

Language comprehension, inference, and alternatives

Dimitrios Skordos & David Barner

University of California, San Diego

CONTACT:

Dimitrios Skordos

Department of Psychology

University of California, San Diego

La Jolla, CA 92093

Email: dskordos@ucsd.edu

Humans have an astonishing ability to comprehend language. Comprehension begins when we perceive a continuous perceptual signal produced by another speaker, assign this signal a phonological form, and parse it into words. These words are arranged into a hierarchical syntactic structure and logical form, which together permit the listener to recover the compositional semantics of the expression. The speed and efficiency with which we achieve these steps, with little apparent effort, belies the richness and complexity of the underlying system. However, as complex as these steps may be, our understanding of them far outstrips our current comprehension of the final step in language comprehension, in which we infer the speaker's intended meaning from the utterance and the context in which it is used. In making this inference, we complete a direct bridge between the speaker's mental states and our own.

Each of these different steps of language comprehension involves a series of informed guesses based on sparse data. The sound signal is generally impoverished, and surface forms often admit multiple syntactic analyses. Similarly, the intentions of speakers are not always transparently reflected by the words that they choose. Instead, the literal meaning of an utterance is often consistent with a large space of possible intentional states, requiring the listener to make rapid abductive guesses about the intentions that best explain the speaker's choice of words in a particular context. Consider, for example, a situation in which you walk into the kitchen area of your office building and overhear part of a discussion consisting of utterance (1) below.

(1) "I ate some of the cookies."

This utterance might have several different intended meanings all derived inferentially from the same underlying semantic content, depending on what the hearer thinks the speaker intends to

communicate. For instance, had you walked in a few moments earlier you might have heard the other communicative partner ask a question that prompted the utterance. Depending on the question asked, the assertion of eating some of the cookies could take on multiple different meanings as in (2) – (4):

- (2) a) “Did you eat all of the cookies?”
 - b) “I ate some of the cookies.” (I did not eat all of the cookies).
- (3) a) “Did you try the cookies and cake I made for you?”
 - b) “I ate some of the cookies.” (I did not try the cake).
- (4) a) “Did you and Jane try the cookies?”
 - b) “I ate some of the cookies.” (Jane did not eat any cookies).

How do humans make these inferences so readily? How do they restrict the hypothesis space of possible meanings in order to converge on the meanings that speakers have in mind? In the present chapter, we discuss this problem, and the fact that the specific inference derived by a listener in a context like this depends critically on the alternative utterances that they believe to be available to the speaker. These alternatives, in turn, depend on the “question under discussion” – e.g., which of the utterances in (2) through (4) best represent the question that the speaker intended to address with their utterance (Roberts, 1996). For instance, given the question in (3a), the alternatives relevant to the conversation include the statement in (3b), but not the alternatives in (2b) or (4b). Likewise, given the question in (4a), the alternative statement in (4b) becomes a relevant alternative, whereas those in (2b) and (3b) do not. Still, even given this

framework, the question remains of how these alternatives are generated by the listener, and how listeners identify which alternatives are relevant to which questions.

The inferences described in examples (2) through (4) constitute different examples of conversational implicature. Among these, the example in (2b) is a classic example of what linguistics call “scalar implicature”, which typically arises when an utterance contains a scalar term like *some* that is enriched via the negation of a stronger, alternative, utterance that contains a scale mate like *all* (e.g., Grice, 1975; Hirschberg, 1985; Horn, 1972). This stronger statement can be said to logically entail the original utterance by virtue of being ‘informationally stronger’ – i.e., because the truth conditions of the original utterance are necessarily satisfied, whenever the truth conditions of the stronger statement are satisfied. For instance, if it is true that “I ate all of the cookies” it is necessarily true that “I ate some of the cookies”. Assuming Grice’s maxim of Quantity (Grice, 1975), which states that communicative partners are expected to be as informative as reasonably expected, the use of the weaker alternative containing *some* is taken to implicate that the speaker does not believe the stronger utterance containing *all* to be true, resulting in the inference “*some* but not *all*.”

In this chapter, we first review studies of communicative inference in language comprehension and then focus on the development of these inferences in children, whose non-adult-like behaviors afford a window into the interaction between semantics and pragmatics. In particular, we will explore evidence from children’s difficulties with scalar implicature and Piagetian class inclusion errors, and argue that although children’s behavior in these cases has been accounted for independently in the past, it is possible to provide a unified account of both sets of phenomena. We will argue that children’s logical reasoning is limited by their ability to spontaneously access and reason about alternative linguistic forms and states of affairs.

The study of inference in language comprehension

Inference in language comprehension has long been studied in adults, beginning with Grice (1969, 1975) and his proposal that communication is a cooperative endeavor between partners, mediated by common sense rules (maxims) that govern conversation. Specifically, Grice proposed that speakers should be truthful, clear, relevant, as well as informative, by providing as much information as needed in an utterance, but no more. Following Grice, subsequent approaches have reinterpreted his intuitions, resulting in a variety of distinct theoretical frameworks. First, one class of theories has attempted to reduce conversational implicature to an optimization process between the expected cognitive ‘gain’ of deriving an inference and the cognitive ‘cost’ incurred in the process of doing so (Sperber & Wilson, 1986/1995; Carston, 1995; Noveck & Sperber, 2007). This approach is part of a ‘contextualist’ theory of inference that generally views inference as effortful, distinct from grammatical processes, and dependent on context. A second class of theories, dubbed Neo-Gricean, is best exemplified by Gazdar (1979), Horn (1972, 2006) and Levinson (1983, 2000), and was later further developed by Sauerland (2004), van Rooij and Schulz (2004) and others. This approach aims, by providing a formal characterization of Grice, to simplify the Maxims in order to preclude inconsistencies or clashes between them, but also to explain the near automaticity of certain inferences like scalar implicatures. To do so, they assume the lexicalization of inferential components like the scalar alternatives <*some, all*> (Horn, 1972), and, in some cases, the generation of scalar inferences by default (Levinson, 2000). Finally, a third class of theory, known as ‘grammatical’ accounts of scalar inference, treats some forms of communicative inference, including scalar implicature, as purely grammatical in nature. On these accounts,

implicatures are triggered by a silent logical operator, akin to the English word *only*, which enforces exhaustification grammatically, rather than pragmatically (Chierchia 2004; Fox 2007; Spector, 2007; Chierchia, Fox & Spector, 2009).

In efforts to differentiate these accounts empirically, studies of mature, adult speakers have probed the time-course of implicature computation (Noveck & Posada, 2003; Bott & Noveck, 2004; Breheny, Katsos & Williams, 2006; De Neys & Schaeken, 2007; Chevallier et al., 2008; Bonnefon, Feeney & Villejoubert, 2009; Huang & Snedeker, 2009, 2011; Grodner, Klein, Carbarry & Tanenhaus, 2010; Bott, Bailey & Grodner, 2012; Bergen & Grodner, 2012; Tomlinson, Bailey & Bott, 2013; Breheny, Ferguson & Katsos, 2013; Chemla & Bott, 2014). A preponderance of evidence supports the idea that deriving these implicatures requires time (and possibly effort) beyond what is involved in computing the core, semantic, meaning of an utterance. However, there is little agreement among researchers about the exact timeline and processes involved, and how these facts adjudicate between different theoretical views. More recent accounts (Degen & Tanenhaus, 2015, 2016, in press), as well as neurolinguistic evidence (Poltzer-Ahles & Gwilliams, 2015; Spsychalska, Kontinen & Werning, 2016), call for more nuanced approaches, where lexical, syntactic, and contextual factors must be integrated into a model of implicature (see Degen & Tanenhaus, this volume, for a review and discussion). Still, almost all theories can currently accommodate this range of facts, with minor modifications.

