasia ([Friedmann, 2006](#); [Grodzinsky, 1984](#); [Grodzinsky, Piñango, Zurif, & Drai, 1999](#); [Zurif & Caramazza, 1976](#)). These tasks indicated that their syntactic abilities were normal (see [Appendix C](#) for a detailed description of the tests we used to assess the participants' syntactic abilities and the participants' good performance on these tasks).

In addition to the 19 patients, a group of 10 healthy adults (3 women and 7 men) with no known language impairment or brain damage served as a control group for the exhaustivity task. Their ages ranged between 29 and 70 years ( $M = 51.8$ ,  $SD = 15.0$ ).

#### 2.1.2. Materials and procedure

Exhaustivity was tested with a question-about-a-picture task. The experimenter showed the participant a picture and then asked a single or a multiple wh-question, while the participant was looking at the picture. In each picture several family members were displayed, and a varying number of them were performing the action asked about. Before the task started, the experimenter showed the participant a picture of all the characters included in the task, and introduced them as mother, father, grandfather, grandmother, boy, and girl to minimize memory effects. The experimenter then presented another family picture, and the participant was requested to name each character (see [Schulz, 2015](#), and [Schulz, Gavarro, Prévost, & Friedmann, 2018](#) for a detailed description and rationale of this task).

The question-about-a-picture task included single and multiple wh questions. It was presented in two versions, each including 30 items. Version A included 10 single wh-questions (see example 1), 10 paired wh-questions, and 5 singleton-answer fillers; Version B included 8 single wh-questions, 12 multiple wh-questions (8 paired, and 4 triple wh-questions, see Example 3), and 10 singleton-answer fillers. Version A included 5 additional items that were paired wh-questions with an overt exhaustivity marker "everyone", which explicitly requires mentioning all (see example 4).<sup>3</sup>

<sup>2</sup> Eight of the participants with aTOMia (DL, ME, NI, DR, JO, ZE, AS, MH) also took part in a different study that involved an extensive test battery assessing the effect of their cognitive difficulty on the use of language, i.e., pragmatics. Their data are reported in [Balaban et al. \(2016\)](#). The performance of these eight patients showed that none of them took into account the interlocutors' intentions or knowledge state in their use of reference terms in conversation.

<sup>3</sup> The two versions both reflect participants' knowledge about exhaustivity in wh-questions to the same degree. They both contain the same



Fig. 1. Picture used in single wh-question condition.

(1) *Mi maxzik kaduregel?*  
 ‘Who hold.3<sup>rd</sup>.sg.mas a soccer ball?’  
 ‘Who is holding a soccer ball?’

In version A the paired questions block appeared first, and the single questions block was second. (The wh-questions with an overt exhaustivity marker were presented as the last block). In version B, the first block included single wh-questions, the second included paired wh-questions, and the third included triple wh-questions.

The singleton fillers were dispersed across the single and multiple test conditions. The fillers in Version A were 5 single questions, in version B these were 4 single, 4 paired, and 2 triple questions. The singleton fillers appeared in each block, according to the question type in this block (single individual in the single questions block, single pair in the paired questions block, single triple in the triple block). The explicit wh-questions were included in version A to assess whether the occurrence of two wh-words per se increase complexity, therefore they contained only one wh-word and a quantifier. These questions also require an exhaustive paired-list as an answer. Eleven participants received Version A; seven participants received Version B, and one participant received both versions.

The inclusion of variations in the number of characters in the pictures and the inclusion of single, paired, and triple wh-questions requiring singleton answers assured that participants could not develop guessing strategies, such as listing all individuals in the picture or consistently responding with the same number of individuals. (Previous studies with healthy adults and children had shown that the number of figures or the number of individuals satisfying the predicate did not affect participants’ performance, which allowed us to introduce variation in these factors to prevent guessing strategies.)

The verbs used in both versions were high-frequency verbs (e.g., drink, eat, read, sit, give). All single wh-questions used the same wh-pronoun *mi* ‘who’ to achieve comparability across items. The paired wh-questions were construed with a fronted animate subject wh-pronoun *mi* ‘who’ and an object wh-pronoun in situ (*et mi* ‘whom’; *ma* ‘what’). Superiority effects were not tested in this experiment, so the order was always: wh<sub>1</sub>-subject wh<sub>2</sub>-non-subject. The triple wh-questions included an adjunct or a prepositional phrase after the object pronoun in situ (*eifo* ‘where’; *le-mi* ‘to whom’). All questions were non-embedded main-clause questions.

Single wh-questions are illustrated in (1) and Fig. 1, paired wh-questions and triple wh-questions are illustrated in (2) and (3), and in Figs. 2 and 3 respectively, and explicit wh-questions are illustrated in (4), with Fig. 4.

In the explicit wh-questions like (4) the object wh-pronoun was fronted, and the subject was a universally quantified NP (the Hebrew equivalent of “everyone”). The universal quantifier was used as an explicit cue for the exhaustivity requirement.

The study was conducted in two public rehabilitation centers in Israel and was approved by the respective Helsinki ethical committees. All participants agreed via written consent to participate in the study. They were tested individually in a quiet room in the rehabilitation centre in which they were hospitalized. All sessions were audio-recorded for later data check against the onsite-coding and for further individual analyses. No response-contingent feedback was given by the experimenter. When the participant failed to supply an answer, the item was repeated once. The control participants were tested in their homes.

### 2.1.3. Coding

Both oral and pointing responses to the single wh-questions were coded as correct if they exhaustively listed all the characters that matched the property asked about (both by listing all the relevant NPs, or by exhausting through exclusion, e.g., *everyone but the girl*).

(footnote continued)

relevant test conditions being administered in the very same manner, and crucially, previous studies have shown that participants are very consistent in their answers (e.g., Fekete, Schulz, & Ruigendijk, 2018; Schulz & Roeper, 2011). The change of version was introduced to allow for future cross-linguistic comparison with languages that were using version B. It was not our aim to compare the versions, which were very similar in the relevant dimensions, so we did not balance or randomize the versions given to the participants: the first participants tested received version A, and the ones who were tested later received version B, the first person tested with B was also tested with A.



Fig. 2. Picture used in paired wh-question condition.

(2) Mi oxel ma?

'Who eat.3<sup>rd</sup>.sg.mas what?'

'Who is eating what?'



Fig. 3. Picture used in a triple wh-question condition.

(3) Mi noten ma le-mi?

'Who give.3<sup>rd</sup>.sg.mas what to-whom?'

'Who is giving what to whom?'



Fig. 4. Picture used in explicit wh-questions items.

(4) Ma kol exad po shote?

'What all one here drink.3<sup>rd</sup>.sg.mas?'

'What does everyone here drink?'

For the paired questions, oral responses that contained exhaustive pairs were counted as correct. For example, the correct answer to question (2) is “*The father is eating a banana, the girl an apple, grandma cake, and mother ice-cream*”. Note that the two characters who are not eating anything do not need to be mentioned. Non-exhaustive responses, i.e., responses that did not mention all matching pairs, were counted as incorrect, as were unpaired responses that listed subjects and objects separately, or responses that listed only subjects or only objects.

## 2.2. Results

### 2.2.1. General results: single vs. paired (and triple) questions

The most important result of this study was the entirely different pattern of responses that the participants with aTOMia provided for the single and paired wh-questions. The participants with aTOMia provided exhaustive responses to only 66% of the single wh-questions (SD = 26%), whereas they provided exhaustive responses to 90% of the paired wh-questions (SD = 14%) and 100% exhaustive responses to the triple wh-questions. (None of the participants mentioned a character that did not perform the action asked about.) The difference between the single and paired questions scores was significant (Wilcoxon signed ranks test,  $T = 5$ ,  $p = .001$ ), and so was the difference between the single and triple questions for the 6 participants with aTOMia who received triple questions (Wilcoxon signed ranks test,  $T = 0$ ,  $p = .03$ ). These comparisons and all following comparisons remained significant after applying FDR correction for experiment-wise multiple corrections (Benjamini & Hochberg, 1995). The two other groups consistently exhausted on all question types: the participants with right-hemisphere brain damage and intact TOM abilities performed at ceiling on both types of wh-questions (100% exhaustive responses). The healthy control group performed very close to perfect, only a single control participant failed to exhaust in response to one single wh-question ( $M = 99.7\%$ ,  $SD = 1.1\%$ ).

