

What Autism Can Reveal About *Every . . . not* Sentences

IRA A. NOVECK, RAPHAËLE GUELMINGER,
NICOLAS GEORGIEFF, AND NELLY LABRUYERE
Université de Lyon and CNRS, France

Abstract

The sentence *Every horse did not jump over the fence* can be interpreted with the negation taking scope over the quantifier (i.e. *not every horse jumped*) or with the quantifier *Every* taking scope over the negation (ultimately providing the reading *no horse jumped*). Beginning with Musolino, Crain and Thornton (2000), much work has shown that while adults typically adopt a *Not every* reading in ‘2-of-3’ contexts (e.g. where 2-of-3 horses jump over a fence), children do not and often produce *None* readings instead. In line with suggestions from Musolino and Lidz (2003, 2006), we propose that this developmental effect relies to a great extent on pragmatic capacities. In the present work, we aim to replicate Musolino *et al.*’s (2000) results with 4-year-olds and adults while including verbally competent autistic participants. Syntactic skills among verbally competent autistic participants are assumed to be unimpaired while their pragmatic deficiencies have been well documented. Our results show an adult preference for the *Not every* reading in 2-of-3 contexts and equivocality among children and autistic participants. This is in line with the expectation that syntax makes the two readings equally available and that adults, unlike young children and autistic participants, are efficient at exploiting the context in order to come up with a single consistent reading.

1 INTRODUCTION

Negations are notorious for creating ambiguous readings. For example, the sentence in (1) can be understood in two different ways.

(1) Every horse did not jump over the fence.

When the negation takes scope over the quantifier (*Not every*), one arrives at the interpretation in (2) and when the quantifier takes scope

over the negation (*Every ... not* or *None*)¹ one arrives at the interpretation in (3).

- (2) Not every horse jumped over the fence.
- (3) All horses are such that they did not jump over the fence.

The present work was initially inspired by Musolino *et al.* (2000, Experiment 2) who reported that adults are more likely than children to accept the *Not every* interpretation, as in (2), after being shown that two of three horses jumped over a fence. Meanwhile, children demonstrate ambiguity between the two readings (Musolino & Lidz 2006, Experiment 1) or a preference for a *None* reading, as in (3) (Musolino *et al.* 2000). After several follow-up investigations, Musolino & Lidz (2003, 2006; Lidz & Musolino, 2002) currently hypothesize that *Every ... not* sentences are syntactically ambiguous but that either performance factors play a role among children (Musolino & Lidz, 2006) or that scope relations favour a *None* reading among children (Musolino *et al.* 2000).² That is, children make their judgements by choosing (at random) one of the two possible readings or because there are reasons that encourage a *None* reading (based on contextual factors or parsing mechanisms). Critical to these claims is the notion that an adult, unlike a child, has access to each reading as well as its potential import and that preferring one (over the other) involves some extra attentional abilities at the point where the utterance is integrated with context.

Musolino & Lidz (2003, 2006) invoke pragmatics in three different ways in order to explain why adults obtain the interpretation in (2) more often than children, even though the parser makes both readings available. One suggestion is the Charity account, or ‘Principle of Charity’ (Musolino & Lidz, 2003), according to which adults seek out a true interpretation; this makes the true *Not every* reading more attractive than a false *None* reading in their tasks. A second suggestion they consider (Musolino & Lidz 2006: 819–822) is what we will call the Metalinguistic Scalar account, according to which the two interpretations arise via the syntax; through a scalar implicature (which we describe below), the weaker interpretation (*Not every*) leads a listener to reject the stronger one (*None*), even if the latter is compatible with the context. Finally, a third suggestion (Musolino &

¹ The reading in (3) can be understood more readily as *None of the horses jumped over the fence* and will be referred to as such (the *None* reading).

² The ‘first parse’ allocates scope because the quantificational expression (*Every horse* in (1)) c-commands the negation (*did not*). In their earlier work, the authors emphasized how children are susceptible to an “isomorphism” effect in which the surface syntax (*Every...not*) determines the preference for the *None* reading.

Lidz, 2006: 842) is that, in minimally elaborated circumstances, adults favour the *Not every* reading because a speaker who means to convey the *None* proposition could have said precisely that by using the more economical form of words, e.g. by saying *None of the horses jumped over the fence*. Assuming that (other things being equal) speakers do not put their audience to gratuitous effort, the fact that the speaker did not use a more economical form of words suggests she meant to convey the *Not every* proposition. In order to account for why adults, and not children, make the *Not every* interpretation, one would have to assume that it is the result of *acquiring* the Manner maxim over the course of language development. In this case, the *Not every* interpretation that emerges from *Every . . . not* sentences may well be viewed as automatic (or as a default). We will refer to this as the Acquired Manner account, in line with Grice's Manner maxim. To round out the picture for the pragmatics of *Every . . . not* sentences, below we present our own Processing account, which is drawn from Sperber & Wilson's (1995) Relevance Theory.