Critically, scalar implicature is not the only test case for exploring how individuals represent and reason about linguistic alternatives in the service of pragmatic inference. For instance, formal semanticists have constructed theories of how language users comprehend sentential focus, which involves reasoning over unspoken alternatives, triggered in context by a focus element (Rooth, 1992, 2016; cf. Kim, this volume). For example, if one utters “*Mary came*

to the party”, sentential focus on “Mary” leads the hearer to consider alternative utterances that represent propositions about implicated individuals who did not actually come (e.g., John). In more recent work (Aloni, 2003; Simons, 2005; Alonso-Ovalle, 2009; cf. Rooth, 2016 for discussion), this type of analysis is extended to a much broader set of phenomena such as disjunction and conditionals, and prior work has made similar arguments for weak negative polarity items like *any* (Krifka, 1995). Finally, Chierchia (2013) extends an analysis based on reasoning with alternatives to all polarity-related phenomena, including strong negative polarity items and free choice inferences. Formal accounts of focus and topic and how these grammatical elements trigger the consideration of alternatives also share an affinity with the notion of the “Question Under Discussion” or QUD (Roberts, 1996, 2004). QUD accounts of pragmatics (Stalnaker, 1979; Roberts, 1996, 2004), describe discourse as proceeding on the basis of conversational goals, primary among which is to discover the ‘current state of affairs’ with regards to what is being discussed, i.e., the topic of the conversation (cf. Hulsey, Hacquard, Fox & Gualmini, 2004; also Conroy, Lidz & Musolino, 2009, for discussion). To discover this information, communicative partners posit and answer explicit and implicit questions and sub-questions relevant to the hypothesized topic. An utterance is considered ‘relevant’ to the QUD, if it provides an answer (full or partial) to it (Roberts, 1996, 2004; van Rooij & Schulz, 2004). It is not difficult to see how a QUD account of inference also requires reasoning with alternative propositions. For instance, if the question under discussion is “What did you eat?”, the utterance “I ate some of the cookies” is likely to trigger consideration of alternatives that involve other foods, such as ‘cake’, ‘sandwich’, etc. If, however, the question under discussion is “Who ate the cookies?” the same utterance would likely trigger consideration of alternatives that involve people, such as ‘John’, ‘Mary’, ‘Jane’ and so on.

Although these accounts can be couched in the language of a Gricean theory of communication, it is important to note that, as Russell (2012) points out, virtually all formal accounts of reasoning with alternatives, including theories linked to the QUD (Gazdar, 1979; Groenendijk & Stokhof, 1984; Sauerland, 2004; van Rooij & Schulz, 2004), as well as those related to focus and ‘alternative semantics’ (Rooth, 1992, 2016) do not actually require Gricean reasoning. Rather, the formal objects posited within these accounts – e.g., exhaustivity operators – can also be thought of as grammatical representations that are independent of any epistemic reasoning that speakers might engage in during conversation. Still, some recent formal approaches motivated by computational linguistics and Bayesian inference incorporate epistemic considerations. These belong to the recently revived area of ‘probabilistic pragmatics’, in which hearers make sense of the speaker’s intended meaning by making probabilistic inferences that take into account alternative utterances the speaker could have used instead (Franke, 2009, 2014; Frank, Goodman & Tenenbaum, 2009; Bergen, Goodman & Levy, 2012; Jäger, 2011; Degen, Franke & Jäger, 2013; Goodman & Stuhlmüller, 2013, Bergen & Goodman, 2015; Bergen, Levy & Goodman, 2016). According to such approaches, language comprehension can be generally seen as a joint inference over (a) a set of linguistic expressions that could have been used, for instance “some of the cookies” vs. “all of the cookies” and (b) a set of possible meanings or intentional states that the expressions might convey, for example, “at least some of the cookies” vs. “some but not all of the cookies” (Frank & Goodman, 2012).

Investigations into how adults make conversational inferences have undoubtedly yielded valuable insights into language processing and comprehension. That said, existing theories are often difficult to differentiate given adult behaviors, such that current debates focus on an increasingly narrow set of phenomena, sometimes disputing millisecond differences in

processing of inferences as evidence for one account or another (Breheny, Katsos & Williams, 2006; Huang & Snedeker, 2009; Grodner, Klein, Carbary & Tanenhaus, 2010; Bott, Bailey & Grodner, 2012, among others). For this reason, researchers have also turned to populations other than typically developed adults, including individuals with linguistic and pragmatic deficits (Noveck, Guelminger, Georgieff & Labruyere, 2007; Loukousa & Moilanen, 2009; Chevallier, Wilson, Happé & Noveck, 2010, Hochstein, Bale & Barner, 2017). One such population, which we explore here, is typically developing children, who often exhibit semantic competence, yet nevertheless fail to make certain forms of pragmatic inference, under a restricted range of circumstances, thus providing unique insight into the interface between semantics and pragmatics.

Developmental evidence: successes and ‘failures’

Despite appearing generally adult-like in their phonology, syntax, and semantics (Gleitman & Newport, 1995), older preschoolers and kindergarteners often have difficulties with inferences involving quantifiers (e.g., *some*, *every*) and logical operators like *or* and the conditional *if... then*. Inhelder and Piaget (1964) first noticed such non-adult-like behaviors in simple reasoning tasks involving universal quantification, which they dubbed “class inclusion”. In a classic example, children were presented with a set of shapes consisting of red squares, as well as red and green circles. When they were asked, “Are all of the squares red?” they often responded “*No*”, which led Inhelder and Piaget to conclude that children’s logical representations were not fully developed. Interestingly enough, some years later, researchers began noticing that children, if anything, responded *more* logically than adults in tasks that involved quantification and which typically elicited scalar inferences like those described earlier.

Smith (1980) for instance, pointed out that preschoolers tended to interpret *some* as “some and possibly all”, consistent with the logical interpretation of the existential quantifier. Paris (1973), along with Braine and Rumain (1981), showed that even kindergarteners preferred an inclusive/logical interpretation of disjunction as “*p* or *q* or both” as opposed to the exclusive interpretation “either *p* or *q*, but not both” preferred by adults. These early findings were confirmed and extended to modals such as *might* and *must* by Noveck (2001), who also highlighted their relevance to the study of inference in language comprehension.

Children’s pattern of difficulties with these inferences is puzzling: Whereas they appear to exhibit logical failures when interpreting universal quantifiers, they appear to be *hyper* logical in other cases, like disjunction, or when interpreting existentially quantified statements. In these latter cases, children’s behavior has been interpreted as a pragmatic failure precipitated by difficulties in Gricean epistemic reasoning. This approach, however, seems at odds with a wealth of recent evidence that much younger children have relatively sophisticated abilities to represent the goals and beliefs of other people. For instance, by 10 months of age, infants engage in joint attention to monitor the interests of speakers as they speak (Corkum & Moore, 1998) and by 18 months can employ joint attention systematically when acquiring and interpreting language (Tomasello, 1995). Both 1- and 2-year-olds show evidence of grasping communicative intentions in non-linguistic tasks (Behne, Carpenter & Tomasello, 2005; cf. Tomasello, Carpenter, Call, Behne & Moll, 2005), and 12-month-olds seem able to consider the epistemic state of others when trying to help them find items in a non-linguistic game (Liszkowski, Carpenter & Tomasello, 2008). More generally, infants have been shown to understand the goals and beliefs of others in a wide variety of tasks (Meltzoff, 1995; Woodward & Markman, 1998; Gergely,

Bekkering & Király, 2002; Southgate, Senju & Csibra, 2007; Onishi & Baillargeon, 2005; Buttelmann, Carpenter & Tomasello, 2009).