### 2.2.2. Exclusion of non-exhaustive responses that may result from neglect

Notably, eight of the 14 patients with aTOMia had left visuo-spatial neglect, a condition that involves a deficit in orienting attention to the left side of space or left visual field, leading to inability to report visual events occurring in the left side, although basic visual abilities are intact.<sup>4</sup> This neglect may have led them to omit figures that appeared on the left side of the picture. To rule out an alternative explanation according to which patients' omissions mainly or exclusively resulted from neglect, we re-analyzed the responses of all the participants who had neglect and excluded from the analysis all non-exhaustive responses that could have resulted from neglect rather than from impaired exhaustivity (see Table 2). Following this reasoning, we excluded responses in which the omissions affected only the one or two leftmost figures in the picture (“left omissions”). If a participant with neglect did not answer exhaustively, but their response only failed due to left omissions, their answer was excluded from the total number of items and was not coded as an error. This lenient coding ensured that only true failures to exhaust were considered in the analysis.<sup>5</sup>

**2.2.2.1. Single vs. paired (and triple) questions: an analysis without left omissions.** The general pattern of responses to single vs. paired questions did not change when left omissions, i.e., errors that may be ascribed to neglect, were excluded. The results remained the same, with failure to exhaust in response to single questions, and exhaustive answers to paired and triple questions. The data excluding neglect errors are summarized in Fig. 5; Table 1 provides data of individual performance in the various conditions.

The group of patients with aTOMia showed a clear difference in their response patterns to different question types: they consistently provided exhaustive answers to paired and triple questions but not to single wh-questions. The participants provided significantly more exhaustive answers to the paired wh-questions ( $M = 93.6\%$ ,  $SD = 8.4\%$ ) and to triple wh-questions ( $M = 100\%$ ) than to the single wh-questions ( $M = 74.1\%$ ,  $SD = 23.1\%$ ) ( $T = 3.5$ ,  $p = .001$ , for the single vs. paired questions comparison;  $T = 0$ ,  $p = .03$  for the single vs. triple questions comparison). Only 3 of the 14 individuals with aTOMia provided exhaustive answers to all single wh-questions.

The brain-damaged individuals with intact TOM provided exhaustive responses consistently and across all question types. The performance of the individuals with aTOMia in comparison to the brain-damaged individuals with intact TOM and to the healthy control group is summarized in Fig. 5. As for single wh-questions, the participants with aTOMia differed significantly from both the group of brain-damaged individuals with intact TOM ( $U = 7.5$ ,  $p = .006$ ) and from the healthy control group ( $U = 16.5$ ,  $p = .001$ ). Regarding paired wh-questions the participants with aTOMia scored similarly to the brain damaged participants with intact TOM ( $U = 20$ ,  $p = .09$ ) and only marginally poorer than controls ( $U = 40$ ,  $p = .04$ ). On the triple wh-questions all groups showed identical performance providing correct exhaustive responses in 100% of the triple questions.

Performance on the filler items requiring singleton answers was 100% correct for participants with aTOMia as well as for the two control groups. On the wh-questions that included an explicit marker of exhaustivity, “everyone”, all participants with aTOMia performed 100% accurately (with the exception of one response of one participant; the two control groups were also 100% correct).

These findings support an account according to which grammar suffices to dictate an exhaustive response for paired (and triple) wh-questions. Consequently, even when TOM is impaired, the grammatical properties of multiple wh-questions prompt exhaustive

<sup>4</sup> Two additional patients with good TOM had left visual field difficulties as well: SA had hemianopia and YI had neglect. They, like all patients with good TOM, performed at ceiling, so their responses did not warrant any reanalysis.

<sup>5</sup> Another indication that failure to exhaust could not have resulted solely from neglect is that 6 of the 19 patients tested in Experiment 1 showed aTOMia but did not suffer from neglect. These participants performed significantly poorer on single questions (88.3% correct,  $SD = 10.2$ ) than the patients without aTOMia and the control participants ( $U = 7.5$ ,  $p = .006$ ;  $U = 16.5$ ,  $p = .001$ , respectively).

**Table 1**

Experiment 1: Mean percent of correct responses to single, paired, triple, and explicit wh-questions excluding left omissions (i.e., omissions of target figures on the leftmost area of the page).

	Participant	TOM score	% exhaustive responses to questions			
			Single	Paired	Triple	Explicit
Participants with aTOMia	EZ <sup>n</sup>	13%	67%	88%	100%	
	DR <sup>n</sup>	31%	50%	100%	100%	100%
	DL <sup>n</sup>	32%	100%	100%		80%
	MI	35%	80%	100%		100%
	NI <sup>n</sup>	37%	86%	100%		100%
	MO	38%	88%	75%	100%	
	ME <sup>n</sup>	41%	71%	89%		100%
	HA	43%	88%	100%	100%	
	JO <sup>n</sup>	47%	38%	100%		100%
	YA	50%	100%	88%	100%	
	MH	53%	100%	100%		100%
	ZE <sup>n</sup>	58%	25%	83%		100%
	AS <sup>n</sup>	59%	71%	100%		100%
EL	60%	75%	88%	100%		
Participants with intact TOM	YI <sup>n</sup>	81%	100%	100%	100%	
	EG	88%	100%	100%	100%	
	SA <sup>H</sup>	89%	100%	100%		100%
	AN	100%	100%	100%		100%
	OF	100%	100%	100%		100%
Control group (n = 10)		–	99.0%	100%	100%	

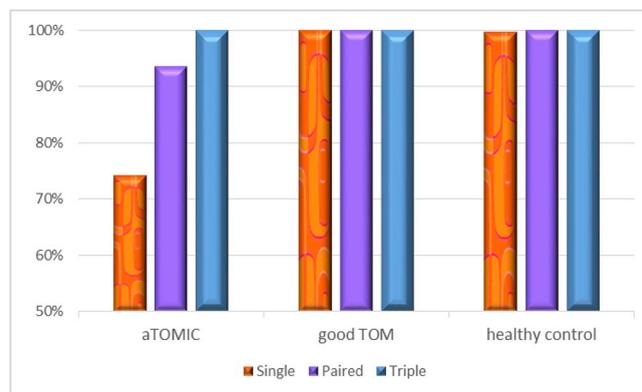
<sup>n</sup> Participants with neglect, <sup>H</sup> Participant with hemianopia

**Table 2**

Experiment 1: Number of omissions of figures on the left and on the non-left in single and multiple (paired/triple) questions of the participants who had both aTOMia and neglect.

Participants	Neglect omissions (leftmost figures)		Other omissions (non-left figure)	
	Single	Multiple	Single	Multiple
DL	0	0	0	0
NI	3	1	1	0
JO	2	0	5	0
ME	3	1	2	1
AS	3	0	2	0
ZE	6	4	3	1
DR	0	0	4	0
EZ	2	0	2	0 <sup>a</sup>
Total	19	6	19	2

<sup>a</sup> EZ also made one mapping error (which did not involve a figure omission).



**Fig. 5.** The performance (% correct exhaustive answers excluding left omissions) of the brain-damaged group with aTOMia, the brain-damaged group with good TOM and the healthy controls in single, paired, and triple wh-questions.

answers. In contrast, single wh-questions do not contain indications of the requirement to exhaust within their sentence structure; providing an exhaustive answer to such questions depends on TOM. More specifically, in order to decide between the two readings of single wh-questions, mention-some and mention-all, the participants must consider the intentions of the person who asked the question. In our experimental setting, all wh-questions require the exhaustive, mention-all, reading, as confirmed by the responses of the healthy adults (see also the previous results on adults in Schulz & Roeper, 2011). Individuals with TOM impairment and preserved linguistic abilities were hence only affected in their ability to provide exhaustive responses to single wh-questions.

In a further analysis, we examined whether participants failing to answer single wh-questions exhaustively mentioned a single character or provided a non-exhaustive 'plural' list (see Schulz, 2015; Schulz & Roeper, 2011). This analysis included the relevant single wh-question items in which the pictures showed more than two characters performing the action asked about. There were 5 items in Version A and 4 items in Version B of this type. This analysis revealed that most of the 24 non-exhaustive answers the participants with aTOMia provided to these items (20/24, 83%) were plural answers and only 4 were singleton answers.

**2.2.2.2. Hemispatial neglect and exhaustivity.** The analysis of the neglect errors of the 8 participants with aTOMia and neglect revealed a further interesting finding. It showed an effect of language on hemispatial neglect, whereby exhaustivity guided exhaustive attention scanning of the picture. These participants omitted the leftmost figures in 16% of their responses (i.e., 25 left omissions out of 160 single and multiple wh-questions). Importantly, these neglect errors were not evenly distributed between the single and multiple wh-questions. In the single wh-questions the neglect error rate was 25% (19 out of 76 items), compared to only 8% neglect errors in the paired wh-questions (6 out of 76 items), and no omissions in the triple wh-questions (0 of 32 items). The rate of left neglect errors in single questions was significantly higher than in multiple questions,  $\chi^2 = 15.2$ ,  $p < .0001$ . The data are summarized in Table 2.

The performance of the participants with impaired TOM clearly shows that neglect is not the main source of their failure to consistently provide exhaustive answers to single wh-questions. First, all participants with aTOMia and neglect who omitted figures on the left also made additional omission errors that were not on the left: they had 19 non-left omissions on single wh-questions (which constituted 40% of their total errors to single questions). Moreover, if the failure in single wh-questions was caused by neglect only, it would be difficult to explain how the same patients with the same neglect in the same task succeeded to mention all relevant characters in multiple questions. Finally, not only the patients with neglect provided non-exhaustive answers: 5 of the 6 aTOMic patients without neglect also provided non-exhaustive answers to single wh-questions.<sup>6</sup> Thus, we conclude that although some of the participants had neglect in addition to aTOMia, their failure to exhaust did not result solely from their neglect.