Relevance theory claims that (a) the speaker's meaning is inferred from the linguistic meaning of the words and expressions used taken together with the specific context at hand and that (b) effort factors play a role in utterance interpretation. We propose that the *Not every* interpretation of (1) in the presented experimental context (where 2-of-3 horses represent a subgroup) is more attractive than a *None* reading to an adult mostly because the 2-of-3 context is more readily consistent with a *Not every* reading. When a participant gets to the word *not* in *Every horse did not jump over the fence*, a *Not every* interpretation with respect to the 2-of-3 context allows a sophisticated hearer to even anticipate the rest of the sentence (*. . . jump over the fence*). With a *None* interpretation, a listener would be hard pressed to anticipate what it is that all of the horses are not doing. On the other hand, children—who are less efficient than adults at processing utterances and integrating them in context—generate the two readings as best they can and essentially choose randomly between them. Note that this account does not say that adults seek out a true interpretation *per se* like the Charity account, but that an interpretation that requires less effortful procedures will be more readily adopted.

As Musolino & Lidz (2006) themselves point out, the findings and suggestions from the papers on *Every . . . not* sentences resonate with the work showing that young children are less likely than adults to produce scalar implicatures (see Noveck 2001; Papafragou & Musolino 2003; Chierchia *et al.* 2004). According to standard neo-Gricean accounts (Horn 1984; Levinson 2000), implicatures arise when

a speaker's use of a relatively weak term allows the addressee to understand that the speaker has chosen not to articulate a stronger, more informative one from the same scale. Consider the weak quantifier *some*. From a semantic point of view, *some* is compatible with *all*, but quite often in conversation *some* will be understood as meaning *some but not all* because the speaker's choice of words is taken to mean that a stronger term from the same scale (*all*) is being avoided. Adults are more likely than children to transform *some* into *some but not all* and arguably because they are more apt at considering the stronger alternative in the first place or at considering the speaker's mental state in context (for summaries, see Noveck 2004; Noveck & Sperber in press).

That said, it is important to note that the two phenomena are distinct in two critical ways. First, a sentence like the one in (1) is genuinely *ambiguous* whereas scalar terms are not. That is, a *some but not all* enrichment is extra-linguistic. At the level of the linguistic code, *some*'s meaning remains unchanged and could suffice for providing responses on scalar implicature tasks even without the enrichment. This distinction—between the linguistic ambiguity for *Every . . . not* sentences and the extra-linguistic enrichment for scalars—explains why one finds different sorts of developmental patterns across these paradigms. Whereas adults 'can be turned into' children and vice versa in tests with the syntactically ambiguous structures (Musolino & Lidz 2003) children consistently remain 'less adult' than adults in the scalar tasks (Noveck 2001; Papafragou & Musolino 2003; Pouscoulous *et al.* in press). The developmental patterns in the scalar paradigm indicate that the pragmatic enrichment there is essentially optional; when it arrives it is part of a more involved pragmatic reading. Second, a *Not every* reading—which is generally preferred among adults in the *Every . . . not* studies—is actually weaker than the *None* reading, whereas the work on scalar implicatures generally shows that older participants aim for more informative readings (*some but not all* is more informative than *some*). It appears then that there is no simple pragmatic rule about adults and the way they abide by informative strength.

In light of the above comparison between *Every . . . not* sentences and weak scalar utterances (as well as the data they produce), we are doubtful about the Metalinguistic Scalar account. This account assumes that adult participants hear (1), generate interpretations (2) and (3), and order them with respect to informativeness before *opting* for the weaker of the two (which leads to the rejection of the stronger of the two); moreover, the background information is considered immaterial (the adoption of the weaker interpretation arises even if the stronger is

contextually plausible). There are three reasons why we consider it unlikely that adults would opt for the weaker reading as a first step. First, it is important to point out that—unlike in the case of scalar implicatures where a weak spoken utterance is privileged *because* the stronger option was not explicitly said—both readings generated by the *Every ... not* sentence are justified by what was said. This leads to the observation that the listener's adoption of the weaker of the two readings in the Metalinguistic Scalar account of (1) appears arbitrary. Second, an account that relies on the resulting *rejection* of the stronger of the two encoded meanings would amount to an inherent contradiction between what is said and one of its direct interpretations.³ Third, the idea that a listener would reject a strong interpretation when it is provided seems unlikely. Many have shown (Hirst & Weil 1982; Moore *et al.* 1990; Noveck *et al.* 1995, Chierchia *et al.* 2004) that children, at least, prefer stronger utterances over weak ones in the absence or presence of validating context. For example, Noveck *et al.* (1995) showed how five-year-olds prefer to follow the advice of the speaker who uses *has to* over another who uses *might* when there is no way to determine which is correct (compare *The peanut has to be under the cup* versus *the peanut might be under the box*) and Chierchia and colleagues have shown that four-year-olds prefer stronger descriptions over weak ones when both options are valid (for example they choose *all* over *some* to describe a scenario that would best be described with *all*). In short, it is not clear what would motivate a listener to adopt the weaker of the two readings as a first step and, furthermore, empirical tests show a general preference for stronger readings. Thus, there is little reason to assume that the Metalinguistic account can work to the extent that it is proposed as a linguistically driven mechanism that automatically adopts a *Not every* reading when the stronger *None* is available as well.