Currently, there is no unified account of this pattern of behaviors, wherein children appear to be impaired or sophisticated logically, across different cases, or pragmatically, in other cases. Instead, for each set of phenomena, a different, isolated, account has been proposed, resulting in a disjointed portrait of inferential development in children. In the section below, we propose that a unified account may in fact be possible, given recent developments in the study of pragmatic development, and in particular the case study of scalar implicature. We explore one proposal for why children struggle with implicature – namely the “access to alternatives” account – and propose that this idea may also be profitably extended to other cases of logical reasoning, to explain why children struggle with universal quantification and the interpretation of conditional statements, while nevertheless having relatively strong epistemic reasoning abilities.

Scalar Implicature

As mentioned earlier, children struggle to compute scalar implicatures with the quantifiers *some* vs. *all* (Smith, 1980; Noveck, 2001; Papafragou & Musolino, 2003; Feeney, Scafton, Duckworth & Handley, 2004; Guasti, Chierchia, Crain, Foppolo, Gualmini & Meroni, 2005; Katsos & Bishop, 2011; Foppolo, Guasti & Chierchia, 2012; Skordos & Papafragou, 2016). Moreover, their difficulties extend to other cases of scalar inference, including the singular-plural distinction (Barner, Chow & Yang, 2009), disjunction (Braine & Romain, 1981; Chierchia et al., 2001; Paris, 1973; Hochstein, Bale, Fox & Barner, 2014; Tieu, Romoli, Zhou & Crain, 2016; Singh et al., 2016), and modals like *might* and *must* (Noveck, 2001; Ozturk & Papafragou, 2015). Unlike Piaget, most researchers have concluded that children have largely

adult-like mental representations of the relevant lexical items and the necessary logical and syntactic competence, but lack other knowledge or abilities required to respond in an adult-like manner (although see Horowitz, Schneider & Frank, in press). Most early accounts, for example, ascribed children's difficulties to domain-general processing limits (Chierchia et al., 2001; Reinhart, 2004; Pouscoulous et al., 2007), while other more recent studies attributed children's apparent 'failures' to their reluctance to reject infelicitous statements that were otherwise logically true (Katsos & Bishop, 2011).

A more recent account argues that children's problems with scalar inference stem from difficulties accessing relevant linguistic alternatives that are required to derive implicatures (Barner & Bachrach, 2010; Barner et al., 2011; Chierchia et al., 2001; Tieu et al., 2016; Hochstein et al., 2014; Skordos & Papafragou, 2016). As we saw previously, to derive the pragmatically enriched inference, "I ate some, but not all of the cookies", from the utterance, "I ate some of the cookies", one needs to consider and negate the relevant alternative, "I ate all of the cookies". If children cannot access the relevant alternative during the computation, the inference cannot be made and children are left with the basic semantic content of the original proposition. In a test of this "access-to-alternatives" account, a study by Barner, Brooks and Bale (2011) showed 4-year-olds simple pictures and asked them questions that required scalar reasoning. In one condition, children were presented with a scene – e.g., three animals including a dog, a cat, and a cow – and all three were sleeping. On critical trials, they were then asked either whether the cat and the dog were sleeping, or, alternatively, whether *only* the cat and the dog were sleeping. In this context, they answered *Yes* to the first question 93% of the time, since it was indeed the case that the cat and the dog were sleeping. Critically, they answered *Yes* only 16% of the time in response to the second question, which included the word *only*, since it was

false that *only* the cat and the dog were sleeping. In stark contrast, children failed to differentiate trials that included *only* from those that did not when questions involved the scalar quantifier *some*. Here, on critical trials, where all animals were sleeping, children responded *Yes* equally often when asked whether some of the animals were sleeping and when asked whether *only* some of the animals were sleeping (in each case around 66% of the time). Barner et al. reasoned that the best explanation of this pattern was that children failed to exhaustify *only some* utterances because they required access to stronger alternatives including the quantifier *all*, which was not contextually available. Crucially, they succeeded with the *only a cat and a dog* cases – which they called “*ad hoc*” scales, because the relevant stronger alternative – i.e., *the cat, the dog, and the cow are sleeping* – could be constructed on the fly from the contextually available images. In this account it is access to alternatives, rather than a blanket pragmatic deficit or processing limitations, that explains children’s classic failures with *some/all* implicatures.

In support of this, subsequent studies have reported that children successfully compute scalar implicatures if the alternatives are provided contextually, as is the case with so-called *ad hoc* implicatures (Hirschberg, 1985). First, in a study by Stiller, Goodman and Frank (2015), 3- and 4-year-olds were shown simple scenes consisting of three possible referents in the form of smiley faces. One of the smileys was bare, a second had glasses, and the third had both glasses and a hat. When told that the experimenter’s “friend” had glasses and asked to guess who that friend was, even 3-year-olds were successful in picking out the smiley that had glasses but no hat, arguably because they were able to derive a scalar implicature from the contextual scale <glasses, glasses & hat>. In another study, Barner, Hochstein, Rubenson and Bale (in press), presented 4- and 5-year-old children with two puppets, one who was described as silly, and who often said strange things, and the other who was described as smart. They then watched as a bear

approached a pair of objects – e.g., an orange and a banana – and either took one, or two of them. At this point, the experimenter told the child that one of the two puppets – the silly or the smart one – had described what had happened, and that they had to guess who it was that had provided the description. On critical trials, when both objects were taken, and the children were told, “Somebody said that the bear took an orange”, and were asked to pick who had said it, they correctly chose the silly puppet, and did so to the same degree that they chose the appropriate puppet on control trials. Critically, in a different study using an identical design, but testing disjunction, children did not succeed. When shown a bear who took both an orange and a banana and told that “Somebody said the bear took an orange or a banana” children systematically chose the smart puppet, despite choosing the appropriate puppet on control trials (Hochstein, Bale, Fox & Barner, 2015). Together, these results suggest that children can compute scalar implicatures when alternatives can be recovered contextually, as in the case of *ad hoc* scales, but not in cases like disjunction, which requires access to a stronger alternative involving conjunction that are not contextually available (e.g., The bear took the orange and the banana).

In the Stiller et al. (2015) study, as well as in *ad hoc* conditions of Barner et al. (2011) and Barner et al. (in press), young children were successful in deriving scalar inferences when the necessary alternatives were made available to them. How exactly alternatives help children however, was left open by these studies. One possibility is that children struggle with implicature because they lack strong associations between members of scales – e.g., that in adults, scalar items like *some* and *all* become strongly associated over time, such that hearing *some* automatically triggers access to *all*. Also, relatively weak knowledge of scales might interact with processing difficulties – children with strong working memory capacities might be able to compensate for weak knowledge of scales in order to access scale mates, while children with

weaker processing capacities might be most impaired by weak knowledge of scales. However, another possibility left open by these studies is that contextually available alternatives help children by hinting at the intended question under discussion: Children may fail to compute implicatures not because they lack the ability to access alternatives in general, but because they don't know which alternatives are contextually relevant – i.e., they struggle to identify the QUD. Recall that if a speaker utters, “I ate some of the cookies”, they might intend to convey, “I ate some, but not all of the cookies”, or, “I ate some of the cookies, but no cake”, or even, “I ate some of the cookies, but Jane didn't eat any”. Crucially, which of the logically possible inferences matches the speaker-intended meaning, is restricted by the alternatives that one considers (<some cookies vs. all cookies>, <cookies vs. cookies & cake>, <speaker vs. speaker & Jane>).