### 3. Experiment 2: interaction of visuo-spatial neglect and exhaustivity in wh-questions

The results from Experiment 1 (Section 2.2.2.2) indicate that grammatical exhaustivity requirement modulates neglect errors. Experiment 2 was designed to address this effect in more detail. To this end, we rearranged the characters on the pictures such that in each picture one or two characters on the leftmost side and one or two figures at the rightmost side matched the property asked about. Hence, an appropriate response to the wh-question always required shifting attention to the left side of the picture.

#### 3.1. Method

##### 3.1.1. Participants

Six additional participants with right hemisphere brain damage and with neglect were tested, three women and three men. They had been diagnosed with neglect according to the standardized battery for the diagnosis of neglect, the Behavioural Inattention Test (BIT, Halligan, Cockburn, & Wilson, 1991) in the rehabilitation center prior to our study.<sup>7</sup> Their ages ranged between 48 and 68 years ( $M = 57$ ,  $SD = 7.6$ ). They were tested at least 2 months post-onset of their brain damage. Five of the six patients with neglect had aTOMia; one participant (RU) had neglect but normal TOM (see Appendix A, Table A2, for participants' details).

##### 3.1.2. Materials

We used the visual material of 18 of the pictures that were used in Experiment 1. The pictures were modified such that there were figures on the leftmost and on the rightmost side of the page performing the action asked about (e.g., eating, sitting, reading). Each picture was presented twice, in two different sessions, once with a single wh-question and once with a paired wh-question. For half of the pictures it was first presented with a single wh-question, and for the other half it was first presented with a paired question (see Fig. 6). One of the single and one of the paired wh-questions were control items requiring singleton answers. Our initial aim in using each picture twice was to allow for a minimal comparison across conditions. However, it turned out that participants often remembered the wh-question from the previous session: sometimes, in their second encounter with the same picture, they responded as they did during its first presentation. This created a situation in which participants gave a paired answer to a single wh-question, which was the required answer to the paired wh-question asked in the first session. Therefore, our main analysis included only the

<sup>6</sup> MH, who is reported in Table 1 as providing 100% exhaustive responses, actually gave 3 non-exhaustive answers to single questions. As these involved an omission of leftmost figures, they do not appear in Table 1, which excludes omissions of the left figure for all patients (with and without neglect).

<sup>7</sup> BIT raw scores were available for YS: 60/146, TM: 117/146, and OK: 101/146.

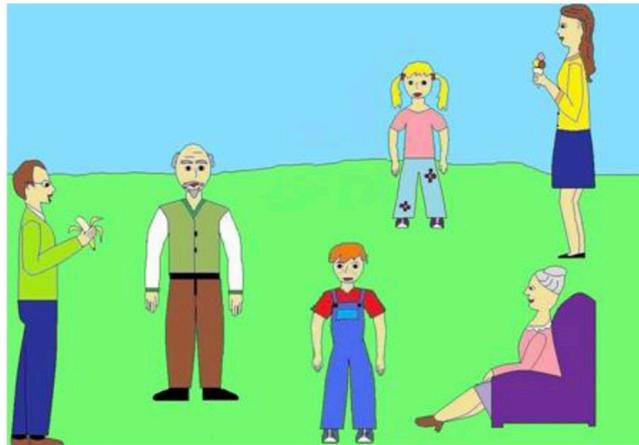


Fig. 6. An example for a picture used in Experiment 2 with the questions *Who is eating what?* and *Who is eating?*.

items presented in the first session, resulting in a total of 108 responses (9 items for each question type for the 6 participants).

### 3.2. Results

As in Experiment 1, the results show a clear interaction between the type of exhaustivity requirement in the question and neglect. When asked a single wh-question, the patients with neglect produced exhaustive answers in only 56% of the cases (SD = 19%). However, when they answered paired wh-questions, which grammatically required exhaustive paired-list answers, the participants with neglect provided correct exhaustive responses to 92% of the questions (SD = 10%). This difference between exhaustive answers given to single and paired questions was significant,  $T = 0, p = .03$ .

This pattern of more omissions on the left in single compared to multiple questions, which we saw in Experiment 1, was replicated here and applied to each of the individual participants (the individual scores are given in Table 3). All patients with neglect omitted more figures on the left side of the picture in single wh-questions than in multiple wh-questions. Left omissions occurred in 40% of their answers to the single wh-questions (16 of a total of 40), but only in 7.5% of their answers to paired wh-questions (3 of 40). This difference was statistically significant for two of the participants, despite the relatively small number of items,  $\chi^2 = 7.27, p = .007$ , for both of them.

In summary, two factors affected these patients' failure to exhaust in their responses to single wh-questions: their aTOMia and their visuo-spatial neglect. Their aTOMia caused omissions of figures that appear on both sides of the picture, and their neglect dictated omissions of figures on the left. Importantly, when they were asked a paired wh-question, neither their aTOMia nor their neglect affected their responses. We take this finding to indicate that the structure of the paired wh-questions drove the participants to exhaust the question domain and complete their shifting of attention all the way to the leftmost part of the picture, resulting in fully exhaustive paired answers.

Recall that the second session was not part of the main analysis. Nevertheless, the participants' responses provided interesting insights into their understanding of exhaustivity in multiple wh-questions. In the second session, in which the same pictures were shown with a different question, we noted an intriguing response pattern. When the participants were asked a single wh-question about the same picture that had been shown before with a paired wh-question, in 14 out of 54 cases the participants responded with a paired answer even though they were asked a single wh-question. For example, when shown the picture in Fig. 5 and asked *Who is*

**Table 3**  
Experiment 2: Individual results of patients with neglect on single and paired wh-questions.

	Participant	TOM Score	% exhaustive answers		Number of responses with omissions of left figures/ total non-exhaustive responses	
			Single wh-questions	Paired wh-questions	Single wh-questions	Paired wh-questions
Participants with neglect and aTOMia	RB	36%	63%	100% <sup>a</sup>	2/3	0
	DS	50%	50%	88%	3/4	1/1
	YS	50%	38%	100%	4/5	0
	TM	50%	88%	100%	0/1	0
	OK	56%	63%	75%	2/3	2/2
Participant with neglect and intact TOM	RU	100%	38%	100%	5/5	0

<sup>a</sup>RB also made one mapping error (which did not involve a figure omission).

eating?, participants answered with an exhaustive paired-list answer like *The mother is eating ice-cream and the father is eating a banana*. These 14 paired-list answers provided by the participants were all exhaustive and did not leave any of the relevant figures unmentioned. In our view this surprising pattern provides an important insight into the trigger of exhaustivity. It is not only the question a person hears that creates the need to exhaust; it seems that once a paired-list answer was planned, the syntactic-semantic structure of the planned response forced the participants to exhaust and shift their attention to the whole picture.

Another interesting result comes from RU, a 61 year old woman with visuo-spatial neglect who had normal TOM abilities. In 62% of her answers to single wh-questions she omitted the figures on the left, but in her responses to paired wh-questions she never omitted figures on the left (100% correct). This contrast shows that the semantic-syntactic requirements of paired wh-questions affect exhaustive search even when the source of omissions is exclusively hemispatial neglect.

### 3.3. Experiment 1 and 2: summary of results

The two experiments show that providing an exhaustive answer to single wh-questions, which do not overtly indicate the requirement to exhaust, can go wrong as a result of TOM impairment. The 19 participants with aTOMia (fourteen in Experiment 1 and five in Experiment 2) offered only 63.2% correctly exhaustive answers to the single questions; this rate was significantly smaller than for multiple wh-questions (91.5%) ( $T = 1.0, p = .0003$ ). This comparison was also significant when we only included responses that could not be explained as neglect errors (i.e., responses in which figures that appeared in the middle or on the right were omitted),  $T = 1.0, p = .0003$ . The performance of the 19 participants with aTOMia on single wh-questions in the two experiments was significantly poorer than that of brain-damaged patients without aTOMia ( $U = 7.5, p = .002$ ) and that of control participants ( $U = 16.5, p = .0002$ ). The large difference in performance between the answers to single and paired questions given by individuals with aTOMia can be explained as follows: single questions include no indication for the need to exhaust and the responder needs to consider the asker's intention, whereas the structure of the paired (and triple) questions provides grammatical directions to exhaust.

Furthermore, we found that hemispatial neglect, which typically causes impaired scanning of the left side of the visual field, is sensitive to the syntactic-semantic properties of wh-questions. These patients missed figures on the left in their answers to single wh-questions, but did not miss figures on the left in their answers to the paired questions. We suggest that the requirement to exhaust the domain, dictated by the paired questions and the mapping requirement, prompted their scanning of the picture to the left (for further discussion see Section 4.2 below).