Regardless of the account that is ultimately supported in this growing literature, our aim here is modest, which is to simply validate the role played by pragmatics in interpreting sentences and to demonstrate that any explanation for the developmental effect cannot rely on syntax alone. We thus investigate the same sentences as Musolino *et al.* (2000) and Musolino & Lidz (2006) and, critically, we add a third group that is known for having pragmatic deficiencies: verbally competent autistic participants.⁴ We include such participants because their

³ Note that *rejecting* a reading differs from ordering readings in terms of preference.

⁴ We do not use the expression Asperger's because the French system does not use the same diagnostic criteria as found elsewhere. However, as will be seen by their responses to control problems, the autistic participants are verbal, high functioning teenagers.

syntactic skills are assumed to be relatively unimpaired (Tager-Flusberg 1981), while their pragmatic deficiencies have been well documented (e.g. see Dennis *et al.* 2001, Happé 1995). Not only are they challenged when confronted with metaphors, idioms, lies and jokes, the literature also shows that verbally competent autistic participants are not as efficient as typically developing controls at using context for more low level tasks like the disambiguation of sentences containing homographs (Happé 1997; Joliffe & Baron-Cohen, 1999). If utterance interpretation of *Every . . . not* sentences requires even a modicum of pragmatic processing, it follows from the previous research that sentences like (1) ought to be difficult for autistic participants, just as they are for young children.

Another issue that we address (indirectly) concerns methodology. In their first investigation, Musolino *et al.* (2000) reported that the *None* reading is preferred among children, even though the two readings are made available by the syntax. Their more recent work (Musolino & Lidz, 2006) went on to show that features of their (Truth Value Judgement or TVJ) task sway responses among children and that (at least part of) the children's preference for *None* readings in the Musolino *et al.* (2000) study is due to background information provided by the task. When a given negative context indicates that first *The horses ruled out jumping over a barn* before two of three jumped over a fence, five year olds' responses indicate that they prefer a *None* interpretation for the test sentence *Every horse did not jump over the fence* (i.e. over 80% say 'No' with respect to the test sentence, arguably because it is not the case that *None* jumped, two did). In contrast, when given affirmative prefacing context that indicates that first *Every horse jumped over a log* before two of three jumped over a fence, five year olds' responses indicate that they are more likely to prefer a *Not every* interpretation for *Every horse did not jump over the fence* (only 40% say 'No' with respect to the test sentence this time; 60% indicate 'Yes' and arguably because they agree that not all of the horses jumped over the fence, one did not).

It appears then that the children are sensitive to contrasting information. When the horses start out by failing, the children largely reject the notion that says the horses failed again (making them likely to adopt the *None* reading, which they then reject as false); when the horses start out by succeeding as a group, the children are more likely to detect that this time not all the horses succeeded (making them more likely here to adopt the *Not every* reading, which they accept as true). Adults, on the other hand, show a consistent preference for *Not every* readings regardless of the background information provided.

The present experiment, using more standard methodology, does not focus participants' attention in any way by presenting prior information. This might well bring about less biased results and potentially reveal the ambiguity among children and autistic participants that is assumed to be inherent in sentences like (1). It is expected that adults will continue to manifest a preference for a *Not every* reading.

2 EXPERIMENT


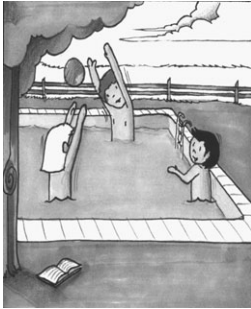
We present one experiment that investigates the way three different groups, 4-year-olds, autistic participants as well as adults, interpret the critical experimental item, as exemplified by (4):

(4) All the children are not in the pool.

In order to address this goal, we present stories in which ultimately 2-of-3 protagonists are shown doing something or in which 2-of-3 objects share a feature. For example, if a story were to concern the sentence above, the critical test sentence would concern the case where two of three children are in the pool. If the *None* reading is adopted, then participants ought to respond negatively but if the *Not every* reading is adopted, they ought to respond positively. Given that the sentence is ambiguous, one ought to expect children and autistic participants to prefer either (a) the initial parse (the *None* reading), or else; (b) to show evidence of being equivocal about the two interpretations. In any case, one should expect typical adults to prefer the *Not every* reading, which leads to a true response in this context.

