How do 3- and 5-year-olds respond to under- and over-informative utterances?

Tiffany Morisseau, Catherine Davies, Danielle Matthews

This paper appeared as:

Morisseau, T., Davies, C. & Matthews, D. (2013) How do 3- and 5-year-olds respond to under- and over-informative utterances? *Journal of Pragmatics* 59 (2013) 26-39
http://dx.doi.org/10.1016/j.pragma.2013.03.007

Abstract

As children learn their native languages, they come to have detailed expectations about how to refer to things. These expectations and the detection of their violations are key to inference-making processes. But what do children do when their expectations are not met? Using reaction-time measures and gaze-direction monitoring in a referential communication task, we investigated whether 3- and 5-year-olds notice the infelicity of under- and over-informative utterances and then seek out further information in order to recover the speaker's intended meaning. We tested how children resolve under-informative instructions such as ''Find the orange'' when there is more than one orange in view. We also tested whether instructions such as ''Find the cat with a tail'', in a context where there is only one, normal-looking cat, would lead them to question why the speaker was over-informative and to seek out further information. Both age groups were sensitive to the ambiguous instructions. Only 5-year-olds were significantly delayed and more likely to check their interlocutor's gaze when responding to over-informative expressions. We discuss how children's spontaneous motivation to resolve violations of expectation, coupled with increased speed of linguistic processing, drives language learning.

1. Introduction

As children learn their native language they develop very detailed expectations about what terms might be used to talk about a given situation. Indeed, some recent accounts consider this ability to predict what will be said given the context to be central to language acquisition (Ramscar et al., 2010). In particular, with regard to the development of referential choice, children need to learn to expect speakers to produce referring expressions that provide just enough useful information for the hearer to identify the referent in question (as adults do: Ariel, 1988). Recent research has demonstrated that, at around 5 years of age, children disprefer both under- and over-informative referential expressions when compared to optimally informative equivalents (Davies and Katsos, 2010). However, there is relatively little research indicating how young children act once they detect an infelicitous referential expression. Taking action is important for two reasons. First, as language comprehenders (in the moment), children need to take steps to accurately ascertain what it is that their interlocutor is talking about (see Sperber and Wilson, 1986/95, for a general theory of this process). Second, as language learners (in the longer term), children need to be able to check why people use the language they do in a given situation in order to adjust their language model. In this study, we tested whether children detect pragmatic infelicities due to under- or over-informativity in a referential communication task (where the instructions were always semantically accurate). Of particular interest is whether, and by what means, they actively seek out further information in order to recover the speaker's intended meaning.

1.1. Comprehending under-informative utterances

A vast body of work has investigated children's comprehension of under-informative expressions. For example, Revelle et al. (1985) found that 4-year-olds were able to detect ambiguous messages in a referential communication task and generally requested clarification before taking action. The 3-year-olds in their study, however, lacked verbal strategies for coping with ambiguity, although they exhibited appropriate and selective monitoring responses for other kinds of problems such as unintelligible or impossible requests (Revelle et al., 1985; see also Flavell et al., 1981). It has been debated whether young children can even detect if a request is intrinsically ambiguous, or if they only discover the communicative problem if it renders them unable to act, as suggested by Markman (1977, 1981) and Robinson and Peter Robinson (1977). Singer and Flavell (1981) investigated the question of how a positive versus negative communicative outcome influenced children's judgments of ambiguity. 5-year-olds' but not 7-year-olds' responses depended on whether the listener said he could not comply with the instruction. These results suggest that, during the early school years, children develop the awareness that an ambiguous message is intrinsically unclear regardless of the listener's response to it. More recently, Nilsen et al. (2008) found that 4-year-old children show sensitivity to ambiguity from another listener's perspective, even when they possess private knowledge that allows them to resolve the ambiguity. This last result was only found when measures of implicit awareness (child's eye movements between possible referents) rather than explicit awareness (request for clarification) were taken (Nilsen and Graham, 2012; Nilsen and Fecica, 2011). This raises the question of whether even younger children are implicitly sensitive to referential ambiguity.