## 4. Discussion

This study explored the nature of the requirement to provide exhaustive answers to wh-questions. To this end we compared multiple wh-questions (*Who eats what?*), which always require an exhaustive list of pairs, and single wh-questions (*Who eats ice-cream?*), which can either receive an answer with an exhaustive list of individuals (mention-all) or a mention-some answer. We were interested in whether the exhaustivity requirement is related to Theory of Mind (TOM) or to purely grammatical constraints, and whether the exhaustivity requirement in the two question types is of the same nature. We examined these issues through the answers provided by adults who have a selective deficit in TOM (aTOMia) following brain damage in the right hemisphere.

### 4.1. aTOMia sheds light on the different sources of exhaustivity requirement in single and paired questions

To shed light on the nature of the exhaustivity requirement we examined patients with aTOMia who have a selective deficit in TOM, but unimpaired grammar. The rationale was that they should perform poorly on tasks that critically rely on TOM abilities, whereas they should perform normally if they can use grammatical, non-TOM-related linguistic abilities to solve the task. This, in turn, would inform us about the nature of the different wh-questions. Questions to which the participants with aTOMia fail to provide appropriate exhaustive answers depend on TOM, and questions to which they provide appropriate exhaustive answers are anchored in grammar rather than in TOM.

The performance of the patients with aTOMia in both experiments revealed a sharp difference between their answers to the single and to the multiple wh-questions. They often failed to give exhaustive answers to single wh-questions, providing mention-some answers in which they named several but not all individuals that fulfill the property in question. In contrast, they consistently exhausted their answers to paired and triple questions, providing almost exclusively exhaustive paired-list answers.

These findings indicate that the sources of the exhaustivity requirements in single and multiple questions are different. In single wh-questions the exhaustivity requirement is not encoded in the grammatical structure of the question, and hence, the decision of whether to provide an exhaustive answer hinges upon TOM abilities, i.e., the ability to consider the speakers' information needs (Nelken & Shan, 2004; van Rooij, 2003).

Theoretical literature maintains that single wh-questions are ambiguous (Nelken & Shan, 2004; or underspecified, van Rooij, 2003) with respect to whether or not they require an exhaustive answer. Under the ambiguity analysis, the quantifier contained in the fronted wh-word is either universal or existential. The interlocutor needs to choose between these two readings, i.e., select one of the quantifiers, resulting in the mention-all or the mention-some reading. Crucially, this decision depends on reading the mind of the asker and on understanding which type of answer the question requires in the given discourse context. In this scenario, a TOM

impairment is detrimental and grammatical knowledge cannot come to the rescue.<sup>8</sup>

In marked contrast, the patients with aTOMia overwhelmingly provided appropriate exhaustive answers to paired and triple wh-questions. This indicates that exhaustivity in multiple wh-questions relies on a different mechanism: in order to achieve exhaustivity in multiple questions the interlocutor does not need to consider the speaker's intention and can rely solely on grammatical computation (be it semantic and/or syntactic). We suggest that the structure of multiple wh-questions includes a grammatical (syntactic and/or semantic) indication of the exhaustivity requirement, which led the patients with aTOMia to provide exhaustive answers.

Our basic idea is this: mapping requires exhausting the domain in the answer; multiple wh-words signal mapping, and therefore, multiple wh-words constitute an overt indication of the grammatical exhaustivity requirement. In the following we sketch this proposal in more detail.

Mapping is a function from a given and identifiable domain, i.e., from the individuals present in the discourse that are involved in the activity denoted by the predicate, to values, i.e., to the objects or individuals that are affected by the activity (see Krifka, 2001). Crucially, this function takes **every** domain member satisfying the predicate in the question and maps it to the set of objects of this predicate. Regarding multiple wh-questions, the calculation of a function from the first, fronted wh-expression (in our case, the subject) to the second, in situ wh-word (in our case, the object) results in exhaustive paired-list answers.

Therefore, the mapping procedure maps **all** relevant individuals in the domain related to the fronted wh-word. Indeed, linguistic analyses suggest a close connection between mapping in multiple wh-questions and the universal quantifier. The fronted wh-expression in multiple wh-questions is assumed to contain a universal quantifier (e.g., Dayal, 1996, 2017; Fox, 2012; Krifka, 2001; Schulz & Roeper, 2011).<sup>9</sup> Mapping, then, requires exhaustivity, and what signals mapping is the presence of two (or more) wh-words relating to the same predicate. Taken together, this means that the presence of two (or more) predicate-related wh-words constitutes an overt indication that exhaustivity is required. This, we suggest, is the basis for the grammatically-based exhaustivity requirement in multiple questions.

Since this requirement to survey the domain in full is rooted in the grammar of multiple wh-questions, individuals with intact grammar can invariably represent the multiple wh-question with an underlying universal quantifier, and hence can exhaust the domain. In the case of the wh-questions that we presented to the participants, this was sufficient for eliciting consistent exhaustive answers to the multiple wh-questions.<sup>10</sup>

According to such an approach, exhaustivity results from different processes in single and in multiple wh-questions. Multiple wh-questions include overt indication of mapping and hence a grammatical basis for exhaustivity. In contrast, single wh-questions do not involve mapping and include no signal dictating exhaustivity. This is in line with the observation mentioned above that single wh-questions can receive different readings, referred to as mention-some and mention-all. The different interpretations of single wh-questions have been argued to arise from the ambiguity between different underlying semantic representations (see Nelken & Shan, 2004; Roeper, Schulz, Pearson, & Reckling, 2007; Schulz & Roeper, 2011) or from different answer values within one semantic representation (e.g., van Rooij, 2003; Zimmermann, 2007). For the purposes of the current study, these two approaches yield identical predictions. For concreteness let us assume that the mention-some reading involves existential quantification and that the mention-all reading involves universal quantification (Nelken & Shan, 2004).<sup>11</sup> For example, if someone asks *Who has a phone?*, if the selected reading is the mention-some reading, the wh-word will contain an existential quantifier. In this case the answer will name one or two people who happen to have a phone. Under the mention-all reading, the wh-word contains a universal quantifier that requires the answer to list all people in the group who have a phone. Notice that the syntactic structure is the same for both readings, but the quantifiers inserted into it differ.<sup>12</sup>

Crucially, the ambiguity between the two readings requires the person answering a single wh-question to consider the interlocutor's point of view and her intentions in posing the question. This will guide the selection between the two possible readings (existential or universal). This is what makes the interpretation of single wh-questions dependent on TOM.<sup>13</sup> Multiple wh-questions,

<sup>8</sup> Note that according to this line of reasoning, we predict that patients with aTOMia, who have a deficit in reading the mind of the other and taking the others' intentions into consideration when providing an answer, will also not give consistent mention-some answers in contexts where this is the felicitous answer. They are predicted to sometimes provide an exhaustive answer instead of a mention-some answer.

<sup>9</sup> In general, multiple wh-questions are assumed to include only one universal quantifier (on the fronted wh-word). This quantifier takes those individuals within the domain that satisfy the predicate: the fronted wh-word defining the domain is assumed to be linked to a contextually given set of referents, a property referred to as 'D-linking' (Pesetsky, 1987). In the case of calculating a one-to-one function as in our experimental set-up, just as with one to-many functions, a full survey of the domain has to be undertaken.

<sup>10</sup> The participants' good grammatical abilities also allowed them to perform the mapping correctly, and to consistently provide appropriate paired responses to multiple wh-questions. This is in striking contrast to children with SLI, who sometimes mention only all subjects or only all objects (see Roeper, 2004; Roeper et al., 2007; Schulz, 2010).

<sup>11</sup> The observation that wh-questions have two readings leaves open whether the exhaustive or the non-exhaustive reading is the default (for the former, see Groenendijk & Stokhof, 1984; for the latter 'non-exhaustivity default', see Beck & Rullmann, 1999; Dayal, 1996, 2017; Hamblin, 1973; Karttunen, 1977; Reis, 1992; see also the overview in Onea & Zimmermann, in press). Here we emphasize the fact that the person answering the wh-question has to decide between the readings.

<sup>12</sup> Focusing on contexts that prompt mention-all answers to single wh-questions, a recent study on the felicity of exhaustive and non-exhaustive answers with German-speaking healthy adults (Fekete et al., 2018) revealed that adults' judgments of non-exhaustive plural answers to single wh-questions pattern with those to *who all*-questions and to plural definite descriptions embedding a restrictive relative clause, which are assumed to be obligatorily exhaustive for semantic reasons. The authors took these parallels to indicate the semantic nature of exhaustivity in single wh-questions.

<sup>13</sup> Note that this proposal is close in spirit to the 'semantic proposal' put forward in Schulz & Roeper (2011), with regard to the role of the interlocutor in deciding between the two readings. As the latter aspect is argued to be crucial, we hold that the notion 'TOM-related' characterizes this approach more adequately.

on the other hand, require a mention-all reading purely on the basis of the syntactic/semantic properties of the question, without consideration of the interlocutor's point of view and her intentions.