To address this idea, Skordos and Papafragou (2016) conducted a set of studies with 5-year-olds. They used an acceptability judgment task (cf. Papafragou & Musolino, 2003), in which children were presented with simple visual scenes featuring novel creatures (“blickets”) who possessed common everyday items. A puppet described the scenes with a statement featuring the quantifier *some* or *all* and children were asked to judge the puppet's utterances. Statements with *all* were true or false and statements with *some* always true, but sometimes infelicitous (critical trials). In a first experiment, the authors tested whether the availability of the stronger scalar alternative *all* could help children derive scalar implicatures when judging utterances containing the weak alternative *some*. Indeed, when trials with *all* and trials with *some* were intermixed in the experimental battery so that *all* was contextually available, children derived scalar implicatures at rates comparable to adults. When, however, the trials with *some* preceded the trials with *all* in a blocked design (such that they never heard *all* before the *some*

trials were completed), children's performance dropped significantly. Having established that priming stronger alternatives helps children, Skordos and Papafragou conducted a second experiment to further investigate why. Their hypothesis was that priming alternatives helps by making it easier for children to understand the 'conversational goal' of the communicative exchange, a Gricean notion analogous to the idea of the QUD. In this experiment, test trials with the quantifier *all* always preceded trials with *some* so that the stronger scalar alternative would always be accessible to children. Critically, in one condition, the conversational goal (QUD) was about whether *all* or *only some* of the "blickets" had an item that the puppet mentioned. This was done by manipulating the false *all*-trials so that the items the "blickets" had were always those mentioned by the puppet and instead the truth of the statement hinged on the number of "blickets" having an item (3/4 or 4/4). This QUD was expected to prime the *some-but-not-all* scalar inference. In another condition, the test trials highlighted a different conversational goal (QUD), namely whether the "blickets" had the item mentioned, or a different one instead. In this case, all four of the blickets always had a certain type of item (so no blicket was left without one) and the truth of the trials hinged on whether the "blickets" possessed the item the puppet had mentioned (e.g., a crayon), or a different one (e.g., an umbrella), such that the QUD concerned the kinds of things that the blickets possessed, instead of how many blickets possessed an item.

Skordos and Papafragou found that despite access to the stronger scalar alternative *all* across the two conditions, children's judgments were critically affected by the QUD. In the condition highlighting the 'quantity' QUD and priming the *some-but-not-all* scalar inference, children overwhelmingly rejected the statement "some of the blickets have a crayon", when all 4 blickets had a crayon. When, however, the QUD was related to the type of item and consequently primed the *crayon-but-not-other-item* inference, children generally accepted the

very same statement as perfectly fine. Even more strikingly, in a third experiment, children derived the *some-but-not-all* inference, even when the experimenters did not make the stronger scalar alternative *all* accessible, so long as the QUD primed a scalar inference, for example with the quantifier *none* in the place of *all*. Since making *none* accessible does not automatically provide access to *all* and since children nevertheless compute implicatures that require accessing *all*, children's difficulty is likely not with the strength of associations between scale mates, but instead lies with their problems identifying the QUD and the inference to be made. Based on these findings, Skordos and Papafragou (2016) concluded that children's primary difficulty with scalar inferences is that they struggle to identify the conversational goal of these communicative exchanges, which can be straightforwardly translated as difficulty identifying the QUD.

Together, these findings raise the possibility that when relevant alternatives are identifiable, children use them in a way similar to adults to make inferences. If so, then contrary to Piaget and some subsequent accounts in linguistics, the logical and computational abilities required by inferential reasoning may be in place by the preschool years (Barner et al., 2011; Stiller et al., 2015). Also, the results suggest that, like adults, children can use contextually accessible alternatives to infer the QUD and consequently the inference they should make among several possible inferences. In contrast to adults, however, children find it more challenging to spontaneously recover the QUD, perhaps because the hypothesis space of possible inferences is too large for them to restrict without help, or perhaps because their ordering of likely hypotheses differs from that of adults (Pierce, 1955; Schulz, 2012). When the hypothesis space is restricted as in Skordos and Papafragou (2016), they do much better, including spontaneously generating the relevant alternatives.

Quantifier spreading

Consistent with the hypothesis that children's inferences are limited by their ability to identify relevant alternative utterances, there is evidence that access to alternatives may also restrict children's ability to make a range of additional logical inferences with quantifiers that are not pragmatic in nature, including the class inclusion examples originally studied by Inhelder and Piaget (1964). According to some accounts, including that of Inhelder and Piaget, children's difficulties with universal quantification reflect deficits in logical reasoning. Others, however, have argued that some form of pragmatic deficit may best explain children's class inclusion judgments (e.g. Crain et al., 1996). Extending the original Piagetian findings, a study by Philip (1991) showed children a scene in which three girls were each riding an elephant in the presence of a fourth, riderless, elephant. Surprisingly, when children were asked, "Is every girl riding an elephant?" they responded "No", sometimes explaining that the fourth elephant did not have a rider. According to some accounts, it was argued that such judgments arise because children interpret the universal quantifier *every* as applying at the sentence level, its scope spreading from the subject to the entire sentence including the verb object, resulting in a symmetrical interpretation equivalent to, "Every girl is riding an elephant, and every elephant is being ridden". Based on this, the phenomenon was dubbed by some as "quantifier spreading" (Roeper & de Villiers, 1993; cf. Brooks & Parshina, this volume). Other accounts proposed that these judgments might arise because children quantify over events (e.g., instances of girls riding elephants) as opposed to entities (Philip, 1991, 1995), or because children otherwise misidentify the syntactic features of the quantifier (Kang, 2001).

Somewhat in the tradition of Piaget, others (Drozd & van Loosbroek, 1999; Drozd, 2001) have suggested that children might have a different semantics for the universal quantifier. Errors

arise because children treat *every* as a “weak” quantifier, similar to *many*, and allow the quantifier to combine with an argument that can be determined contextually, sometimes switching to the verb object (Westerståhl, 1985). For example, consider the sentence in (5):

(5) Many Finns are World Rally Championship drivers.

(a) A large proportion of the Finnish population are WRC drivers.

(b) A large proportion of the WRC drivers are Finnish.

The canonical reading of (5) corresponds to (5a), which we know to be false (it is not true that a large number of the set of Finnish citizens are WRC drivers). Yet people judge (5) to be true with a reading like (5b), effectively allowing the quantifier *many* to combine with the object of the sentence as opposed to the subject. In this case, adults appear to disregard the ordinary “proportional” meaning of the quantified sentence, as computed by syntax and semantics, and use their pragmatics to make an inference about what the speaker means (cf. Cohen, 2001), in a way not dissimilar to what children might be doing with *every*.

Other accounts propose that children’s difficulty does not arise in semantics or syntax per se, but rather in their interface, when children attempt to map the complex semantics of quantifiers to their surface structure. For example, Geurts (2003) argues that although children have an adult-like “strong” semantics of *every*, processing difficulties associated with the complexity of quantified expressions lead children to effectively treat them as weak quantifiers instead. Specifically, overwhelmed by the complexity of these sentences, children instead use context to restrict quantification, allowing the quantifier to combine with either the subject or the object of the sentence (e.g., girls or elephants), depending on which is contextually salient. Similar arguments that processing demands may impair children are relatively common in the literature (Freeman et al., 1982; cf. Brooks & Sekerina, 2005) and as we already saw appear

frequently in developmental studies of scalar implicature as well. In both cases, however, no direct evidence was offered to support this possibility, and no detailed causal models were put forward to explain precisely how processing limits might cause quantifier spreading errors or scalar implicature failures.

An Alternative

In recent work, we have begun to explore the idea that children’s problems with class inclusion stem from the same types of difficulties they experience with implicature. Specifically, children fail to interpret them in adult-like ways because they struggle to identify the QUD, and consequently do not access relevant alternatives that are required for interpreting the utterance. Further, the QUD can be identified in one of two ways: (1) via explicit mention of relevant linguistic alternatives, or (2) by mention of different contextual states of affairs that render specific linguistic alternative sets relevant. For example, in a context including a man, a man with sunglasses, and a man with a hat and sunglasses, then “the man with the sunglasses” rules out reference to the man who is wearing both sunglasses and a hat. However, the same utterance is compatible with this man if there is no man who is wearing *only* sunglasses. In this context, “the man with sunglasses” is compatible with the hat-wearing man, because now, due to context, the stronger statement involving a hat is no longer considered as a linguistic alternative. Such inferences can be modeled in a neo-Gricean epistemic framework, like the Rational Speech Act account (Frank & Goodman, 2012), but are also compatible with non-epistemic logics, so long as they incorporate models of both utterances and possible states of affairs / referential domains. Critically, our proposal is that by adopting this type of joint inference, the same framework

which explains children's difficulties with scalar implicature can be extended to explain quantifier spreading, too.

