Understanding irony in autism: The role of context and prosody

Iris Scholten, Eerin Engelen & Petra Hendriks
University of Groningen

1 Introduction

Irony is a figure of speech that can be used to express the opposite of what is literally said. For example, the sentence “That was fun!” implies that the speaker had a great time, but when intended ironically (such as after a boring party) the same sentence expresses the exact opposite. The point of irony is to indicate that a proposition that a speaker may normally endorse is in fact not endorsed by the speaker, for example because it is false or might be unlikely given the situation (Wilson & Sperber, 2012). To make sure the hearer will understand the ironic intention of the speaker, speakers can use cues to get their intention across. For example, they can use a specific facial expression or body language. In this study, we will focus on two linguistic cues: context and prosody.

The context in which something is said is considered to be one of the most important cues for the recognition of irony (Jorgensen et al., 1984; Capelli et al., 1990; Giora, 1997). Context can set up a particular expectation, which is in conflict with the content of the ironic statement. This conflict can help the hearer to recognize the ironic intention of the statement. As an example, consider the following situation: John is on holiday. He discovers that his bags have gotten lost at the airport and the hotel he booked is full. When he says to Mary: “This must be my lucky day!,” this utterance is so obviously in conflict with the context that Mary should normally be able to recognize the irony.

A second important cue is the prosody used in the ironically intended expression. According to Bryant (2010), there is a typical prosody that implies the intention of irony and therefore could invite the recognition of irony by hearers. This typical prosody in ironic expressions involves two high peaks: one peak around the second word (usually the verb, in languages such as English and Dutch) and one peak at the end of the expression. There may be other prosodic features that are associated with an ironic intention, such as an exaggerated monotone intonation or overly enthusiastic exclamations, but these features might be harder to recognize (Winner, 1988).
An important factor in the use of irony is the ambiguity of the ironically intended expression. Even though context and prosody can provide very clear cues for the ironic interpretation, it is still possible that the hearer does not pick up on the ironic intention and instead interprets the expression literally. The clearer the discrepancy between the speaker’s description of the situation and the actual state of affairs, the easier it is to recognize the irony (Wilson & Sperber, 2012). The same may be true for prosody: when the intonational pattern described above is used, it might become easier to recognize the ironic intention. If prosody is indeed a clear indicator for irony, context might not even be needed to recognize irony and prosody may be enough to understand the speaker’s intention.

Understanding irony is a skill that seems particularly difficult for individuals with Autism Spectrum Disorder (ASD). ASD is a congenital neural developmental disorder that is characterized by qualitative deficits in social interaction and communication and by limited, repetitive or stereotypical behaviors, interests or activity patterns (APA, 2000), ranging from mild to severe (APA, 2013). It is claimed that individuals with ASD do not have difficulties with language per se, but rather with the pragmatic functions of language (Baron-Cohen, 1997; Happé, 1997; Norbury & Bishop, 2002; Martin & McDonald, 2004). As a consequence, they may have problems in taking advantage of the contextual cues that indicate indirect, figurative or ironic language use. Furthermore, they may have difficulties grasping the suprasegmental aspects of language, such as prosody, rhythm and accents (Rutter et al., 1992). These aspects of language are very important for the understanding of irony, since contextual and intonational cues contribute to the recognition of irony. A deficit in understanding these cues may therefore lead to problems in recognizing ironic intentions.

Various studies have shown difficulties in the understanding of irony in children and adults with ASD in comparison to typically developing peers (Kaland et al., 2002; Adachi et al., 2004; MacKay & Shaw, 2004; Martin & McDonald, 2004; Wang et al., 2006). According to Kaland et al. (2002), children with ASD are unable to recognize the ironic intention of the speaker when being asked about why someone says something. MacKay and Shaw (2004) found that children with ASD, when being asked for someone’s true reason for saying something like “Great job!”, often give a literal meaning or merely rephrase the expression produced.
Several explanations have been proposed for the difficulties children and adults with ASD have in recognizing irony. For example, these difficulties may be due to a deficit in Theory of Mind (Baron-Cohen, Leslie & Firth, 1985). Theory of Mind (ToM) is the ability to understand and predict behavior based on one’s own beliefs and the beliefs of others (Wimmer 1983; Perner & Wimmer, 1985). To understand irony, it is important that the hearer is able to apply higher-order ToM reasoning. First-order ToM reasoning (ToM-1) is the ability to attribute beliefs, thoughts and desires to someone else and to assume that these beliefs, thoughts and desires influence this person’s behavior. For example, John is able to apply first-order ToM reasoning if he understands that Mary utters the sentence “That was fun!” when talking about a party because Mary believes the party was great. Higher-order reasoning (ToM-2 and further) involves the beliefs someone else has about another person and their predictions about this other person. For example, Mary is able to apply higher-order ToM reasoning if she understands that John believes that she believes the party was great. Thus, ToM-2 is needed to be aware of the fact that someone else has beliefs about you (Verbrugge & Mol, 2008). Crucially, the beliefs of this other person might be different from your own beliefs. That is, while John may believe that Mary believes the party was great, Mary might in fact have found the party quite boring. Understanding irony requires the ability to apply higher-order ToM reasoning because it requires understanding a thought about an attributed thought (Happé, 1993). When Mary ironically says “That was fun!”, she mentions a thought and at the same time expresses her attitude towards this thought. Understanding the irony in this utterance requires that the listener not only understands the thought, but also understands the speaker’s attitude towards this thought.