In the current study all answers involved mapping between pairs of one agent and one theme. It would be interesting in future studies to include one-to-many contexts (e.g., by asking *Who is holding what?* about a picture depicting one person holding a watermelon and a strawberry and another person holding a tomato and a carrot).

#### 4.2. The effect of grammar on attention shift in neglect

The findings from aTOMia shed light on the intricate relations between Theory of Mind and exhaustivity in two respects. First, individuals with impaired TOM did not consistently provide exhaustive answers to wh-questions in which only TOM dictates exhaustivity of the answer. Second, grammatical requirements can override the limitation caused by TOM impairment: the TOM-impaired patients provided exhaustive answers to the multiple wh-questions, because their intact grammar led them to do so.

Our findings reveal a similar bipartite effect of grammar on another domain: attention shift to the left. Individuals with hemispatial neglect often fail to shift their attention to the left. This was evident in the responses the individuals with neglect provided to single wh-questions. When the exhaustivity requirement was solely driven by the need to understand the speaker's intention in this discourse context, it was not strong enough to force them to shift attention to the left. However, the requirement in multiple wh-questions, driven by grammar (i.e., the need to exhaust the domain), was enough to force the patients with neglect to shift their attention to the left.

Such an effect of grammar on shifting attention to the left in patients with neglect has also been found with respect to the syntactic structure of the sentence, and to the syntactic requirements of verbs. [Friedmann, Tzailer-Gross, and Gvion \(2011\)](#) found that patients with neglect who omit the left side of sentences, do so less often when the omission renders the sentence ungrammatical. Even on the single word level, the morphological structure of the word was found to dictate how often the left side of the word would be neglected in neglect dyslexia ([Reznick & Friedmann, 2015](#)). These findings can be further explored in the context of rehabilitation. The presentation of linguistic stimuli that encourage the patient to shift her attention to the neglected half of the visual field may offer a distinct contribution to the treatment of this condition.

#### 4.3. The dissociation between impaired TOM and preserved grammar and its relation to language

The results of the two experiments thus show a clear dissociation between a preserved grammatical ability and an impaired ability in considering others' intentions when tailoring responses to wh-questions. A similar distinction was found in previous studies that focused on other types of linguistic abilities that require both types of processes ([Balaban, Belletti, & Ariel, 2016](#); [Balaban, Friedmann, & Ariel, 2016](#); [Balaban and Friedmann, 2014](#)). In one study we showed that participants with aTOMia demonstrated preserved ability to comply with the syntactic binding principles associated with C-Command configuration and an inability to comply to a discursive principle aimed to regulate the distribution of proper names ("repeated name penalty", [Ariel, 2001](#); [Gordon, Grosz, & Gilliom, 1993](#); [Gundel, Hedberg, & Zacharski, 1993](#)). Participants with aTOMia ruled out the possibility that two proper names refer to the same referent in C-Command configurations (e.g., "Rachel promised that Rachel will stop smoking" was judged as inappropriate, compared to "Rachel promised that she will stop smoking"). On the other hand, when the two proper names were presented in a configuration that did not violate the syntactic principle C, the same participants judged it as acceptable (e.g., "Tami registered to a gardening school because Tami loves trees and flowers" was judged as acceptable as the sentence "Tami registered to a gardening school because she loves trees and flowers") ([Balaban, Belletti, et al., 2016](#)).

## 5. Conclusion

The two experiments shed light on the effect of TOM impairment on exhaustivity in wh-questions and on the interaction between TOM, attentional, and linguistic abilities. The results also bear on the linguistic question of the nature of the exhaustivity requirement. The individuals with TOM impairment due to brain damage failed to exhaust their answers to single wh-questions, but consistently provided exhaustive answers to paired and triple wh-questions. This asymmetry suggests that TOM alone dictates exhaustivity in single wh-questions, whereas in paired and triple wh-questions the grammatical structure of these multiple wh-questions dictates the exhaustivity requirements. As a result, the intact grammar of patients with aTOMia saved them from making errors and triggered exhaustive answers to multiple wh-questions. A similar effect of grammar was found with respect to attention: hemispatial neglect led to omissions of figures on the left of the picture in single wh-questions. However, in multiple wh-questions their grammar dictated a complete search of the picture and hence a shift of attention all the way to the left, and in turn reduced left omissions.

These novel findings from adults with aTOMia following brain damage provide new empirical evidence regarding the nature of the exhaustivity requirement in wh-questions. Our results show that exhaustivity in single and multiple wh-questions does not rely solely and uniformly on TOM, as in this case these individuals should have failed answering both single and multiple wh-questions, contrary to our findings. Furthermore, the results indicate that exhaustivity does not rely solely on grammar in single and in multiple wh-questions, because if this were the case, these individuals should have succeeded with both question types, which they did not. Our results speak in favor of an approach according to which the source of the exhaustivity requirement in single wh-questions is different from that in multiple wh-questions. Whereas identifying the exhaustivity requirement in single wh-questions crucially relies on TOM, the exhaustivity requirement in multiple wh-questions can be identified through grammar alone (independent of whether TOM plays a role in multiple wh-questions as well). This is why patients with aTOMia fail to exhaust appropriately when they answer

single wh-questions, but succeed in providing exhaustive answers to paired and triple wh-questions, in which exhaustivity is determined via the semantics/syntax of the question.

In summary, the two experimental studies offer a unique insight into to the nature of the exhaustivity requirement in wh-questions. The present findings suggest that exhaustive responses to single wh-questions rely on the need to consider the askers' intentions, whereas the exhaustivity requirement in paired and triple wh-questions is encoded grammatically.

**Acknowledgment**

This research was supported by the Israel Science Foundation (grant no. 1296/06, Friedmann), and by the Human Frontiers Science Program (RGP0057/201, Friedmann).

**Appendix D. Supplementary data**

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jneuroling.2019.100853>.

**Appendix A. background information about the participants**

Table A1  
Participants in Experiment 1

Participant	Gender	Age	Spoken Language	Education	Neglect	Handedness	Lesion site and Etiology	Months post onset
DL	F	50	Hebrew, English	High school	yes	R	Ischemic infarct in the right hemisphere	4
NI	M	64	Hebrew, Arabic	High school	Yes	R	Ischemic infarct involving right fronto-temporal-parietal areas	6
MI	F	57	Hebrew, Romanian	Elementary	no	R	Recurrent right parietal parasagittal meningiomas, right parietal craniotomy and vertebrobasilar stroke	5
JO	M	51	Hebrew	High school	Yes	R	Ischemic infarct - thalamus, internal capsule hypodense area in basal nuclei on the right	3
ME	M	48	Hebrew	High school	yes	R	Ischemic infarct in the right MCA territory	8
MH	M	55	Hebrew, Arabic	High school	No	R	Ischemic infarct in the right MCA territory, complete block of right ICA	2
AS	M	36	Hebrew	High school	yes	R	Craniotomy for removal of right parieto-frontal-temporal tumor	7
ZE	M	65	Hebrew, English, Polish, Yiddish	Academic	Yes	R	Ischemic infarct in the territory of the right MCA	9
SA	F	62	Hebrew, English	Academic	Hemi-anopsia	L	Right occipital infarct. Spontaneous intraparenchymal hemorrhage	24
AN	F	32	Hebrew	High school	no	R	Ischemic infarct in the right MCA territory	4
OF	M	47	Hebrew	Academic	no	R	Ischemic infarct in the right MCA territory. Right periventricular infarct	7
HA	F	54	Hebrew	High school	no	R	Right parietal craniotomy and tumor removal	4
EZ	F	67	Spanish/Hebrew	High school	yes	L	Ischemic infarct in the right temporal pole	3
EL	M	53	Hebrew	Academic	no	R	Right fronto-temporal ischemic infarct	3
DR	F	63	Hebrew	High school	yes	R	Ischemic infarct in the right MCA territory	5
YA	M	54	Hebrew	High school	no	R	Recurrent Ischemic infarct in the right MCA territory	3
YI	M	45	Hebrew	High school	yes	R	Right parietal hemorrhage, and older bilateral basal ganglia infarcts	3
MO	M	63	Hebrew	High school	no	R	Ischemic infarct in the right MCA territory	3
EG	F	41	Hebrew	Academic	no	R		1

Table A2  
Participants in Experiment 2

Participant	Gender	Age	Spoken Language	Occupation	Hand-edness	Lesion site and Etiology	Months post onset
DS	M	60	Hebrew	Salesman	R	Right parieto-occipital infarct	2
RB	F	68	Hebrew	Cleaning	R	Infarct in basal ganglia and Right MCA territory	2
OK	M	48	Hebrew	Supermarket (butcher)	R	Right fronto parietal hemorrhage, with occipital lobe involvement	3
RU	F	61	Hebrew (came from Morocco – 2 years old)	Caregiver for the elderly	R	Right ICA occlusion. Right parietal infarct	3
YS	M	49	Hebrew	Policeman	R	Large right parieto-occipital hemorrhage	2
TM	F	57	Hebrew	Caregiver for the elderly	R	Ischemic infarct in the right MCA territory	2

## Appendix B. The aTOMia Battery

The aTOMia battery includes items testing eight aspects of TOM, two items per category. Each category presents a type of social scenario that requires inference of intentions, beliefs, feelings, and expectations of another person, and identification of gaps in knowledge and point of view (for additional information see Balaban, Friedmann, & Ziv, 2016).