First, consider the case of scalar implicature. In order to comprehend an utterance like "Some of the boys have bicycles", one needs to consider both the alternative linguistic forms that the speaker could have uttered, but also the alternative meanings (or states of affairs) that the utterance is compatible with, which might be provided contextually. Critically, these are related. For example, in a context including a group of boys and a group of girls where all of the boys have bicycles and all of the girls have skateboards, salient alternatives to "Some of the boys have bicycles" might involve skateboards (e.g., "Some of the boys have skateboards"). However, in a context in which some of the boys and all of the girls have bicycles, such alternatives might be less salient than an alternative such as "All of the boys have bicycles". This is because in each case, the context provides a particular dimension of contrast (e.g., between kind of vehicle in the first, and number of vehicles in the second). Similarly, the relevant alternatives can also be elicited via direct mention, making the QUD explicit – e.g., if a speaker asks "What kinds of things do the boys have?" or "Do all of the boys have bicycles?" Context makes available different states of affairs, and therefore elicits different QUDs, resulting in the licensing of different linguistic alternatives.

To see how this might work for children in the case of quantifier spreading, it is useful to consider how Crain and colleagues explain children's judgments (Crain et al., 1996). According to Crain et al., children appear to have difficulties when asked to judge universally quantified sentences because the tasks used in the literature often fail to satisfy what they call the condition of 'plausible dissent'. For a question like, "Is every girl riding an elephant?" to be felicitous, the affirmative response (*Yes*) must have been in doubt (in the past) and not clearly incontrovertible.

For this to happen, an alternative outcome (e.g., a girl riding a donkey) should be conceivable to the participant. If such an outcome is not accessible, children are left to infer that other interpretations of the question must be relevant and respond based on those instead (e.g., the numerical correspondence between girls and elephants). This explanation can be straightforwardly recast as children having difficulties identifying the QUD. In quantifier spreading cases, while adults identify the QUD as being about the number of girls riding elephants because they consider an alternative in which not all three girls are on elephants, children may fail because they cannot generate this alternative state of affairs (i.e., the possible intended meaning), leading them to instead base their judgment on another dimension, such as the mismatch between the scene and their prior expectations given the utterance – i.e., that there should be a set of girls riding elephants, but no extra characters, or extra animals. As evidence for the condition of ‘plausible dissent’, Crain and colleagues (1996) show that when children are provided with a plausible alternative outcome (e.g., “some farmer feeds a dinosaur” for “every farmer feeds a donkey”) in the form of short background stories, children’s responses become more adult-like.

The notion of ‘plausible dissent’ captures an important intuition: The evaluation of statements inviting inferences in natural language requires more than merely processing the utterance and checking that it matches a semantic model. In addition, it also involves inferring the QUD and thus the dimension along which an inference should be made: Is it the number of girls riding elephants that matters, the kinds of things being ridden, or the number of elephants being ridden? Given clear contextual cues regarding a state of affairs that would make the utterance false (i.e., a basis for plausible dissent), children might identify the relevant dimension, and thus the QUD, thereby allowing them to make adult-like inferences. In support of this idea, a recent

study by Skordos, Myers, and Barner (under review) attempted to address this question by priming different possible alternative states of affairs and their corresponding QUDs. In this study, preschoolers were shown simple cartoon pictures in which three girls were each riding an elephant (or driving a car, riding a bicycle, etc.) and an extra elephant (car, bicycle, etc.) was present in the scene. Children were asked to answer the question, e.g., “Is every girl riding an elephant?” In this baseline condition, children answered *Yes* only 33% of the time, a result consistent with prior work. In a second, experimental condition, the critical trials were preceded by scenes in which only two of the three girls were riding elephants, while the third girl was on foot and two more elephants were without riders. These trials were intended to provide children with alternative states of affairs that might indicate to children the QUD – i.e., whether these three girls were riding elephants. When children were asked whether every girl was riding an elephant, they overwhelmingly answered *No* to the priming trials and subsequently answered *Yes* to the critical trials 78% of the time; a large and significant improvement compared to the baseline and much closer to adult performance.

Still, even in this scenario children differed modestly from adults, leaving open the possibility that other factors contributed to their difficulties. In the condition just described, children were assisted by being explicitly provided an alternative state of affairs that afforded a basis for plausible dissent. In the language of joint inference with which we began this section, these children were provided with a broader set of meaning hypotheses against which an utterance might be evaluated, which in this case played the role of indicating the dimension along which the speaker intended the utterance to be evaluated – i.e., whether all three girls were riding elephants. However, in the absence of such an explicitly provided alternative, children might generate their own, and might do so in ways that also contribute to their non-adult-like

behaviors. One such possibility is raised by Drozd and van Loosbroek (1999), who suggest that children may struggle with domain restriction. In particular, upon seeing a riderless elephant, they may infer that, since elephants in this world generally belong to girls, there must exist another girl who is not on her designated elephant, such that it is legitimate to conclude that not every girl is riding an elephant (since the missing girl is not). To test this, Skordos et al. (under review) tested children in a condition like the previous ones, but this time instead of priming children with a situation in which the utterance was false, they instead introduced children to three girls, by name, emphasizing that they would be the only three girls under discussion. At test, when shown, as before, the three girls riding four elephants, children judged that every girl was riding an elephant 49% of the time, which appears larger than the 33% of the baseline, but was not statistically reliable. Consequently, we concluded that at least in this dataset, domain restriction did not play a reliable role in children's judgements: Children did not primarily make errors because they *overgenerated* alternative states of affairs, imagining situations not intended by the speaker. A follow-up experiment, that combined the domain restriction cues (evidence that only three girls existed) with the false alternative cue (an example of only two girls riding) found that children performed near ceiling on the task (92%), and thus almost never showed the type of error originally reported by Philip (1991). Together, these data suggest that a combination of plausible dissent and domain restriction can likely account for all of the variability that differentiates children from adults.

Based on these recent findings, it seems plausible that the phenomenon of quantifier spreading can be given a treatment similar to that originally applied to scalar implicature: The difficulties children have when reasoning with quantified statements are in large part a result of their difficulty restricting the inferential hypothesis space in language comprehension. As

suggested above, on this view, language comprehension involves a kind of abductive inference, wherein the listener attempts to identify the meaning of an utterance from what is initially a large (if not infinite) space of possible alternative meanings. The hypothesis space is only partially narrowed by the literal semantics of utterances, with the result that the listener must consider alternative linguistic forms that the speaker could have used but didn't, as well as the possible states of affairs which the utterance might plausibly be taken to encode. Children's difficulties in narrowing the inferential hypothesis space can be mitigated by priming linguistic alternatives that are only consistent with a particular QUD, or by presenting alternative scenes that focus children's attention to that QUD, by making a particular dimension of the scenes salient. In both the case of implicature and quantifier spreading, once children have identified the QUD, they readily narrow their inferences, access any required linguistic alternatives, and make judgments that are almost perfectly adult-like.