Studies that tested children on a ToM-2 task and a separate irony task confirm the suggestion that the understanding of an ironic intention and ToM-2 reasoning are closely related (Winner & Leekam, 1991; Happé, 1993; Filippova & Astington, 2008): children who fail on a ToM-2 task are also less capable of understanding ironic expressions. It is well-established that children with ASD have more difficulty with ToM reasoning than typically developing children (e.g., Baron-Cohen et al., 1985; Happé, 1993, 1994; Baron-Cohen, 1995, 1997, Filippova, & Astington, 2008; Massaro et al., 2012). Therefore, their poor performance in understanding irony could be caused by their difficulties with ToM. Alternatively, their difficulties in
understanding irony may also be due to their problems in understanding the cues for irony.

In contrast to children’s understanding of irony, not much is known about adolescents’ understanding of irony. Adolescents with ASD are expected to be linguistically more advanced than children with ASD. Does this mean that they are fully capable of using linguistic cues such as context and prosody to recognize the ironic intention of the speaker?

The present study aims to investigate whether adolescents with ASD are able to recognize and understand irony in the same way as their typically developing peers. In particular, we wish to find out whether they use the linguistic cues of context and prosody in the same way. To this end, we carried out an irony recognition task with a group of Dutch-speaking adolescents with ASD and a control group of typically developing adolescents, in which we manipulated context and prosody.

2 Methods

2.1 Participants
Thirteen adolescents with Autism Spectrum Disorders (mean age 15.5, age range 14–20, 10 male) were recruited from Scholengemeenschap De Ambelt in Zwolle (a school for secondary special education, cluster 4). The inclusion criteria for the sample were based on parental information about the clinical diagnosis that was confirmed by the participants. One additional adolescent was tested but later excluded from the analysis because of lack of confirmation of the clinical diagnosis. There were no participants with a double diagnosis, such as the combination of ASD and ADHD. The control group consisted of fourteen typically developing adolescents (mean age 14.4, age range 11–20, 5 male); twelve of these adolescents were recruited through Scouting Group Don Bosco in Geldrop and two others were recruited through the researchers’ personal network. All parents and/or caretakers of the participants gave written informed consent for their participation in this study.

2.2 Design and materials
The experiment manipulated two factors: context (inviting an ironic versus a non-ironic interpretation) and prosody (inviting an ironic versus a non-ironic interpretation), yielding a design with four conditions (see Table 1). There were eight
items per condition, resulting in 32 items in total. From these items, four counterbalanced randomized lists were constructed. Each participant heard 16 items in total and 4 items per condition.

The 32 items consisted of short stories followed by three test questions. Each story started with an introductory sentence, followed by a concrete event that further specified the situation, and concluded with an evaluative statement. The experiment tested these items in four conditions: the neutral condition, the prosody condition, the context condition, and the combination condition. In the neutral condition, neither the story context nor the prosody of the evaluative statement invited an ironic interpretation of the evaluative statement. In the prosody condition, the evaluative statement had a prosody that invited an ironic interpretation, while the story context was compatible with a non-ironic interpretation of the evaluative statement. In the context condition, the story context invited an ironic interpretation of the evaluative statement, while the prosody of the evaluative statement was neutral. In the combination condition, finally, both the story context and the prosody of the evaluative statement invited an ironic interpretation of the evaluative statement.