**(a) First order false belief.** Two short scenarios were presented based on the original false belief task (Wimmer & Perner, 1983). Both scenarios described a change of location of an item. The change took place in the absence of the protagonist. We asked the participants two questions: a fact question: Where is the item? and a false belief question: Where will she/he look for that item? The appropriate answer was that the item was where it was moved to and that the protagonist will look for it in the place s/he left it.

**(b) Second-order false belief.** Two short scenarios were composed based on items used to test typically developing children (Sullivan, Zaitchik & Tager-Flusberg, 1994). Three characters were presented in the story, each of them having different knowledge about the others' state of knowledge regarding an object. After each scenario we asked two fact questions to assure understanding of the scenario, a yes/no TOM question about one of the character's understanding of others' understanding of the situation and a justification question. This task tested the ability to consider what one character knows about a different character's state of knowledge.

**(c) White Lies.** Two items tested the ability to understand a situation in which a protagonist acted as if she was surprised, although she was not truly surprised, in order to save face. Both items were based on the surprise stories presented by Happé (1994). Three questions were asked after each item: a fact question, a yes/no TOM question, and a justification question.

**(d) Knowledge gaps.** Two scenarios were presented. In each social situation two characters were presented and a misunderstanding between them was described (Happé et al., 1999). The misunderstanding resulted from a misperception of one character about the second character's intention. The participants were asked a fact question about the story and then they were asked to explain the action or comments of the two characters. The appropriate answer was one that explained the misunderstanding.

**(e) Teaching Initiation.** Two scenarios were composed based on the relation between TOM and the act of initiating teaching (Ziv & Frye, 2004). Initiation of teaching requires the teacher to appreciate that there is a knowledge gap between him/her and another person, and to have an intention to bridge that gap. Three questions were asked after each item: a fact question, a yes/no TOM question, and a justification question.

**(f) Faux pas.** Two short scenarios were presented to test the ability to appreciate that a person's rude remark resulted from her/his lack of knowledge about the situation (Baron-Cohen et al., 1999). To answer these items appropriately, the participants need to infer the naïve intention of the speaker and to consider the emotional reaction that is expected in the embarrassing social situation. After each scenario we asked a fact question, a yes/no TOM question, and a justification question.

**(g) Hiding Emotions.** Two items tested the understanding of situations in which the protagonist was either happy or felt embarrassment while interacting in a social situation that called for hiding the emotions. We asked a fact question about the true feelings of the protagonist. We then asked whether the protagonist will expose her feeling via her/his facial expression or will show a straight face and a justification question to explain why she would show/ hide her feelings.

**(h) Cartoon.** Two cartoons that did not include words were presented. The key to understanding them was grasping the figures' thoughts and intentions. First, we asked the participants to describe what they saw in each picture. Then we asked what they thought the cartoon was about and if they found it funny in any way (Happé et al., 1999). The cartoons were included in the battery to ensure the TOM assessment did not rely solely on items presented in a verbal form.

## Appendix C. Assessment of the participants' syntax

Our interpretation of the results of Experiment 1 relies on the idea that the grammatical properties of the multiple wh-questions forced the participants to exhaust, even in the absence of input from ToM. We directly assessed these participants' grammatical abilities through syntactic structures that are known to be clinical markers for grammatical impairments in agrammatic aphasia, and are directly related to the grammar of wh-questions. We used tasks that examined their comprehension of sentences derived by Wh-movement: relative clauses and wh-questions, their interpretation of pronouns, and their production of embedded sentences.

### 1. Comprehension of sentences derived by Wh-movement: sentence-picture matching

Relative clauses and wh-questions were chosen because these structures are the best clinical markers in Hebrew for syntactic impairments in various populations (Friedmann, Yachini, & Szterman, 2015), and because the current study tested Wh questions. Previous studies that used the task we use here to assess relative clause and Wh-question comprehension in Hebrew-speakers were sensitive to syntactic impairments of individuals with agrammatic aphasia (Friedmann, 2008; Friedmann & Gvion, 2003, 2012; Friedmann & Shapiro, 2003), children with developmental syntactic impairment (syntactic SLI, Levy & Friedmann, 2009; Friedmann & Novogrodsky, 2004, 2007, 2011), and orally-trained children with hearing impairment (Friedmann & Szterman, 2006, 2011; Friedmann, Szterman, & Haddad-Hanna, 2010), as well as pre-school typically developing children who have not yet mastered these structures (Friedmann, Belletti, & Rizzi, 2009; Belletti, Friedmann, Brunato & Rizzi, 2012; Friedmann, Rizzi, & Belletti, 2017). Sixteen participants participated in the Wh-movement sentence-picture matching tasks (11 of them with aTOMia): Ten of the participants were tested on relative clauses, and six on Wh questions (both tasks were taken from the BAFLA syntactic test battery, Friedmann, 1998).

**The relative-clause comprehension task** (BAFLA ZT, Friedmann, 1998) was comprised of 60 semantically reversible sentences:

20 simple SVO sentences (e.g., ‘The boy is washing the penguin’), 20 subject-relative clauses (e.g., ‘Show me the boy that is washing the penguin’), and 20 object-relative clauses (e.g., ‘Show me the penguin that the boy is washing’). The task included 20 picture-pairs, each was shown three times, each time with a different type of sentence. While the participant heard the sentence, she was shown a page with two pictures and was asked to point to the picture that matched the sentence. In one picture the roles matched the sentence; in the other the roles were reversed.

**The Wh questions comprehension task** (BAFLA ZST, [Friedmann, 1998](#)) was comprised of 20 semantically reversible Wh questions: 10 subject *which* questions (e.g., ‘Which grandpa is tickling the boy?’) and 10 object *which* questions (e.g., ‘Which girl is the grandma drawing’). The task included 20 pictures. While the participant heard the sentence, she was shown a page with the picture and was asked to point to the figure that matched the sentence. The picture included 3 figures, two of the same kind (e.g., two elephants), and another of a different kind (e.g., a lion). Each picture included figure A1 (a person or an animal of type A) performing an action on figure B (a person or an animal of a different type), which was performing the same action on figure A2 (another person or animal of type A) (e.g., an elephant spraying a lion spraying another elephant).

**Results:** The average score of the 16 participants was very high ( $M = 97\%$ ,  $SD = 3.0\%$ ). On object relative and object question structures, which are based on syntactic movement that changes the canonical word order, the participants with aTOMia performed very high ( $M = 92\%$ ,  $SD = 4.9\%$ ). For individual results see Table C.

## 2. Pronoun interpretation: sentence-picture matching

To examine the syntactic aspects of the participants' pronoun interpretation, and more specifically, their ability to apply the syntactic constraints on pronoun interpretation ([Chomsky's, 1981](#) binding-principles), we used another sentence-picture matching test. The interpretation of pronouns under binding relations is another domain that was found to be sensitive to syntactic impairments in aphasic patients ([Ruigendijk, Vasić, & Avrutin, 2006](#); [Vasić, 2006](#)), and is a late-developing ability in typically-developing Hebrew-speaking children ([Friedmann, Novogrodsky, & Balaban, 2010](#); [Ruigendijk, Friedmann, Novogrodsky, & Balaban, 2010](#)). Pronoun comprehension was chosen because, like question comprehension, pronouns have grammatical, as well as discursive, TOM-related aspects ([Balaban, Friedmann & Ariel, 2016](#)).

**The pronoun interpretation task** ([Ruigendijk et al., 2010](#)) included 72 sentences with pronoun or reflexive object: 24 coordinated sentences (‘The girl and the cow met and then the cow licked herself/her’, ‘cow’ in Hebrew is feminine), 24 sentences with sentential complements (‘The girl said that the cow is licking herself/her’), and 24 relative-clauses (‘The cow that met the girl licked herself/her’). Half of the sentences of each type included a pronoun and the other half included a reflexive. The two NPs in each sentence always had the same gender. Each sentence was presented with two pictures, one picture matched the sentence, the other presented an incorrect antecedent (the other NP in the sentence) as the object.

**Results:** Eleven of the participants participated in this task, 8 of them were diagnosed with aTOMia. The total average score was high ( $M = 97.9\%$ ,  $SD = 3.8\%$ ). All the participants performed 90% correct and above, seven of them answered 100% correct. The individual results are detailed in Table C.

## 3. Production of embedded sentences

The third syntactic measure was the percentage of sentences that included sentential complements of verbs in the answers provided during the administration of the aTOMia battery. This analysis is informative because the representation and processing of wh-questions require, aside from exhaustivity resolution, the construction of the syntactic tree up to its highest nodes (the left periphery of the sentence, [Rizzi, 1997](#)). So impaired construction of the high syntactic nodes may lead to wh-question impairment (see [Friedmann, 2001, 2002, 2005](#)). We analyzed the responses to the aTOMia battery and counted the number of justifications (out of 14) that included an overt (obligatory) complementizer in CP and a sentential complement (e.g., ‘She forgot that the chocolate was in the drawer’). This type of embedded structure is included in the justifications that people with no language impairments give to the TOM questions.