Concluding remarks

In this chapter, we initially discussed the importance of pragmatic inference involving alternatives for language comprehension, highlighting that utterances are multiply ambiguous and can lead to different interpretations as a result of different inferences that can be made. We reviewed the problem of restricting this inferential hypothesis space and presented a brief overview of theoretical and empirical work on adults that argues for the role of alternative linguistic forms and alternative meanings in restricting the hypothesis space and guiding both scalar inferences but also contrastive inferences triggered by focus elements. We then turned to developmental evidence and presented two characteristic case studies from language acquisition, scalar implicature and quantifier spreading. In both cases, children who are otherwise competent

users of language appear to struggle when interpreting sentences including quantifiers. We argued that in both scalar implicature and quantifier spreading, children's problems are closely linked to difficulties in reducing the inferential hypothesis space, while matching what is said to what is meant. We argued that a substantial proportion of those problems are caused by the fact that children often misidentify the QUD, which leads them to consider irrelevant alternatives and make non-adult-like inferences. When relevant alternatives are made salient and the QUD is appropriately identified, children use it to make inferences in an adult-like manner. While alternatives can be used within the course of an experiment to lead children to the appropriate QUD and restrict inference, further work is need to understand why it might be that children struggle to identify the QUD, forcing them to consider the wrong alternatives and make non-adult-like inferences.

References:

- Akhtar, N. & Tomasello, M. (2000). The social nature of words and word learning. In R. M. Golinkoff, et al. (eds.), *Becoming a word learner: a debate on lexical acquisition*. Oxford: Oxford University Press.
- Aloni, M. (2003). Free choice in modal contexts. In M. Weisgerber (Ed.), *Proceedings of the Conference "sub7 – Sinn und Bedeutung"*, (pp. 25–37). University of Konstanz. Vol. 114 of *Arbeitspapiere des Fachbereichs Sprachwissenschaft*.
- Alonso-Ovalle, L. (2009). Counterfactuals, correlatives, and disjunction. *Linguistics and Philosophy*, 32, 207–244.
- Baldwin, D. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language*, 20, 395-418.
- Bale, A. & Barner, D. (2013). Grammatical alternatives and pragmatic development. In Anamaria Falaus (Ed.), *Alternatives in Semantics, 'Studies in Pragmatics, Language and Cognition'*. (pp. 238-266). Palgrave Macmillan.
- Barner, D., Brooks, N., & Bale, A. (2011). Accessing the unsaid: The role of scalar alternatives in children's pragmatic inference. *Cognition*, 188, 87-96.
- Barner, D., & Bachrach, A. (2010). Inference and exact numerical representation in early language development. *Cognitive Psychology*, 60(1), 40-62.
- Barner, D., Chow, K., & Yang, S., J. (2009). Finding one's meaning: A test of the relation between quantifiers and integers in language development. *Cognitive Psychology*, 58, 195–219.

- Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science*, *8*(6), 492-499.
- Bergen, L. & Goodman, N. (2015). The strategic use of noise in pragmatic reasoning. *Topics in Cognitive Science*, *7*, 336-350. DOI: 10.1111/tops.12144
- Bergen, L., & Grodner, D. J. (2012). Speaker knowledge influences the comprehension of pragmatic inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(5), 1450.
- Bergen, L., Goodman, N., & Levy, R. (2012). That's what she (could have) said: How alternative utterances affect language use. In N. Miyake, D. Peebles, & R. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 120-125). Austin, TX: Cognitive Science Society.
- Bergen, L., Levy, R., & Goodman, N. (2016). Pragmatic reasoning through semantic inference. *Semantics and Pragmatics*, *9*(20). DOI: <http://dx.doi.org/10.3765/sp.9.20>
- Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT Press.
- Bonnefon, J.-F., Feeney, A., & Villejoubert, G. (2009). When some is actually all: Scalar inferences in face-threatening contexts. *Cognition*, *112*(2), 249–258.
- Bott, L., & Noveck, I. A. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of memory and language*, *51*(3), 437-457.
- Bott, L., Bailey, T. M., & Grodner, D. (2012). Distinguishing speed from accuracy in scalar implicatures. *Journal of Memory and Language*, *66*(1), 123-142.
- Braine, M., & Romain, B. (1981). Children's comprehension of “or”: evidence for a sequence of competencies. *Journal of Experimental Child Psychology*, *31*, 46-70.

- Breheny, R., Ferguson, H. J., & Katsos, N. (2013). Taking the epistemic step: toward a model of on-line access to conversational implicatures. *Cognition*, *126*(3), 423-440.
- Breheny, R., Katsos, N., & Williams, J. (2006). Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition*, *100*(3), 434-463.
- Brooks, P.J., & Parshina, O. (in press). Quantifier Spreading. In C. Cummins & N. Katsos (Eds.), *Handbook of Experimental Semantics and Pragmatics*. Oxford: Oxford University Press.
- Brooks, P.J., & Sekerina, I. (2005). Shortcuts to quantifier interpretation in children and adults. *Language Acquisition*, *13*, 177-206.
- Brooks, N., Audet, J., & Barner, D. (2013). Pragmatic inference, not semantic competence, guides 3-year-olds' interpretation of unknown number words. *Developmental Psychology*, *49*(6), 1066-1075.
- Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, *112*, 337-342.
- Carston, R. (1995). Quantity maxims and generalised implicature. *Lingua*, *96*, 21-44.
- Chemla, E., & Bott, L. (2014). Processing inferences at the semantics/pragmatics frontier: Disjunctions and free choice. *Cognition*, *130*(3), 380-396. DOI: 10.1016/j.cognition.2013.11.013
- Chevallier, C., Wilson, D., Happé, F., & Noveck, I. (2010). Scalar inferences in autism spectrum disorders. *Journal of autism and developmental disorders*, *40*(9), 1104-1117.

- Chevallier, C., Noveck, I. A., Nazir, T., Bott, L., Lanzetti, V., & Sperber, D. (2008). Making disjunctions exclusive. *The Quarterly Journal of Experimental Psychology*, *61*(11), 1741–1760. DOI:10.1080/17470210701712960
- Chierchia, G., Crain, S., Guasti, M., T., Gualmini, A., & Meroni L. (2001). The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures. *Proceedings from the Annual Boston University Conference on Language Development*, *25*, 157-168.
- Chierchia, G., Fox, D., & Spector, B. (2009). Hurford's Constraint and the Theory of Scalar Implicatures. In: P. Egré & G. Magri (Eds.), *Presuppositions and Implicatures*, (pp. 47–62). Proceedings of the MIT-Paris Workshop. Cambridge, MA: MIT Working Papers in Linguistics.
- Chierchia, G., (2004). Scalar implicatures, polarity phenomena and the syntax / pragmatics interface. In: Belletti, A. (Ed.), *Structures and Beyond*, (pp. 39-103). Oxford University Press, Oxford.
- Chierchia, G., (2013). *Logic in Grammar: Polarity, Free Choice, and Intervention*. Oxford University Press
- Clark, E. V., & Amaral, P. M. (2010). Children build on pragmatic information in language acquisition. *Language and Linguistics Compass*, *4*, 445-457. DOI: 10.1111/j.1749-818X.2010.00214.x
- Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition*, (pp. 1–33). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Clark, E. V. (1990). On the pragmatics of contrast. *Journal of Child Language*, *17*, 417–431.