Evidence from the word learning literature suggests that children have the pragmatic wherewithal to detect ambiguity and seek clarifying information from an interlocutor. For example, it is established that 2- to 4-year-olds check the gaze of their interlocutor in an attempt to establish which of two possible objects she is referring to with a novel word (Nurmsoo and Bloom, 2008; see also Diesendruck, 2005; Grassmann et al., 2009). Indeed, even 13-month-old infants check the line of regard of an interlocutor more when a novel word is produced with two novel objects in view rather than one (Vaish et al., 2011). It is thus possible that young children can exploit indications of the speaker's epistemic state in order to resolve ambiguous referring expressions. However, there are several reasons to think that the application of these early pragmatic skills to reference resolution in other settings may not be straightforward (Matthews et al., 2006). First, the use of a novel word renders the presence of a communication gap that needs to be repaired obvious to the child. In contrast, when a familiar word is used ambiguously in the presence of two similar objects, resolution is more challenging because it requires the child to detect the ambiguity per se. Moreover, when utterances and the contexts they are produced in become more complex, the number of inferences a child could draw at any one moment increases exponentially. Whereas in a word learning task the child is presented with one novel word and one or two objects, in referential communication tasks (and arguably in many everyday interactions), children are presented with complex descriptions made up of familiar words and structures (drawn from a set of many possible alternative descriptions) and large arrays of objects. It is an empirical question as to how children respond to inferences about the speaker's intent under such circumstances.

1.2. Comprehending over-informative utterances

That a referring expression has to contain enough information to identify a referent is not the only expectation one might have regarding the informativity of a sentence: utterances that contain too much information are also infelicitous. Although far less research has investigated children's sensitivity to over-informativity, Davies and Katsos (2010) demonstrated that 5-year-olds tend to judge object descriptions with redundant adjectives as non-optimal. As a result of this growing expertise, 5-year-old children consider the addition of some adjectives to a canonical description to mark a contrastive function. For example, Huang and Snedeker (2008) adapted the referential paradigm used by Sedivy et al. (1999) and found that 5-year-olds match a modified expression to a target object more rapidly in trials containing a contrast set. Likewise, a recent study of 3-year-olds' adjective comprehension suggests that even at this age there is some sensitivity to over-informativity. Bannard et al. (2013) had children observe one adult use a novel adjective to request a toy from another adult (e.g., ''Give me the dilsige duck''). If the use of the adjective was justified by a contrastive motive (e.g., there were two ducks in view, one of which was marked), then children were more likely to imitate the use of the adjective than if there was no reason to produce it (i.e., because there was only one duck). Thus, children as young as 3 years of age were more reluctant to imitate an adjective if they thought its use would be over-informative. This is in line with findings that, in simplified contexts, children as young as 2 years of age tend to avoid producing redundant information in their referring expressions (Matthews et al., 2007, 2012). Therefore it seems to be the case that from 3 years of age, children are able to make a simple association between the presence of a contrast set and the use of a modified expression (be this with an adjective or prepositional phrase).

The question we are interested in is how this sensitivity to over-informativity affects language comprehension, that is, how a child responds to the presence of additional information that is not relevant to the task at hand given the visual context. In referential communication tasks, over-informativity never prevents the hearer from identifying the target object. In this sense, the child does not need to repair the utterance to respond correctly to her interlocutor's instruction. However, as a language learner and social partner, the child may generate expectations as to what their interlocutor should have said. If the child notices that these expectations are not met, it would be in her interest to establish why. But do young children go so far as to seek out additional information (e.g., look for a contrasting referent) that would justify the otherwise redundant description? If they do, this would be a striking illustration of children's desire to have a good fit between their predicted language use and what they hear around them in order to ensure the success of their communicative interactions.

