The development of irony comprehension and epistemic vigilance

Introduction. Irony (e.g., uttering "You're so kind" to criticise someone who has not provided the expected support) has been found to be a relatively late acquisition: several studies suggest that children only start grasping ironical utterances from around the age of 6. However, some studies have suggested that a sensitivity to some aspects of irony (e.g., speaker's using a characteristic tone of voice/facial expression) may arise earlier than this (AuthorX and AuthorY, 2021). The current experiment takes as its starting point the relevance-theoretic *echoic* account of verbal irony (Wilson & Sperber, 2012) and addresses the development of the recognition of irony based on the sensitivity to the inappropriateness of what is said. The aim was to investigate whether children's epistemic vigilance towards utterance content (Sperber et al. 2010) might help them detect the ironical speaker's dissociative attitude towards the proposition literally expressed by her utterance. An utterance echoing a thought that is very inappropriate is likely to be recognized as ironic because (1) it is more likely to be the target of a dissociative attitude, and (2) because of epistemic vigilance mechanisms, the utterance is less likely to be interpreted literally since one is unlikely to assume that the speaker intends her audience to accept such an inappropriate thought. Furthermore, since children are found to often provide literal interpretations of ironical utterances in experimental settings (AuthorX & Author Y, 2020), we hypothesised that the absence of a literal option - and using videos which revealed the ironical speaker's facial expression - would improve children's irony comprehension.

Design. We designed a novel irony task which does not require a verbal response. Participants saw short movies involving a young woman and a hand puppet (see Figure 1). The hand puppet has to guess what is on a picture (e.g., "a completely full glass") and the woman either praises his guess or mocks it ironically. The sentences in both the literal and the ironical condition are identical with respect to wording (e.g., "'Yes, this is really a completely full glass"), but they differ in the speaker's attitude: sincerely happy versus ironical intonation and facial expressions. Based on these audiovisual cues, the participants had to choose which of two images (e.g., a full or an empty glass) is depicted on the card in the speaker's hand. The two pictures represent different points of a scale: a literal option (e.g., a full glass), a medium option (a half full glass) or an extreme option (an empty glass). All participants watched 12 videos varying in utterance type (literal, ironic), picture combination (literal-extreme, literal-medium, medium-extreme) and scale (e.g., full-empty, happy-sad, hard-soft). We measured picture choice and eye gaze while the sentence unfolds.

Figure 1: Screenshot from a video stimulus

Participants. Eighty-nine Norwegian-speaking children between 3 and 7 years of age (range: 3.08 to 7.33 years; mean age: 5.12 years; 41 female) and a control group of 20 adults (range: 20 to 53 years; mean age: 28.5; 16 female) participated in the study.

Results
Picture-selection. The accuracy of picture-selection in children was 89 percent for literal utterances and 12 percent for ironical utterances; for adults it was 97 for literal utterances and 85 percent for ironical utterances. Put differently, children selected the picture most closely aligned with a literal interpretation, regardless of whether a literal or ironical utterance was presented. To give an example, when hearing the sentence "Yes, this is really a completely full glass", children overwhelmingly chose the picture depicting a full glass – and in case this literal picture option was not available, they picked the half full glass over the empty glass.
We analyzed children's picture-selection data with mixed effects logistic regression in R (version 4.1.2.), using the binary response accuracy of picture selection as a dependent variable. The model includes random intercepts for subjects and items, as well as age, condition (irony, literal) and picture combination (literal-extreme, literal-medium, medium-extreme) as fixed effects factors. Children selected the correct picture more often in the literal condition than the irony condition ($\beta = 4.482$, $z = 18.136$, $p < .001$). Age was weakly significant ($\beta = 0.211$, $z = 1.972$, $p = 0.049$), mostly driven by the fact that older children tended to be more accurate than younger children in the literal condition, albeit not the irony condition. The type of picture combination did not affect children's accuracy of picture selection.

Gaze. We analysed the proportion of looks while the utterance unfolds (plus 300 ms afterwards) to four areas of interest: the two pictures as well as the faces of speaker and addressee. Both children and adults preferred to look at the picture closest to a literal interpretation, regardless of utterance type (irony, literal). However, when comparing the looks to the two pictures in both conditions, calculating a difference score, children's preference for the picture that is most in line with a literal interpretation turned out to be stronger in the literal condition than in the irony condition ($p = .006$), similar to adults. Furthermore, children looked more at the speaker in the irony condition compared to the literal condition (0.54 vs. 0.45, $p = .020$).

Discussion and conclusion
The offline data from the picture-selection task show no evidence of irony understanding in children aged 3 to 7 and, surprisingly, no improvement of irony understanding with age. With just 12 percent correct picture choice in the irony condition, children were substantially below the chance level of 50 percent, showing a strong bias to interpret ironical utterances literally. This was the case even when the ironical utterance was echoing a thought that was very inappropriate in the context, suggesting little help from epistemic vigilance mechanisms. Removing the picture representing the literal interpretation did not improve children's irony comprehension, as they tended to pick the picture closest to the literal option on the scale. By contrast, the gaze data reveal that children looked less at the "literal" picture in the ironical condition compared to the literal condition. Furthermore, children pay special attention to the facial expressions of an ironical speaker. Both findings could indicate a sensitivity for irony, not captured by the offline results. A possible explanation for the observed divergence between offline and online measures could be the high demands of the picture-selection task, requiring children to infer the state of the world solely based on the speaker's tone of voice and facial expressions. The fact that in standard narrative paradigms the majority of children is able to understand irony by the age of 6 (e.g., AuthorX and AuthorY, 2021), our goal to create a simple and sensitive irony task was not successful. However, with our new methodology we were among the first to study the role of facial expressions in children's interpretation of irony, something further studies should explore in more detail.

References. AuthorX and AuthorY, 2020, 2021; Sperber, D. et al. Epistemic vigilance. *Mind & Language, 25*(4), 359–393; Wilson, D. & Sperber, D. 2012. Explaining relevance. *Meaning and Relevance,* Cambridge UP 2012.