- Cohen, A. (2001). Relative Readings of Many, Often, and Generics. *Natural Language Semantics*, 9(1), 41–67. DOI: 10.1023/A:1017913406219
- Condry, K. & Spelke, E. (2008). The Development of Language and Abstract Concepts: The Case of Natural Number. *Journal of Experimental Psychology: General*, 137(1), 22-38. DOI: 10.1037/0096-3445.137.1.22
- Conroy, A., Lidz, J., and Musolino, J. (2009). The fleeting Isomorphism Effect. *Language Acquisition*, 16(2), 106-117.
- Corkum, V., & Moore, C. (1998). The origins of joint visual attention in infants. *Developmental psychology*, 34(1), 28.
- Crain, S., Thornton, R., Boster, C., Conway, L., Lillo-Martin, D., & Woodams, E. (1996). Quantification Without Qualification. *Language Acquisition*, 5(2), 83-153. DOI: 10.1207/s15327817la0502_2
- De Neys, W., & Schaeken, W. (2007). When people are more logical under cognitive load. *Experimental Psychology (formerly Zeitschrift für Experimentelle Psychologie)*, 54(2), 128-133.
- Degen, J., & Tanenhaus, M. K. (2015). Processing scalar implicature: A constraint-based approach. *Cognitive Science*, 39(4), 667–710. DOI:10.1111/cogs.12171
- Degen, J. & Tanenhaus, M. K. (2016). Availability of alternatives and the processing of scalar implicatures: A visual world eye-tracking study. *Cognitive Science*. 40(1), 172-201. DOI: 10.1111/cogs.12227
- Degen, J. & Tanenhaus, M. K. (in press). Constraint-based pragmatic processing. In C. Cummins & N. Katsos (Eds.), *Handbook of Experimental Semantics and Pragmatics*. Oxford: Oxford University Press.

- Degen, J., Franke, M., & Jäger, G. (2013). Cost-based pragmatic inference about referential expressions. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 376-381). Austin, TX: Cognitive Science Society.
- Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology*, 37, 630–641.
- Drozd, K. F. (2001). Children's Weak Interpretations of Universally Quantified Questions. In M. Bowerman and S. C. Levinson, eds., *Language Acquisition and Conceptual Development*, Cambridge University Press, Cambridge, England.
- Drozd, K. & van Loosbroek, E. (1999). Weak quantification, plausible dissent, and the development of children's pragmatic competence. In A. Greenhill, H. Littlefield & C. Tano (Eds.), *The Proceedings of the 23d Annual Boston University Conference on Language Development* (pp. 184-195). Somerville MA: Cascadilla Press.
- Feeney, A., Scafton, S., Duckworth, A., & Handley, S. (2004). The story of some: everyday pragmatic inference by children and adults. *Canadian Journal of Experimental Psychology*, 58:2, 121-132.
- Foppolo, F., Guasti, M.T., & Chierchia, G. (2012). Scalar implicatures in child language: Give children a chance. *Language Learning and Development*, 8, 365-394. Psychology Press.
- Fox, D. (2007). Free Choice Disjunction and the Theory of Scalar Implicatures. In: U. Sauerland and P. Stateva (Eds.), *Presupposition and Implicature in Compositional Semantics*. New York: Palgrave Macmillan.
- Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, 336(6084), 998-998.

- Frank, M., Goodman, N. D., & Tenenbaum, J. B. (2009). Using speakers' referential intentions to model early cross-situational word learning. *Psychological Science*, *20*(5), 578–585.
- Franke, M. (2009). *Signal to Act: Game Theory in Pragmatics*. Doctoral Dissertation, Universiteit van Amsterdam.
- Franke, M. (2014). Typical use of quantifiers: A probabilistic model. In Bello, P., Guarini, M., McShane, M., & Scassellati, B. (Eds.), *Proceedings of the 36th Annual Meeting of the Cognitive Science Society* (pp. 487-492). Austin, TX: Cognitive Science Society.
- Freeman, N., Sinha, C. & Stedmon, J. (1982). All the cars – which cars? From word meaning to discourse analysis. In M. Beveridge (Ed.), *Children Thinking through Language* (pp. 53-74). London: Edward Arnold.
- Gazdar, G. (1979). *Pragmatics: Implicature, presupposition, and logical form*. New York: Academic Press.
- Gergely, G., Bekkering, H., & Király, I. (2002). Developmental psychology: Rational imitation in preverbal infants. *Nature*, *415*, 755-756. DOI: 10.1038/415755a
- Geurts, B. (2003). Quantifying kids. *Language Acquisition*, *11*, 197–218.
- Gleitman, L. R., & Newport, E. L. (1995). The invention of language by children: Environmental and biological influences on the acquisition of language. In L. R. Gleitman & M. Liberman (Eds.), *Language* (Vol. 1). An invitation to Cognitive Science (2nd ed.). Cambridge: MIT Press.
- Goodman, N. & Stuhlmüller, A. (2013). Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science*, *5*, 173–184. DOI: 10.1111/tops.12007
- Grice, H., P. (1969). Utterer's Meaning and Intentions. *Philosophical Review* *68*, 2, 147-177.

- Grice, H., P. (1975). Logic and conversation. In P. Cole, and J., L. Morgan (Eds.) *Syntax and semantics: Speech acts*, Volume 3, (pp. 41–58). New York: Academic Press.
- Grodner, D., Klein, N., Carbary, K., Tanenhaus, M. (2010). “Some,” and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition*, *116(1)*, 42-55. DOI: 10.1016/j.cognition.2010.03.014
- Groenendijk, J. and Stokhof, M., 1984, *Studies in the semantics of questions and the pragmatics of answers*, Ph.D. Thesis, University of Amsterdam.
- Guasti, M. T., Chierchia, G., Crain, S., Foppolo F., Gualmini, A., & Meroni L. (2005). Why children and adults sometimes (but not always) compute implicatures. *Language and Cognitive Processes*, *20*, 667-696.
- Hirschberg, J. (1985). *A theory of scalar implicature*. Doctoral diss., University of Pennsylvania.
- Hochstein, L., Bale, A., & Barner, D. (2017). Scalar implicature in absence of epistemic reasoning? The case of autism spectrum disorder. *Language Learning and Development*, 1-17.
- Hochstein, L., Bale, A., Fox, D., & Barner, D. (2014). Ignorance and inference: Do problems with Gricean epistemic reasoning explain children’s difficulty with scalar implicature? *Journal of Semantics*, *33(1)*, 107-135. DOI: 10.1093/jos/ffu015
- Horn, L., R. (1972). *On the semantic properties of the logical operators in English*. Doctoral diss., UCLA.
- Horn, L. R. (2006). *The border wars: A neo-Gricean perspective. Where Semantics Meets Pragmatics: The Michigan Papers*, (pp. 21-48). Oxford: Elsevier.
- Horowitz, A. C., Schneider, R. M., Frank, M. C. (in press). The trouble with quantifiers: Explaining children's deficits in scalar implicature. *Child Development*.

- Huang, Y. T., & Snedeker, J. (2009). Semantic meaning and pragmatic interpretation in five-year olds: Evidence from real time spoken language comprehension. *Developmental Psychology, 45*, 1723–1739.
- Huang, Y. T. & Snedeker, J. (2011). Logic and conversation revisited: Evidence for a division between semantic and pragmatic content in real-time language comprehension. *Language and Cognitive Processes, 26(8)*, 1161-1172.
- Hulsey, S., Hacquard, V., Fox, D., & Gualmini, A. (2004). The question answer requirement and scope assignment. In *Plato's problems: Papers on language acquisition. MIT Working Papers in Linguistics 48*. Cambridge, MA: MITWPL.
- Jäger, G. (2011). Game-theoretical pragmatics. In J. F. van Benthem & A. ter Meulen (Eds.), *Handbook of logic and language* (2nd ed., pp. 467–491). London: Elsevier.
- Kang, H.-K. (2001). Quantifier spreading: Linguistic and pragmatic considerations. *Lingua, 111*, 591-627.
- Katsos, N., & Bishop, D. V. (2011). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition, 120(1)*, 67-81.
- Katsos, N., Cummins, C., Ezeizabarrena, M., Gavarró, A., Kraljević, J. K., et al. (2016). Cross-linguistic patterns in the acquisition of quantifiers. *Proceedings of the National Academy of Sciences of the United States of America, 113(33)*, 9244–9249, DOI: 10.1073/pnas.1601341113
- Kim, S., C. (in press) Experimental investigations on focus. In C. Cummins & N. Katsos (Eds.), *Handbook of Experimental Semantics and Pragmatics*. Oxford: Oxford University Press.