1.3. Testing children's response to under- and over-informative utterances

In the current study we looked at the actions children take when a referring expression does not meet their pragmatic expectations. Given that, by the age of 3, children are already processing complex noun phrases quite rapidly (Fernald et al., 2010), we predicted that 3- and 5-year-old children would be sensitive to both types of infelicities, resulting in longer reaction times and increased gaze checking on infelicitous trials. In addition, we predicted that the ability to take action by asking for clarification (in the under-informative cases) and visually searching the referential context (in the over-informative ones) would be stronger in the older group, presuming that knowledge of how to deal with problematic messages increases with age (Sonnenschein, 1982; Lloyd et al., 1998). Furthermore, looking at children's reactions to both under- and over-informative utterances allowed us to test whether individual differences in response to these two types of infelicity were correlated. This is currently an important question in developmental pragmatics since studies show that children with language disorders do not tend to have pragmatic impairments across the board in the way one might expect (Bishop and Adams, 1991; Norbury et al., 2004; Norbury, in press). This raises the question as to whether common cognitive and social-cognitive abilities underlie the processing of different types of referring expression.

3- and 5-year-old comprehenders were asked to move pictures around a grid in order to match a target configuration. Children responded to verbal instructions in a 2 (Modification: Modified vs. Non-modified instructions) 2 (Contrast: presence of a Contrast set vs. No Contrast) design. Modified instructions included a prepositional phrase modification (e. g., ''Move the cat with a tail'') whereas Non-modified instructions did not (''Move the cat''). The contrast variable refers to the presence of either one or two objects in the grid that were potential referents of the head noun. Because a certain level of redundancy is to be expected in adult language (e.g., Viethen and Dale, 2007; Nadig and Sedivy, 2002), we chose prepositional phrase modifications that described a typical feature of the referent. This reduced the probability that the modification would be interpreted as being used for descriptive (rather than contrastive) purposes (Karmiloff-Smith, 1979:45). That is, from an adult point of view, the prepositional phrases we used would be entirely redundant unless there was a contrast object that lacked that feature in question. Of course, children are more likely than adults to hear redundant descriptions of objects as their caregivers attempt to teach them new words. Nonetheless, based on previous research, we expected even young children to potentially find these descriptions over-informative. Our question was what action they would take upon detecting such an infelicity.

2. Methods

2.1. Participants

Sixty typically developing, monolingual, French-speaking children participated in the study. There were thirty 3-year-olds (15 males and 15 females) with an average age of 3 years 2 months (range = 2;10-3;9) and thirty 5-year-olds (14 males and 16 females) with an average age of 5 years 3 months (range = 4;10-5;9). Within each age group, subjects were randomly assigned to one of two lists, which differed with respect to which items were administered in the experimental condition versus the control condition. As a result, all children heard the exact same instructions but in different contexts that rendered them either felicitous or infelicitous, depending on the list they were assigned to. All children were tested in a quiet room at their school in Lyon, France.

2.2. Design

The experiment had a 2 x 2 factorial design. The variable Modification refers to whether the head noun referring to the target was modified with a prepositional phrase (Modified condition) or not (Non-modified condition). The variable Contrast refers to whether the target object was one of two objects of the same nominal category in the grid (Contrast condition) or was the only object of its category in the grid (No contrast condition). Both variables were administered within-subjects. Crossing these factors resulted in four types of critical instructions:

- Under-informative, e.g., ''the banana'' in a context of there being two bananas (one peeled and one unpeeled);
- Optimal without modification, e.g., ''the train'' in a context of there being only one train;
- Over-informative, e.g., ''the cat with a tail'' in the context of there being only one normal-looking cat;
- Optimal with modification, e.g., ''the house with windows'' in the context of there being two houses, one without windows.

Items were counterbalanced such that, for the study as a whole, each item was heard equally often as optimal and non-optimal.

2.3. Materials

A wooden grid with 16 (4 x 4) slots was fixed on a table and used to present the stimulus pictures, which depicted familiar objects. One slot always remained empty, which allowed the participant to displace one object at a time. The colored stimulus images were printed on white cards and placed in plastic holders so that they could stand on their own (see Fig. 1). Four sets of cards were created, each corresponding to a round of the game. All children completed four rounds and the order of the rounds was randomized. Categories of objects (e.g., animals or toys) were balanced across the rounds.