**Results:** The overall average percent of justifications that included well-formed embedded complements for the 19 participants was 56% ( $SD = 16\%$ ). The participants with aTOMia produced these sentences on 55% ( $SD = 17\%$ ) of the justifications and the participants without aTOMia produced them in 60% ( $SD = 14\%$ ) of the justifications. The difference in rate of spontaneously producing this complex structure correctly between the groups was not significant ( $U = 28.5$ ,  $p = .29$ ). The individual rates of production of these sentences is detailed in Table C.

In sum, we found that the participants with aTOMia we tested did not have a syntactic impairment. When TOM was not a prerequisite for a language task, they showed good language abilities.

Table C  
Individual results in the three syntactic tasks.

Participant	Wh-movement comprehension %correct	Pronoun interpretation: Binding principles A + B % correct	Use of embedded sentences in the justifications of aTOMia battery
DL	93%	100%	50%
NI	95%	100%	50%
JO	98%		75%
ME	93%	96%	50%
AS	100%	97%	50%
ZE		100%	63%
MH	90%		69%
MI		99%	56%
EZ	85% <sup>q</sup>		44%
HA			25%
YA	100% <sup>q</sup>		31%
MO	100% <sup>q</sup>	100%	44%
EL	95% <sup>q</sup>		81%
DR	98%	90%	75%
SA	100%	100%	50%
YI	100% <sup>q</sup>		44%
AN	100%	100%	75%
OF	100%	100%	75%
EG	100% <sup>q</sup>		56%

<sup>q</sup> Participants who participated in the Wh-question comprehension task.

## References

- Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, 7(4), 268–277.
- Apperly, I. A. (2011). *Mindreaders: The cognitive basis of "Theory of mind"*. Psychology Press.
- Apperly, I. A., Samson, D., 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(10), 1773–1784.
- 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.
- Balaban, N., Belletti, A., Friedmann, N., & Rizzi, L. (2016). Disentangling principle C: A contribution from individuals with brain damage. *Lingua*, 169, 1–20. <https://doi.org/10.1016/j.lingua.2015.09.004>.
- Balaban, N., Biran, M., & Sacher, Y. (2019). Theory of Mind and referring expressions after traumatic brain injury. *Aphasiology*. <https://doi.org/10.1080/02687038.2019.1637812>.
- Balaban, N., & Friedmann, N. (2014). Disentangling responses to wh-questions: TOM and syntactic abilities. In C. Contemori, & L. Dal Pozzo (Eds.). *Inquiries into linguistic theory and language acquisition. Papers offered to Adriana Belletti* (pp. 1–14). Siena: CISCL Press.
- Balaban, N., Friedmann, N., & Ariel, M. (2016). The effect of theory of mind impairment on language: Referring after right hemisphere damage. *Aphasiology*, 30(12), 1424–1460. <https://doi.org/10.1080/02687038.2015.1137274>.
- Balaban, N., Friedmann, N., & Ziv, M. (2016). Theory of mind impairment after right hemisphere damage. *Aphasiology*, 30(12), 1399–1423. <https://doi.org/10.1080/02687038.2015.1137275>.
- Balaban, N., Ziv, M., & Friedmann, N. (2010). *aTOMia battery: Assessments of ToM abilities. Language and brain lab*. Tel-Aviv University.
- 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.
- Bartels-Tobin, L. R., & Hinckley, J. J. (2005). Cognition and discourse production in right hemisphere disorder. *Journal of Neurolinguistics*, 18, 461–477.
- Beck, S., & Rullmann, H. (1999). A flexible approach to exhaustivity in questions. *Natural Language Semantics*, 7, 249–298.
- Belletti, A., Friedmann, N., Brunato, D., & Rizzi, L. (2012). Does gender make a difference? Comparing the effect of gender on children's comprehension of relative clauses in Hebrew and Italian. *Lingua*, 122(10), 1053–1069. <https://doi.org/10.1016/j.lingua.2012.02.007>.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57, 289–300.
- Blake, M. L. (2006). Clinical relevance of discourse characteristics after right hemisphere brain damage. *American Journal of Speech-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-Language Pathology*, 16, 331–342.
- Bošković, Z. (1998). *Wh-phrases and wh-movement in slavic. Position paper presented at the comparative slavic morphosyntax workshop*. Bloomington, Indiana <http://citeseerx.ist.psu.edu/viewdoc/download?Doi:10.1.1.72.9721>.
- Brady, M., Armstrong, L., & Mackenzie, C. (2006). An examination over time of language and discourse production abilities following right hemisphere brain damage. *Journal of Neurolinguistics*, 19(4), 291–310.
- Brownell, H., & Stringfellow, A. (1999). Making requests: Illustrations of how RBD can affect discourse production. *Brain and Language*, 68, 442–465.
- Chantraine, Y., Joannette, Y., & Ska, B. (1998). Conversational abilities in patients with right hemisphere damage. *Journal of Neurolinguistics*, 11(1), 21–32.
- Cheang, H. S., & Pell, M. D. (2006). A study of humour and communicative intention following right hemisphere stroke. *Clinical Linguistics and Phonetics*, 20(6), 447–462.
- Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- Danckert, J., Stöttinger, E., Quehl, N., & Anderson, B. (2012). Right hemisphere brain damage impairs strategy updating. *Cerebral Cortex*, 22(12), 2745–2760.
- Dayal, V. (1996). *Locality in wh-quantification: Questions and relative clauses in Hindi*. Dordrecht: Kluwer.
- Dayal, V. (2017). Multiple wh questions. In M. Everaert, & H. van Riemsdijk (Eds.). *The blackwell companion to syntax 3* (pp. 275–327). (2<sup>nd</sup> ed.). Malden: Wiley-Blackwell.
- Fekete, I., Schulz, P., & Ruigendijk, E. (2018). Exhaustivity in single bare wh-questions: A differential-analysis of exhaustivity. *Glossa: A Journal of General Linguistics*, 3(96), 1–32. <https://doi.org/10.5334/gjgl.549>.