- Krifka, M. (1995). The semantics and pragmatics of polarity items. *Linguistic analysis*, 25, 209–257.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press, Cambridge.
- Levinson, S. (2000). *Presumptive meanings*. Cambridge, MA: MIT Press.
- Lidz, J. & Musolino, J. (2002). Children’s Command of Quantification. *Cognition*, 84, 113–154.
- Liszkowski, U., Carpenter, M, & Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition*, 108, 732-739.
- Loukusa, S., & Moilanen, I. (2009). Pragmatic inference abilities in individuals with Asperger syndrome or high-functioning autism. A review. *Research in Autism Spectrum Disorders*, 3(4), 890-904.
- Markman, E. M., & Wachtel, G. F. (1988). Children’s use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, 20, 121–157.
- Meltzoff, A. (1995). Understanding the Intentions of Others: Re-Enactment of Intended Acts by 18-Month-Old Children. *Developmental Psychology*, 31(5), 838-850.
- Noveck, I. & Posada, A. (2003). Characterizing the time course of an implicature: An evoked potentials study. *Brain and Language*, 85, 203-210. DOI:10.1016/S0093-934X(03)00053-1
- Noveck, I. A., Guelminger, R., Georgieff, N., & Labruyere, N. (2007). What autism can reveal about every... not sentences. *Journal of Semantics*, 24(1), 73-90.
- Noveck, I., A., & Sperber, D. (2007). The why and how of experimental pragmatics: The case of ‘scalar inferences’. In N. Burton-Roberts (Ed.) *Advances in Pragmatics*. Basingstoke: Palgrave.

- Noveck, I. (2001). When children are more logical than adults: Experimental investigations of scalar implicatures. *Cognition*, 78, 165-188.
- Onishi, K.H., Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308, 255-258.
- Ozturk, O., & Papafragou, A. (2015). The acquisition of epistemic modality: From semantic meaning to pragmatic interpretation. *Language, Learning and Development*, 11:3, 191-214, DOI: 10.1080/15475441.2014.905169
- Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics pragmatics interface. *Cognition*, 86, 253-282.
- Paris, S. 1973. Comprehension of language connectives and propositional logical relationships. *Journal of Experimental Child Psychology*, 16: 278–291.
- Philip, W. (1991). *Spreading in the Acquisition of Universal Quantifiers*. WCCFL 10: 359-373.
- Philip, W. (1995). *Event quantification in the acquisition of universal quantification*, Doctoral dissertation. University of Massachusetts, Amherst.
- Inhelder, B. & Piaget, J. (1964). *The Growth of Logical Thinking from Childhood to Adolescence*. New York: Basic Books.
- Peirce, C. S. (1955). *Abduction and induction*. Philosophical writings of Peirce, 11.
- Politzer-Ahles, S. & Gwilliams, L. (2015). Involvement of prefrontal cortex in scalar implicatures: evidence from magnetoencephalography. *Language, Cognition and Neuroscience*, 30(7), 853-866. DOI: 10.1080/23273798.2015.1027235
- Pouscoulous, N., Noveck, I., Politzer, G., Bastide, A. (2007). Processing costs and implicature development. *Language Acquisition*, 14, 347-375.

- Reinhart, T. (2004). The processing cost of reference set computation: Acquisition of stress shift and focus. *Language Acquisition*, 12(2), 109-155.
- Roberts, C. (1996). Information structure: towards an integrated theory of formal pragmatics. *OSU Working Papers in Linguistics 49, Papers in Semantics*, (pp. 91—136).
- Roberts, C. (2004). Information structure in discourse. *Semantics and Pragmatics*, 5, 1 – 69.
- Roeper, T. & De Villiers, J. (1993). Emergence of bound variable structures. In E. Reuland & W. Abraham (Eds.), *Knowledge and Language: From Orwell's Problem to Plato's Problem* (pp. 105–139). Dordrecht: Kluwer.
- Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, 1:1, 75–116.
- Rooth, M. (2016). Alternative semantics. In C. Féry & S. Ishihara (Eds.), *Oxford Handbook of Information Structure* (pp. 19-40). Oxford: Oxford University Press.
- Russell, B. (2012). *Probabilistic Reasoning and the Computation of Scalar Implicatures*. Unpublished doctoral dissertation, Brown University.
- Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and Philosophy*, 27, 367–91.
- Schulz, L. E. (2012). Finding new facts; thinking new thoughts. Rational constructivism, *Advances in child development and behavior*, 42.
- Simons, M. (2005). Dividing things up: the semantics of or and the modal or interaction. *Natural Language Semantics*, 13, 271–316.
- Singh, R., Wexler, K., Astle-Rahim, A., Kamawar, D., & Fox, D. (2016). Children interpret disjunction as conjunction: Consequences for the theories of implicature and child development. *Natural Language Semantics*, 24(4), 305-352. DOI: 10.1007/s11050-016-9126-3

- Skordos, D., & Papafragou, A. (2016). Children's derivation of scalar implicatures: alternatives and relevance. *Cognition*, *153*, 6–18.
- Skordos, D., Myers, A., & Barner, D. (under review). Access to alternatives and universal quantification in language acquisition.
- Smith, C., L. (1980). Quantifiers and question answering in young children. *Journal of Experimental Child Psychology*, *30*, 191–205.
- Southgate, V., Senju, A., & Csibra, G. (2007). Action Anticipation Through Attribution of False Belief by 2-Year-Olds. *Psychological Science*, *18*(7), 587-592.
- Spector, B. (2007). Aspects of the pragmatics of plural morphology: On higher-order implicatures. In U. Sauerland & P. Stateva (Eds.), *Presupposition and Implicature in Compositional Semantics*, (pp. 71–120). New York: Palgrave Macmillan.
- Sperber, D., & Wilson, D., (1986/1995). *Relevance: Communication and Cognition*, 2nd ed. 1995. Cambridge, MA: Harvard University Press.
- Spychalska, M., Kontinen, J., & Werning, M. (2016). Investigating scalar implicatures in a truth-value judgment task: Evidence from event-related brain potentials. *Language, Cognition and Neuroscience*, *31*(6), 817-840. DOI: 10.1080/23273798.2016.1161806
- Stalnaker, R. 1979. Assertion. In: Cole, P. (Ed.), *Syntax and Semantics 9*, (pp. 315-332). Academic Press, New York.
- Stiller, A., Goodman, N. D., & Frank, M. C. (2011). Ad-hoc scalar implicature in adults and children. In *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*, Boston, MA.

- Tieu, L., Romoli, J., Zhou, P., & Crain, S. (2016). Children's knowledge of free choice inferences and scalar implicatures. *Journal of Semantics*, 33(2), 269–298.
DOI:10.1093/jos/ffv001
- Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origins and role in development*, (pp. 103-130). Hillsdale, NJ: Erlbaum.
- Tomasello, M., Carpenter, M., Call, J., Behne, T. & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28(5), 675–735. DOI: 10.1017/S0140525X05000129
- Tomlinson, J., Bailey, T., & Bott, L. (2013). Possibly all of that and then some: Scalar implicatures are understood in two steps. *Journal of Memory and Language*, 69, 18-35.
- van Rooij, R. & K. Schulz (2004). Exhaustive interpretation of complex sentences. *Journal of Logic, Language and Information* 13(4): 491-519.
- Westerståhl, D. (1985). Determiners and context sets. In J. van Benthem & A. ter Meulen (Eds.), *Generalized quantifiers in natural language* (pp. 45-72). Dordrecht: Foris Publications.
- Woodward, A., & Markman, E. (1998). Early word learning. In W. Damon (Ed.), *Handbook of child psychology: Vol. 2. Cognition perception and language*, (pp. 371–20). Hoboken, NJ: John Wiley & Sons, Inc.
- Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, 24, 220-251.