Table 1. Design of the experiment, with a sample item for each condition (translated from Dutch)

<table>
<thead>
<tr>
<th>– Prosody</th>
<th>+ Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>– Context</td>
<td>+ Context</td>
</tr>
<tr>
<td>John’s long-time wish is to get a scooter. Today is his birthday. When he enters the garage, he sees a brand new, shiny scooter. He says to his parents: “What a great gift.”</td>
<td>Sara has a job interview at the local grocery store today. She feels relaxed and responds to the questions very well. She gets hired. When she comes home, she says to her father: “It went very badly.”</td>
</tr>
<tr>
<td>Tim spends the entire summer working in a clothing store. The customers constantly muddle the clothes and there is never a moment of relaxation. When his girlfriend stops by he says to her: “This is the worst job ever.”</td>
<td>Peter promises his wife to clean the house. When he tries to dust the mantelpiece he accidentally knocks over the favorite vase of his wife. It shatters into a thousand pieces as it hits the floor. His wife hears the noise and says: “Great, well done.”</td>
</tr>
</tbody>
</table>
Prosody was manipulated by distinguishing between two patterns of pronunciation for the evaluative statement at the end of the stories. In one pattern, the sentence was uttered with a typical ironic intonation in which there were two high peaks: one around the second word and one at the end of the sentence (see Bryant, 2010). In the other pattern, the sentence was uttered with neutral intonation. The intonation patterns were verified using Praat, a computer software package to analyse speech, and were pre-tested with seven typically developing adults, who listened to the statements out of context and rated these statements on a five-point-scale for level of irony. Statements that were rated as not clearly ironic or not clearly non-ironic (with average ratings between 1 and 4) were recorded again with a more distinct pronunciation.

Context was manipulated by distinguishing between stories that are consistent with the subsequent evaluative statement (e.g., the story at the top left in Table 1) and stories that are inconsistent with the evaluative statement (e.g., the story at the top right in Table 1, in which the positive expectation of the story is inconsistent with the negative value of the evaluative statement). The first type of story invites an ironic interpretation of the evaluative statement, whereas the second type of story does not. We pre-tested these stories with six typically developing adult participants, who listened to the stories without evaluative statements and were asked to indicate which emoticon matched best with the main character of the story. Stories for which not at least five out of the six participants chose the target emoticon were adapted. We thus made sure that the stories in our experiment were all unambiguously interpreted as either positive or negative.

All stories were followed by three questions: a question about the emotion of the main character in the story, a first-order ToM question (ToM-1) about the emotion of the main character and a second-order ToM question (ToM-2) about the belief of the secondary character about the emotion of the main character. For example, after the story at the bottom left in Table 1 about Tim, the following three questions were asked:

(1) Which emoticon do you think matches best with Tim?
(2) Do you think Tim thinks this is the worst job ever?
(3) Does Tim’s girlfriend think that Tim thinks this is the worst job ever?
Participants were instructed to answer the first question by pointing to one of four emoticons, which were presented on a piece of paper. These emoticons were selected on the basis of a pre-test: an online questionnaire. In this pre-test, the respondents (n = 93, all different from the participants in the present study) were presented with three emoticons and one emotion and were instructed to select the emoticon that they thought represented the emotion most accurately. A sample question (translated from Dutch) was: Which emoticon expresses the emotion ANGRY best, according to you? The participant had to choose between three emoticons from the Emoji of smartphones that can be used to indicate the emotion mentioned, in this case angry. For the present study, the four emoticons were chosen that - according to the results of the pre-test - best represented the four emotions used in the test: happy, angry, scared and sad (see Figure 1 for the black and white versions of the colored emoticons).

![Emoticons](image-url)

**Figure 1.** Emoticons that were used in the test, from left to right: happy, angry, scared and sad.

The answer to the two ToM-questions, illustrated by (2) and (3) above, could be yes or no. Participants’ responses to each of the three question types (i.e., the choice of emoticon and the yes/no answers) were scored as either ironic or not ironic. For example, if the participant chose the sad, angry or scared emoticon in response to the statement “This is the worst job ever” in the context in Table 1, that would be scored as not ironic. On the other hand, if the participant chose the happy emoticon, that would be scored as ironic. Likewise, if the participant answered yes on the ToM-1 question in (2) or the ToM-2 question in (3), that would be scored as not ironic, and if they answered no, that would be scored as an ironic response. The responses per condition and per question type were analysed separately. This resulted in a mixed design with 12 variables. Performance on these variables was based on 4 items each and was converted into percentages of ironic responses.
In typically developing adolescents, we expect the combination condition to lead to more ironic responses than the other three conditions, because two cues that are important for recognizing and interpreting irony are present in this condition. If adolescents with ASD ignore contextual cues, prosodic cues, or both, when listening to utterances that are intended ironically, as is suggested by the literature, we expect them to give fewer ironic responses in the condition employing these cues than their typically developing peers. For both groups, least ironic responses are expected on the neutral condition, because this condition provides no cues for an ironic interpretation. As context is considered a stronger cue than prosody, both groups are also expected to give more ironic responses in the context condition than in the prosody condition. If context is a prerequisite for an ironic interpretation, the prosody condition may in fact not invite any ironic responses at all.