- Ferstl, E. C., Walther, K., Guthke, T., & von Cramon, D. Y. (2005). Assessment of story comprehension deficits after brain damage. *Journal of Clinical and Experimental Neuropsychology*, 27(3), 367–384.
- Fox, D. (2012). Multiple wh-questions: Uniqueness, pair-list and second order questions. *Class notes for MIT seminars*.
- Friedmann, N. (1998). *BAFLA: Friedmann battery for agrammatism*. Tel Aviv: Tel Aviv University.
- Friedmann, N. (2001). Agrammatism and the psychological reality of the syntactic tree. *Journal of Psycholinguistic Research*, 30(1), 71–90. <https://doi.org/10.1023/A:1005256224207>.
- Friedmann, N. (2002). Question production in agrammatism: The tree pruning hypothesis. *Brain and Language*, 80, 160–187. <https://doi.org/10.1006/brln.2001.2587>.
- Friedmann, N. (2005). Degrees of severity and recovery in agrammatism: Climbing up the syntactic tree. *Aphasiology*, 19, 1037–1051. <https://doi.org/10.1080/02687030544000236>.
- Friedmann, N. (2006). Speech production in Broca's agrammatic aphasia: Syntactic tree pruning. In Y. Grodzinsky, & K. Amunts (Eds.). *Broca's region* (pp. 63–82). New York, NY: Oxford University Press.
- Friedmann, N. (2008). Traceless relatives: Agrammatic comprehension of relative clauses with resumptive pronouns. *Journal of Neurolinguistics*, 21, 138–149. <https://doi.org/10.1016/j.jneuroling.2006.10.005>.
- 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 limitation in aphasia: A dissociation between semantic-syntactic and phonological reactivation. *Brain and Language*, 86, 23–39. [https://doi.org/10.1016/S0093-934X\(02\)00530-8](https://doi.org/10.1016/S0093-934X(02)00530-8).
- Friedmann, N., & Gvion, A. (2012). Intervention and locality in agrammatic aphasia. In V. Bianchi, & C. Chesì (Eds.). *Enjoy linguistics! Papers offered to luigi Rizzi on the occasion of his 60th birthday* Siena, Italy: CISCL Press 9788890794308.
- Friedmann, N., & Novogrodsky, R. (2011). Which questions are most difficult to understand? The comprehension of wh questions in three subtypes of SLI. *Lingua*, 121, 367–382. <https://doi.org/10.1016/j.lingua.2010.10.004>.
- Friedmann, N., Rizzi, L., & Belletti, A. (2017). No case for Case in locality: Case does not help interpretation when intervention blocks A-bar chains. *Glossa: A Journal of General Linguistics*, 2(1), 1–18. <https://doi.org/10.5334/gjgl.165>.
- 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. [https://doi.org/10.1044/1092-4388\(2003\)023](https://doi.org/10.1044/1092-4388(2003)023).
- Friedmann, N., & Sztzman, R. (2011). The comprehension and production of wh questions in children with hearing impairment. *Journal of Deaf Studies and Deaf Education*, 16(2), 212–235.
- Friedmann, N., Sztzman, R., & Haddad-Hanna, M. (2010). The comprehension of relative clauses and Wh questions in Hebrew and Palestinian Arabic hearing impairment. In J. Costa (Ed.). *Language acquisition and development: generative approaches to language acquisition* (pp. 157–169). Newcastle, UK: Cambridge Scholars Press.
- Friedmann, N., Tzailer-Gross, L., & Gvion, A. (2011). The effect of syntax on reading in neglect dyslexia. *Neuropsychologia*, 49(10), 2803–2816.
- Friedmann, N., Yachini, M., & Sztzman, R. (2015). Relatively easy relatives: Children with syntactic SLI avoid intervention. In E. Di Domenico, C. Hamann, & S. Matteini (Eds.). *Structures, strategies and beyond. Studies in honour of Adriana Belletti* (pp. 303–320). Amsterdam, The Netherlands: John Benjamins, Linguistik Aktuell series. <https://doi.org/10.1075/la.223.14fri>.
- Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London B Biological Sciences*, 358(1431), 459–473.
- Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron*, 50(4), 531–534.
- Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, 17(3), 311–347.
- Grodzinsky, Y. (1984). The syntactic characterization of agrammatism. *Cognition*, 16, 99–120.
- 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. <https://doi.org/10.1006/brln.1999.2050>.
- Groenendijk, J., & Stokhof, M. (1984). On the semantics of questions and the pragmatics of answers. In F. Landman, & F. Veltman (Eds.). *Varieties of formal semantics*. Dordrecht: Foris.
- Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69, 274–307. <https://doi.org/10.2307/416535>.
- Hagstrom, P. A. (1998). *Decomposing questions* Ph.D. dissertation. MIT.
- Halligan, P. W., Cockburn, J., & Wilson, B. (1991). The behavioural assessment of visual neglect. *Neuropsychological Rehabilitation*, 1, 5–32.
- Hamblin, C. L. (1973). Questions in montague English. *Foundations of Language*, 10, 41–53.
- 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., Brownell, H., & Winner, E. (1999). Acquired 'theory of mind' impairments following stroke. *Cognition*, 70, 211–240.
- 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.
- Karttunen, L. (1977). Syntax and semantics of questions. *Linguistics and Philosophy*, 1, 3–44.
- Kriifka, M. (2001). For a structured meaning account of questions and answers. In C. Fery, & W. Sternefeld (Eds.). *Audiatur vox sapientia. A festschrift for armin von Stechow* (pp. 287–319). Berlin: Akademie Verlag.
- Levy, H., & Friedmann, N. (2009). Treatment of syntactic movement in syntactic SLI: A case study. *First Language*, 29, 15–50. <https://doi.org/10.1177/0142723708097815>.
- 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(3), 451–466.
- Myers, P. S. (2001). Towards a definition of RHD syndrome. *Aphasiology*, 15, 913–918.
- Nelken, R., & Shan, C.-C. (2004). *A logic of interrogation should be internalized in a modal logic for knowledge. Paper presented at the Conference on Semantics and Linguistic Theory (SALT 14)*. Chicago [www.semanticsarchive.net/Archive/mEwOTFIO/nelken-shan-salt14.pdf](http://www.semanticsarchive.net/Archive/mEwOTFIO/nelken-shan-salt14.pdf).
- Onea, E., & Zimmermann, (in press). Questions in discourse: An overview. In K. von Heusinger, E. Onea, & M. Zimmermann (Eds.), *Questions in discourse*. CRISPI series. Leiden: BRILL.
- Pesetsky, D. (1987). Wh-in-situ: Movement and unselective binding. In E. Reuland, & A. ter Meulen (Eds.). *The representation of (in)definiteness* (pp. 98–129). MIT Press Chapter 5.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169–226.
- Reis, M. (1992). The category of invariant alles in wh-clauses: On syntactic categories vs. quantifying particles in German. In R. Tracy (Ed.). *Who climbs the grammar tree* (pp. 465–492). Tübingen: Niemeyer Verlag.
- Reznick, J., & Friedmann, N. (2015). Evidence from neglect dyslexia for morphological decomposition at the early stages of orthographic-visual analysis. *Frontiers in Human Neuroscience*, 9(497), 1–19. <https://doi.org/10.3389/fnhum.2015.00497>.
- Rizzi, L. (1997). The fine structure of the left periphery. In L. Haegeman (Ed.). *Elements of grammar. handbook of generative syntax* (pp. 281–337). Dordrecht: Kluwer.
- Roeper, T. (2004). Diagnosing language variations: Underlying principles for syntactic assessment. *Seminars in Speech and Language*, 25, 41–55.
- Roeper, T., Schulz, P., Pearson, B., & Reckling, I. (2007). From singleton to exhaustive: The acquisition of wh-. In M. Becker, & A. McKenzie (Vol. Eds.), *Proceedings of the 3rd conference on the semantics of underrepresented languages in the americas (SULA): University of Massachusetts occasional papers in linguistics: Vol. 33*, (pp. 87–102). (Amherst, MA).
- van Rooij, R. (2003). Questioning to resolve decision problems. *Linguistics and Philosophy*, 26(6), 727–763. <https://doi.org/10.1023/B:LING.000004548.98658.8f>.

- van Rooij, R. (2004). Questions and relevance. In R. Bernardi, & M. Moortgat (Eds.). *Questions and answers. Proceedings of 2nd CoLogNET-ElsNET symposium* (pp. 96–107).
- Ruigendijk, E., Friedmann, N., Novogrodsky, R., & Balaban, N. (2010). Symmetry in comprehension and production of pronouns: A comparison of German and Hebrew. *Lingua*, 120(8), 1991–2005. <https://doi.org/10.1016/j.lingua.2010.02.009>.
- Ruigendijk, E., Vasić, N., & Avrutin, S. (2006). Reference assignment: Using language breakdown to choose between theoretical approaches. *Brain and Language*, 96(3), 302–317.
- Schulz, P. (2010). Some notes on semantics and SLI. In A. Castro, (Ed.). *language acquisition and development. Proceedings of GALA 2009* (pp. 391–406). Cambridge: Cambridge Scholars Press.
- Schulz, P. (2015). Exhaustivity. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.). *Methods for assessing multilingual children: Disentangling bilingualism from Language Impairment* (pp. 76–94). (Book Series: Multilingual Matters).
- Schulz, P., Gavarró, A., Prévost, P., Friedmann, N., et al. (2018). *Children's understanding of exhaustivity in wh-questions: Evidence for universal acquisition principles across 15 languages*. Manuscript: Goethe University Frankfurt.
- Schulz, P., & Roeper, T. (2011). Acquisition of exhaustivity in wh-questions: A semantic dimension of SLI? *Lingua*, 121, 383–407.
- Sherratt, S., & Bryan, K. (2012). Discourse production after right brain damage: Gaining a comprehensive picture using a multi-level processing model. *Journal of Neurolinguistics*, 25(4), 213–239.
- Sperber, D., & Wilson, D. (2004). Relevance theory. In L. Horn, & G. Ward (Vol. Eds.), *The handbook of pragmatics: Vol. 26*, (pp. 607–632). Oxford: Blackwell Publishers.
- Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Developmental Psychology*, 30(3), 395–402. <https://doi.org/10.1037/0012-1649.30.3.395>.
- Surian, L., & Siegal, M. (2001). Sources of performance on theory of mind tasks in right hemisphere damaged patients. *Brain and Language*, 78, 224–232.
- Vasić, N. (2006). *Pronoun comprehension in agrammatic aphasia: The structure and use of linguistic knowledge. Vol. 140*. LOT.
- Villiers, J. de, Ning, C., Liu, X. L., Zhang, Y. W., & Jiang, F. (2018). The acquisition of exhaustive pairing in multiple Wh-Questions in Mandarin. *Journal of Psycholinguistic Research*, 47(6), 1369–1389. <https://doi.org/10.1007/s10936-018-9588-1>.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103–128.
- Zimmermann, M. (2007). Quantifying question particles in German: Syntactic effects on interpretation. In E. Puig-Waldmueller (Ed.). *Proceedings of Sinn und Bedeutung (SuB 11)* (pp. 627–641). Barcelona.
- Zurif, E., & Caramazza, A. (1976). Psycholinguistic structures in aphasia: Studies in syntax and semantics. In H. Whitaker, & H. A. Whitaker (Vol. Eds.), *Studies in neurolinguistics: Vol. I*, (pp. 260–292). New York: Academic Press.
- Ziv, M., & Frye, D. (2004). Children's understanding of teaching: The role of knowledge and belief. *Cognitive development*, 19(4), 457–477.