Aside from items that measure comprehension, we also introduce two controls. To verify that the sentence itself is not the source of difficulty in this task, we present it in a context where both the *None* reading and the *Not every* reading yield the same (false) conclusion. This can be done for the sentence in (4) by providing a second context, one in which all (3-of-3) children are in the pool. One can safely conclude that the sentence itself is not problematic if participants are consistent in their treatment of (4) under such conditions. A second control is to present the same sentence while removing negations, whose presence may be a source of difficulty (see Prado & Noveck 2006). Evaluating a sentence such as *All the children are in the pool* successfully and in the two conditions mentioned earlier (where 2-of-3 v. where 3-of-3 are in the pool) would demonstrate that both children and verbally competent autistic participants can readily evaluate the quantifying aspect of these sentences. With respect to affirmative

Table 1 Truth value of a story's final slide as a function of the quantified statements (QN and QP) and context (2-of-3 and 3-of-3)

Statement	Presentation condition	
	2-of-3	3-of-3
		
QN: All the children are not in the pool.	?	F
QP: All the children are in the pool.	F	T

Notes. The ? indicates that this is the main test statement and that the truth value depends on the participant's reading. It would be true with a *Not > every* reading (*Not all the children are in the pool*) and false with an *Every > not* reading (*None of the children are in the pool*).

sentences like *All the children are in the pool*, participants ought to say false when two of three children are in the pool and to say true when all three are. These controls are not expected to cause difficulties for any of the participants and also serve to obscure the purpose of the task. Table 1 summarizes the task. Again, our main focus is to verify Musolino and colleagues' developmental effect with respect to sentences such as (1) in 2-of-3 contexts with our own materials and to determine the extent to which verbally competent autistic participants resemble young children.

2.1 Method

2.1.1 Participants Nineteen 4 year olds (average age: 4 years and 7 months) from a public school in Lyon, 20 adults recruited from the Université de Lyon, and 15 verbally competent autistic participants from a Lyonnais clinic (ITTAC) participated. The characteristics of the verbally competent autistic participants are summarized in Table 2. All the participants were French native speakers.

Table 2 Characteristics of the verbally competent autistic participants

Number of participants	Mean CA	Mean MA	IQ total	IQ verbal	Per cent passing Sally Ann task
15 (13 male; 2 female)	16;3	8;9	53.94	57.87	0.67

Notes. CA, chronological age; MA, mental age. Mental age was calculated from *Echelle d'Intelligence pour Enfant de Wechsler*, 3rd edition.

2.1.2 Materials Ten stories (in French) were prepared that were accompanied by drawings in colour by a professional artist following our specifications. Each of the stories concerned three people or three things and each was designed to have four standard frames completed by one of two possible outcomes. It follows that each story presented five frames to each participant (see Appendix for examples).

2.1.3 Test sentences In one condition (the 2-of-3 version) the final frame shows that two of the three people or objects make up a subgroup. For example, one story describes three friends who play together. The final frame in this condition concludes with two in the pool and one reading a book (see Table 1 and the Appendix). After each story, the experimenter presents a quantified statement to be evaluated, to which participants are required to say *True* or *False*. For example, for the Pool story in the Appendix, where the children are playing by a pool and then two of them decide to jump in, the critical experimental statement is (4), which we present in French, as (5).

(5) Tous les enfants ne sont pas dans la piscine.

A *Not every* reading of negation (*Not all children are in the pool*) in this 2-of-3 context ought to prompt a true response and a *None* reading (*None of the children are in the pool*) a false response. We now turn to the alternative context in which quantified statements were presented.

Stories were prepared so that they could appear in an alternative version where all three protagonists (or objects) follow through on something (the 3-of-3 condition). Thus, in the conclusion of the example above, all three children end up in the pool. In this case, the statement in (4) is false regardless of one's interpretation. Note that the syntax is obviously the same; only the context has changed. This control aims to verify that the sentence itself is not a source of difficulty. If the sentence itself is hard to process then it would be

difficult to provide consistently correct responses even in the 3-of-3 context.

All the stories were prepared so that they could be part of either condition. The stories were broken up into 4 groups so that a different random set of five stories represented the 2-of-3 condition in each with the remainder in the 3-of-3 condition. To obscure the purpose of the experiment, the above sort of statement was intermixed with positively quantified statements as well (see (7) below). For example, the above story could be followed by *All the children are in the pool* (which is false in the 2-of-3 context and true in the 3-of-3 context).