2.3 Procedure

All stories were recorded using Adobe Audition and played during test sessions using iTunes on a laptop with speakers. The participants listened to 2 practice stories and 16 experimental stories in a quiet room; the students of De Ambelt were tested in a room at school and the scouting youth was tested in a room in the scouting building. Two researchers were present during the test sessions. One of the researchers made notes on the scoring forms and operated the laptop and the voice recorder. All sessions were recorded on a voice recorder. We started with a pre-test assessing whether the participants were familiar with the emoticons used in the test. Next, participants listened to the pre-recorded stories while looking at pictures. The pictures did not display any of the characters in the stories, that could be associated with emotions, but merely showed emotion-neutral objects mentioned in the story to help the participants focus on the task. Test sessions took approximately 12 minutes.

3 Results

Table 2 and Table 3 list the means and standard deviations on the four conditions and three question types for adolescents with ASD and the control group of adolescents without ASD.
Table 2. Mean percentages of ironic responses and standard deviations for the ASD group (n=13) per condition (Neutral, Prosody, Context, Combination) and question type (Emotion, ToM-1, ToM-2).

<table>
<thead>
<tr>
<th>Condition</th>
<th>Emotion</th>
<th>ToM-1</th>
<th>ToM-2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Neutral</td>
<td>0</td>
<td>0</td>
<td>3.9</td>
</tr>
<tr>
<td>Prosody</td>
<td>17.3</td>
<td>23.7</td>
<td>34.6</td>
</tr>
<tr>
<td>Context</td>
<td>75.0</td>
<td>17.7</td>
<td>76.9</td>
</tr>
<tr>
<td>Combination</td>
<td>57.7</td>
<td>15.8</td>
<td>65.4</td>
</tr>
</tbody>
</table>

Table 3. Mean percentages of ironic responses and standard deviations for the control group (n=14) per condition (Neutral, Prosody, Context, Combination) and question type (Emotion, ToM-1, ToM-2).

<table>
<thead>
<tr>
<th>Condition</th>
<th>Emotion</th>
<th>ToM-1</th>
<th>ToM-2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Neutral</td>
<td>0</td>
<td>0</td>
<td>3.6</td>
</tr>
<tr>
<td>Prosody</td>
<td>17.9</td>
<td>22.8</td>
<td>26.8</td>
</tr>
<tr>
<td>Context</td>
<td>80.4</td>
<td>17.5</td>
<td>82.1</td>
</tr>
<tr>
<td>Combination</td>
<td>80.4</td>
<td>14.4</td>
<td>87.5</td>
</tr>
</tbody>
</table>

The results are also shown graphically in Figure 2 and Figure 3 below.
Figure 2. Mean percentages of ironic responses and standard deviations for the ASD group (n=13) per condition (Neutral, Prosody, Context, Combination) and question type (Emotion, ToM-1, ToM-2).

Figure 3. Mean percentages of ironic responses and standard deviations for the control group (n=14) per condition (Neutral, Prosody, Context, Combination) and question type (Emotion, ToM-1, ToM-2).
A mixed ANOVA was performed with Group (ASD, control) as the between-subjects factor, and Condition (neutral, prosody, context, combination) and Question Type (emotion, ToM-1, ToM-2) as within-subjects factors. There were significant main effects for Condition (F(3,75) = 186, p < .001, η² = .881) and Question Type (F(2, 50) = 5.81, p = .005, η² = .188) on the mean percentages of ironic responses. There were significant interactions between Condition and Question Type (F(3, 94.9) = 6.15, p < .001, η² = .197) (Greenhouse-Geisser corrected) and between Condition and Group (F(3, 75) = 3.45, p = .021, η² = .121). Both groups interpreted utterances that were accompanied by contextual cues or a combination of contextual and prosodic cues as more ironic than utterances that were not accompanied by any of these cues or were accompanied by prosodic cues only. Furthermore, for all four conditions, both groups had more ironic responses on the ToM-1 questions than on the ToM-2 questions.