Finally, as a way to verify comprehension of each story, a quantified statement was always accompanied by a factual statement that also required an evaluation. One factual statement accompanied each story and it was permanently linked to a story.⁵ In summary, participants—across the task—were confronted with three *sorts* of statements as participants were presented each of the ten stories:

- (6) Quantified assertion preceding negation (QN):
All the children are not in the pool.
- (7) Quantified (positive) assertion without negation (QP):
All the children are in the pool.
- (8) Factual statement that could be true or false:
The children in the pool were playing with a ball.⁶

Participants always received either a QN statement or a QP statement along with a single factual statement. Overall, each participant received six QN statements (with half being in the 2-of-3 condition and half in the 3-of-3 condition) and four QP statements (again, with half being in each context condition) that were distributed among the ten stories so that we could do analyses based on participants as well as stories.

2.1.4 Procedure Each participant was engaged individually by the second author. She read a prepared text as she presented the story's pictures at sensible, prescribed times. Each frame was accompanied by one or two short sentences. The factual statement was the first test item presented and it was followed by one of the quantified statements.

⁵ Seven factual statements were designed to lead to a true evaluation and 3 to a false one.

⁶ This was the question for the story presented in Table 1. To appreciate a false statement, see Appendix.

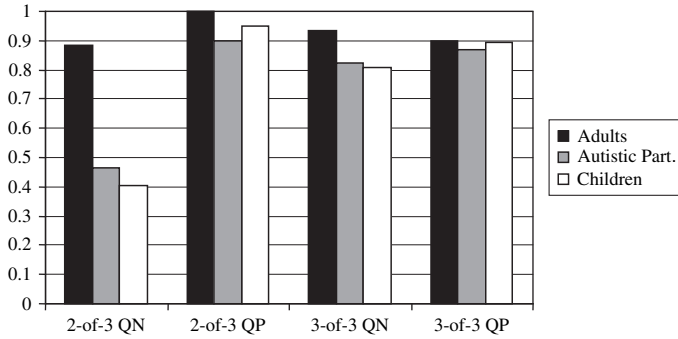


Figure 1 Mean percentage of correct responses among adults, autistic participants and 4-year-old children to quantified sentences that were affirmative (QP) or negative (QN) and across situations where two-of-three characters (or objects) were part of a completed action or where all three were. Adults' responses (reflecting a *Not every* reading) were considered correct in the 2-of-3 QN case.

2.2 Results

We analysed the results with respect to choice proportions (with a correct response to the critical QN question being the adult 'true' response) and we analysed the results both with respect to individuals and to stories.⁷ We first summarize the control items to demonstrate that all the participants generally understood the stories and the quantified statements before turning to the sentence type of interest, QN in a 2-of-3 context.

Rates of correct performance were high for all of the control problems. Adults', verbally competent autistic participants', and children's rates of correct responses to the factual statements were $M = 0.98$, $M = 0.96$, and $M = 0.86$, respectively.⁸ As can be seen in Figure 1, rates of correct responses to the QP statements were comparably high for adults, verbally competent autistic participants, and children ($M = 0.95$, $M = 0.88$, and $M = 0.92$ overall, respectively). Finally, rates of correct responses to the QN statement in the 3-of-3 condition were high and comparable among adults, verbally competent autistic participants and children ($M = 0.93$, $M = 0.82$, and $M = 0.81$, respectively). Figure 1 displays the mean percentage of

⁷ By convention we refer to F values obtained with participants as a random factor as F_1 (or t_1) and with items as a random factor as F_2 (or t_2).

⁸ Rates of correct responses to the factual statements were significantly higher among the adults than the children, $t_1(37) = 3.06$, $p < .005$, $t_2(18) = 3.61$, $p < .005$, but not significantly higher than among the autistic participants). The significant differences were due to the fact that adults' responses were at ceiling. Children's responses to factual statements were still well above levels predicted by chance.

correct responses to the two sorts of quantified statements across the two conditions.

In contrast, the QN statement in the 2-of-3 condition did not yield comparable rates of ‘true’ responses. As can be seen in Figure 1, rates of ‘true’ responses were higher for the adults than for either the children or the autistic participants ($M = 0.88$, $M = 0.45$, $M = 0.40$, respectively). A one-way ANOVA showed a significant main effect due to cohort group, $F_1(2, 51) = 13.199$, $p < 0.001$, $F_2(2,27) = 10.7$, $p < 0.001$. Adults were significantly more likely to respond *true* to this statement than the verbally competent autistic participants $t_1(33) = 3.96$, $p < 0.001$; $t_2(18) = 3.68$, $p < 0.005$ and the children, $t_1(37) = 5.87$, $p < 0.001$; $t_2(18) = 4.48$, $p < 0.001$. Children and verbally competent autistic participants were statistically indistinguishable and both were at chance levels (p 's > 0.5).

Slightly over a quarter of the children (5 of 19) consistently preferred the *None* reading (i.e. in all three trials) and slightly over 10% (2 of 19 children) consistently preferred the *Not every* reading, with the remainder choosing inconsistently. Six of 15 autistic participants (40%) consistently preferred the *None* reading and three (20%) consistently preferred the *Not every* reading, with the remainder choosing inconsistently. Sixty-five percent of the adults consistently chose the *Not every* reading and the remainder chose this reading in 2 (of 3) trials.

3 DISCUSSION

This experiment was designed to replicate Musolino *et al.*'s (2000) and Musolino and Lidz (2006) developmental effect with *Every ... not* sentences, and, by introducing verbally competent autistic participants to this paradigm, we aimed to verify that typical adult interpretations of such sentences rely on pragmatic processes. We did indeed replicate the effect among typically developing (children and adult) participants, but in a way that was slightly different than what was found in Musolino *et al.* (2000) and in Musolino and Lidz (2006). In the present study, children adopted either reading seemingly randomly (with a slight non-significant preference for *None* readings), reflecting the ambiguity of the sentence, and adults—as one would predict—preferred the *Not every* interpretation. Also, we showed that the verbally competent autistic participants and the children are very similar with respect to their responses. Responses to the control items indicate that neither the quantifier *all* nor the complex QN sentence is problematic for children and autistic participants and that they have a very high level of

comprehension overall. This shows that the QN statement presents a challenge for the non-adults uniquely in the 2-of-3 cases, and arguably because it is ambiguous. Overall, we take this as evidence in favour of the idea that a typical adult's *Not every* reading involves some amount of pragmatic processing. Below, we consider how to account for the pragmatic processing of these *Every . . . not* sentences by first reviewing the two less controversial accounts suggested by Musolino and colleagues, the Acquired Manner and Charity accounts, as well as our own Processing account. This is followed by the presentation of another hypothetical test case that can eventually serve to differentiate the Acquired Manner account from the other two.

According to the Acquired Manner account, the *Not every* reading of sentence (1) occurs practically by default because sophisticated speakers reject the *None* reading. If the speaker wanted the listener to treat, e.g. *All the children are not in the pool* as *None of the children of in the pool*, the speaker would have put it that way to start with. Of course, this account assumes that only the adults in the present study will have acquired this capacity. According to the Charity account, adults are more apt at adopting an interpretation that keeps the utterance true and, in fact, adults appear to do that. According to the Processing account we applied, the *Not every* reading in a 2-of-3 context is preferred over a *None* reading because the former entails less effort when integrating it into the context. That all three accounts should be consistent with the data is not surprising.

Here we entertain an experimental scenario which could test between the Acquired Manner account and the two others. It would involve presenting an *Every . . . not* test sentence in a context where no protagonist is doing what is described in the verb phrase (consider a case where 0 of 3 children can be found in a pool). Such a situation makes a *None* reading of the *Every . . . not* sentence (*Every child is not in the pool*) true semantically. In fact, it makes the *Not every* reading true as well (*None* entails *Not every*). However, according to the Acquired Manner account, adults routinely interpret *Every . . . not* sentences with a *Not every* reading (while presumably bypassing *None* readings). This reading (*Not every child is in the pool*) has the potential to prompt a false response in 0-of-3 contexts by way of a scalar implicature because *Not every* is under-informative when *None* is appropriate.⁹ The Acquired

⁹ Musolino and Lidz (2006, Experiment 2; see also Pouscoulous, Noveck, Politzer & Bastide, in press) have provided evidence showing that adults, unlike children, are more likely to reject statements that are literally presented as *Not every* in scenarios that express "None" and for the pragmatic reason just stated (*Not every* is underinformative with respect to a scenario describing *None*).

Manner account should predict a noticeable proportion of false responses (at least at rates comparable to those found when an utterance is literally expressed as *Not every*). The Charity account would predict that adults would adopt a reading that would make the utterance true, which is delivered by the *None* reading. Our own account would predict that the adults would prefer the *None* reading in this hypothetical 0-of-3 case and for two reasons. First, it is more immediately consistent with the presented context (here anticipating the rest of the sentence *in the pool* at the word *not* is consistent with a *None* reading). Second, even if participants were to consider a *Not every* reading, it does not automatically prompt a scalar implicature (see Noveck 2001, Bott & Noveck, 2004; Pouscoulous *et al.* in press).¹⁰

Choosing an appropriate reading for an *Every . . . not* sentence differs markedly from the interpretation procedure concerning scalars. The work on scalar terms has shown that the semantic interpretation of weak scalars (such as *some*) can suffice. It is only when a text or a task encourages participants to narrow the interpretation (of *some* to *some but not all*) that reading times or response times slow down on experimental tasks (Bott & Noveck 2004; Noveck & Posada 2003; Breheny *et al.* 2006). For example, when presented ‘underinformative’ items such as *Some cows are mammals*, those participants who answer true do so at a pace (roughly 800 msec) that is comparable to that of control items (including to false controls, such as *All mammals are cows*), but those who answer false to ‘underinformative’ items (indicating that they invoked a narrower meaning for *some*) take roughly 1400 msec; i.e. significantly longer than all the controls (Bott & Noveck, 2004, Experiment 1). Data from developmental studies and reaction time studies support the notion that pragmatic enrichment in scalar implicature tasks is optional or that it arrives as part of a more involved procedure. In the case of *Every . . . not* sentences, it is in the addressee’s interest to choose one reading or otherwise the sentence remains ambiguous. Unlike for the case of scalars, *Every . . . not* sentences require disambiguation and this depends on pragmatics.