To further inspect the interaction effect of Group with Condition, four one-way MANOVAs (Field, 2009) were run separately for each condition with Group as the fixed factor and the three question types as dependent variables. There was a significant difference in ironic responses based on the participant’s diagnosis for the combination condition (F(3, 23) = 4.69, p = .011; Wilk’s Λ = .621, η² = .379). The group with ASD gave significantly less ironic responses on the combination condition than the control group. Follow-up tests on the Combination Condition furthermore revealed that Group had a significant effect on the emotion question (F(1, 25) = 15.18, p = .001, η² = .378) (Bonferroni corrected). Post hoc tests revealed that the group with ASD gave significantly less ironic responses on the emotion question in the combination condition (M = 57.7, SD = 15.8) than the control group (M = 80.4, SD = 14.5). From this we can conclude that for adolescent with ASD the presence of both contextual and prosodic cues leads to significantly less ironic interpretations than for adolescent without ASD, especially on the emotion question.

4 Discussion

In this study, we investigated whether adolescents with ASD have difficulty understanding irony. Hypothesizing that individuals with ASD are less capable of recognizing and interpreting irony than their typically developing peers, we furthermore wanted to find out in what way their recognition of irony depends on
linguistic factors. To investigate this, we compared adolescents with ASD and typically developing adolescents on their interpretation of short stories in which prosodic and contextual cues for irony were manipulated.

If young individuals with ASD have difficulty understanding irony, we expect them to recognize the ironic intention in our stories less well than their typically developing peers. We found that, overall, the adolescents with ASD did not recognize the ironic intention less often than their typically developing peers. However, they did so when the ironic intention was indicated by both prosody and context. In that case, they gave fewer ironic responses on the emotion question than typically developing adolescents. Thus, adolescents with ASD have more difficulty than their typically developing peers to recognize the ironic intention of a statement that has an ironic prosody and at the same time is preceded by a context that is inconsistent with the positive or negative value of the statement.

Could the observed lower performance by the adolescents with ASD be attributed to their suboptimal use of prosodic or contextual cues? Both the adolescents with ASD and their typically developing peers interpreted stories in which the only cue to the speaker’s ironic intention was the prosodic structure of the sentence differently than they did stories without any cues. The presence of prosodic cues led to more ironic interpretations than the absence of any linguistic cues. Although we did not expect the prosody of a sentence alone to lead to an ironic interpretation, we found that prosody can invite an ironic interpretation. Furthermore, we found that adolescents with ASD, like their typically developing peers, use such prosodic cues in their interpretation of irony.

Also the presence of contextual cues was found to lead to more ironic interpretations compared to when there were no cues, both for adolescents with ASD as for typically developing adolescents. In both groups, the percentage of ironic interpretations was much larger when irony was signalled by context than when it was signalled by prosody. In fact, for both groups the combination of prosodic and contextual cues did not lead to more ironic interpretations than the presence of only contextual cues. This suggests that for adolescents with ASD as well as for adolescents without ASD the most important cue for recognizing irony is context. Thus, the difference between adolescents with ASD and their typically developing peers in the recognition of an ironic intention of a statement does not seem to be due to their insensitivity to prosody or context, which are the two most important
linguistic cues for irony. Despite their sensitivity to prosodic and contextual cues for irony, it is possible that adolescents with ASD are less efficient in using these cues or perhaps have difficulty integrating two different cues.

Even in the conditions with the highest percentages of ironic interpretations, the adolescents’ ironic interpretations generally did not rise above 80%. An exception are the responses by the typically developing adolescents on the ToM-1 questions in the context condition and the combination condition. Our study did not include an adult group, so we cannot be certain whether adolescents in general are not adult-like yet in their recognition of irony and their use of linguistic cues, or whether their performance with irony is adult-like. However, it is quite likely that adults are not perfect in their recognition of the ironic intentions of a speaker either.

Adolescents with ASD gave fewer ironic responses than their typically developing peers on emotion questions, but not on ToM-1 or ToM-2 questions. This does not mean that adolescents with ASD have no difficulty with ToM reasoning. In our study, the responses to the two ToM questions are dependent on the response on the emotion question. Therefore, our study did not test participants’ ToM reasoning independently of their recognition of irony. To further investigate the relation between the recognition of irony and ToM reasoning, participants should be tested on an irony task as well as a separate ToM task. We leave this for further research.

Acknowledgments

We thank Scholengemeenschap De Ambelt in Zwolle and Scouting Group Don Bosco in Geldrop for their hospitality. In particular, we greatly appreciate the help of René Reith, Rianne Harms and the other teachers of De Ambelt and Kirsten van der Geer of the scouting group. Furthermore, we are grateful to all participants and parents for their cooperation in this study. Laurie Stowe and Wander Lowie are thanked for their advice on statistical issues, and two anonymous reviewers for their useful comments.

References

differentiate high functioning pervasive developmental disorder from attention deficit/hyperactivity disorder. *Brain and Development*, 26(5), 301-306.