One might ask why the children in Musolino and colleagues’ studies are not consistently equivocal (as a group) like the children in the current study. We suspect that the difference is due to experimental procedures. Whereas Musolino *et al.* (2000) presented stories in which three characters were first shown failing to do something (e.g. in the horse-jumping

¹⁰ Whereas the Charity account argues that pragmatics aims for true responses, our Processing account does not rule out a response indicating a *Not every* reading with an implicature (i.e. a false response) in a 0-of-3 context. Our account would only expect the more effortful interpretation to be less preferred. Further work would be required to distinguish these two accounts experimentally.

scenario, all three declined jumping over a barn before considering the fence instead), our task did not. In general, truth-value judgment tasks require its participants to integrate prior information in the form of base rates of success or failure from a previous event. Our task places no particular focus on what the three protagonists or objects did as a group prior to the ending. The outcomes concerning the three were largely incidental to the conclusion. Nevertheless, in all of these studies one finds that adults come up with a *Not every* reading more routinely than children.

To summarize, the results from this study confirm that adults rely on pragmatic processes to determine which of two readings is most applicable when hearing *Every . . . not* sentences. A contribution of this work is that it puts one in the position to extend the list of processing difficulties in autism to syntactic ambiguities. This modest claim is supported by the autistic participants' consistent performance with respect to the comprehension questions and the multiple unambiguous control sentences (making these participants comparable to both the children and adults) and their being inconsistent solely with respect to QN sentences in the 2-of-3 contexts (making them comparable solely to the children). It remains to be seen how to best describe the pragmatic abilities that distinguish healthy adults from both typically developing children and autistic participants. Nevertheless, this work confirms that there is much to be gained by viewing *Every . . . not* sentences through a pragmatic lens.

APPENDIX

Two examples stories, translated from French, each with a 2-of-3 and 3-of-3 ending, its factual item (test of comprehension) and quantifier test items.

Pool

- (1) It is vacation. Thomas, Eric and Sandrine are playing with a ball in the garden (the picture features a pool).
- (2) It is very hot in the sun, the kids are exhausted and they are perspiring a lot.
- (3) Thomas and Sandrine want to go swimming. They get undressed to go inside the pool. Eric does not feel like going swimming.
- (4) Thomas and Sandrine jump into the pool while Eric sits down next to a tree.
- (5) (2-of-3) It is a lot of fun to play with the ball in the pool. Eric prefers to read a book in the shade.
- (5) (3-of-3) It is a lot of fun to play with the ball in the pool. Eric decides to join them.

Fact: The children in the pool are playing with a ball. (True).

QN: All of the children are not in the pool (*Tous les enfants ne sont pas dans la piscine*).

QP: All of the children are in the pool (*Tous les enfants sont dans la piscine*).

Eggs

(1) Stephane spends his vacations at his grandmother's, who owns a farm.

(2) He spends his time playing with the hens and the chicks.

(3) He feeds them and cleans the chicken coop every morning.

(4) Three hens are in the process of brooding (*couver*), the three eggs should not take much longer before they hatch.

(5) (2-of-3) The next day, while cleaning the coop, Stephane notices that two new chicks were born. There is still one more egg.

(5) (3-of-3) The next day, while cleaning the coop, Stephane notices that the three new chicks were born.

Fact: The child is wearing (long) pants (False; he is wearing shorts).

QN: All the chicks are not born (*Tous les poussins ne sont pas nés*).

QP: All the chicks are born (*Tous les poussins sont nés*).

Acknowledgements

The authors wish to thank Robyn Carston, Bart Geurts, Julien Musolino, Anne Reboul and three anonymous reviewers for their comments with respect to an earlier version of this paper, Jeffrey Lidz and Sandra Waxman for conversations during this work, as well as the team RD&P in Lyon (Jean-Baptiste Van der Henst, Mathilde Bonnefond, Coralie Chevallier, and Jérôme Prado) for their helpful suggestions on all phases of the work. We also wish to thank those who participated in the experiment as well as the ITTAC in Villeurbanne for their collaboration. This work was supported by a grant from the *Fondation de France* awarded to the first and third authors and coordinated by Marc Jeannerod. Correspondence concerning this article should be addressed to Ira Noveck, L2C2, Institut des Sciences Cognitives, 67 Blvd. Pinel, 69675 Bron, France (email: noveck@isc.cnrs.fr).

IRA A. NOVECK

*Laboratoire sur le Langage
le Cerveau et la Cognition
Université de Lyon 1 and CNRS*

RAPHAËLE GUELMINGER

*Laboratoire sur le Langage
le Cerveau et la Cognition
Université de Lyon 1 and CNRS*

NICOLAS GEORGIEFF

*Centre de Neurosciences Cognitives
Université de Lyon 1 and CNRS*

NELLY LABRUYERE

*Centre de Neurosciences Cognitives
Université de Lyon 1 and CNRS*

REFERENCES

- Atlas, J. (1989) *Philosophy Without Ambiguity*. Clarendon Press. Oxford.
- 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.
- Carston, R. (2002) *Thoughts and Utterances*. Blackwell. Oxford.
- Chierchia, G., Guasti, T., Gualmini, A., Meroni, L. & Crain, S. (2004) 'Adult's and children's semantic and pragmatic competence in interaction'. In I.A. Noveck and D. Sperber (eds), *Experimental Pragmatics*. Palgrave Macmillan. Basingstoke.
- Dennis, M., Lazenby, A.L. & Lockyear, L. (2001) 'Inferential language in high-function children with autism'. *Journal of Autism and Developmental Disorders* **31**(1):47–54.
- Happé, F. (1995) 'Understanding minds and metaphors: Insights from the study of figurative language in autism'. *Metaphor and Symbol* **10**(4):275–295.
- Happé, F. (1997) 'Central coherence and theory of mind in autism: Reading homographs in contexts'. *British Journal of Developmental Psychology* **15**:512–521.
- Hirst, W. & Weil, J. (1982) 'Acquisition of the epistemic and deontic meaning of modals'. *Journal of Child Language* **9**:659–666.
- Horn, L. R. (1984) 'Toward a new taxonomy for scalar inference'. In D. Schiffrin (ed), *GURT*. Georgetown University Press. Washington DC.
- Joliffe, T. & Baron-Cohen, S. (1999) 'A test of central coherence theory: Linguistic processing in high-functioning adults with autism or Asperger syndrome: Is local coherence impaired?' *Cognition* **71**:149–185.
- Levinson, S. (2000) *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. MIT Press. Cambridge, MA.
- Lidz, J. & Musolino, J. (2002) 'Children's command of quantification'. *Cognition* **84**(2):113–154.
- Moore, C., Pure, K. & Furrow, D. (1990) 'Children's understanding of the modal expression of speaker certainty and uncertainty and its relation to the development of a representational theory of mind'. *Child Development* **61**:722–730.
- Musolino, J., Crain, S. & Thornton, R. (2000) 'Navigating negative quantificational space'. *Linguistics* **38**(1): 1–32.
- Musolino, J. & Lidz, J. (2003) 'The scope of isomorphism: Turning adults into children'. *Language Acquisition* **11**(4):277–291.
- Musolino, J. & Lidz, J. (2006) 'Why children are not universally successful with quantification'. *Linguistics* **44**(4):817–852.
- Noveck I. A. (2001) 'When children are more logical than adults: Experimental

- investigations of scalar implicature'. *Cognition* **78**(2):165–188.
- Noveck, I. A. (2004) 'Pragmatic inferences related to logical terms'. In I.A. Noveck and D. Sperber (eds), *Experimental Pragmatics*. Palgrave. Basingstoke. 301–322.
- Noveck, I. A, Ho, S. & Sera, M. (1996) 'Children's understanding of epistemic modals'. *Journal of Child Language* **23**(3):621–643.
- Noveck, I. A. & Sperber, D. (in press). 'The why and how of experimental pragmatics: The case of "scalar inferences" '. In N. Burton-Roberts (ed), *Advances in Pragmatics*. Palgrave. Basingstoke.
- Papafragou, A. & Musolino, J. (2003) 'Scalar implicatures: experiments at the semantics-pragmatics interface'. *Cognition* **86**(3):253–282.
- Pouscoulous, N., Noveck, I. A., Politzer, G. & Bastide, A. (in press). 'Processing costs and their impact on the development of scalar implicature'. *Language Acquisition*.
- Prado, J. & Noveck, I. A. (2006) 'How reaction times can elucidate matching effects and the processing of negation'. *Thinking and Reasoning* **12**(3):309–328.
- Sperber, D. & Wilson, D. (1995) *Relevance: Communication and Cognition*. Basil Blackwell. Oxford.
- Tager-Flusberg, H. (1981) 'On the nature of linguistic functioning in early infantile autism'. *Journal of Autism and Developmental Disorders* **11**:45–56.

Submitted: 15.05.06

Accepted: 01.08.06