Each set of cards was created such that the grid contained two key contrast sets and two key singletons. For the purposes of counterbalancing, two versions of each set were created differing only in terms of two items such that a key singleton in version 1 became a member of a contrast set in version 2 and vice versa. As mentioned above, participants were randomly assigned to one of two lists. Experimental items that were under- or over-informative for participants assigned to list 1 were control (optimal) items for participants assigned to list 2 and vice versa. Thus, differences observed between conditions could not be due to the instruction being inherently odd. The remaining items in each set of cards were the same across subjects, and they were always referred to in an optimal way, whether they were singletons or part of a contrast set. The critical contrast object was always located in the grid right next to the target object.

Fig. 1. Example of experimental set-up.

2.4. Procedure

Children were tested individually. The experimental session took place during normal class time, which children left one at a time to participate in the experiment, and lasted approximately 20 min. Children sat at the table in front of the grid. Paper hands were placed on the table and children were asked to put their hands on the paper hands to show they were ready before each trial. This ensured that reaction times (time taken to move the hand from the table to a slot in the grid) would be measured from the same point for all trials. The distance between the edge of the table and the grid differed between the two age groups due to height differences. Their data sets are thus analyzed separately. The experimenter was seated near the left side of the table, so that the child would have to turn her head 45° to look at her. A Logitech webcam (QuickCam Vision Pro 9000) was placed at 90° to the grid on the same side of the table, at a distance such that it captured both the grid and the child. In this way, both hand movements and looks to the experimenter could be recorded.

Children were told that they were going to play a game with instructions that can be glossed in English as follows: we are going to play a game together in which you will have to move objects around this grid. The goal is to move these objects so that the grid becomes exactly the same as the one shown in this booklet. For each trial, I will tell you which object to move, and you will pick up the correct object and put it in the slot that remains empty. After you are finished with one object, you must always put your hands on the paper hands in front of you''.

In a **warm-up phase**, children heard three instructions, all of which were optimal. If necessary, the experimenter would help the child by showing her the target object, the empty slot or where to put her hands after the trial.

The **test phase** consisted of four rounds, each including four critical instructions (one item for each of the four types of critical instructions) and five fillers (all of which were optimal). Instructions to move objects within each round were exactly the same across subjects and uttered in the same tone and in the same order. The experimenter uttered her instructions while looking at the printed grid she kept in her hands, and then looked at the child until she had chosen a card (i.e., she did not look at the target object).

In the case of under-informative instructions, there was no 'right' response (as two objects were potentially being referred to, one of which was arbitrarily termed the target for data analysis purposes). If the child asked for clarification, the experimenter answered with the initial instruction adding in the relevant modifier. If a child picked up one of the objects that met the description without asking for clarification, the experimenter treated this as a correct response. Some children did not ask for clarification verbally but did so by looking continuously at the experimenter as if for further information, sometimes without having selected any object, sometimes after having touched one of the objects that met the description. Looking at the experimenter without picking-up an object was considered an attempt to express inability to respond to the instruction. In such cases, the experimenter waited three seconds before completing her request with the relevant modification. Finally, if the child had difficulty finding where to put the target object, the experimenter would give her details about the location of the empty slot. The full list of instructions for each grid is presented in Appendix 1. In a **tidy-up phase**, after the completion of each round, the child was invited to help the experimenter put away the pictures. During this phase, while all the cards were still visible, the experimenter asked the child to name the four items that had been referred to with critical instructions (two infelicitous and two control instructions), by saying ‘‘Quelle est cette carte?’’ (‘‘What is this card?’’). This allowed us to check whether children would reproduce the infelicitous descriptions they had heard or spontaneously use an optimal description.

2.5. Coding

Off-line coding was conducted using the multimedia annotator ELAN, developed by the Technical Group of the Max Planck Institute for Psycholinguistics (Lausberg and Sloetjes, 2009).

Each test trial was coded according to whether or not the child reacted verbally and whether or not she looked at the experimenter. Reaction times to pick up the target object were also measured.

Gazes to the experimenter were coded from the offset of the instruction to the moment the child’s hand left the grid after putting the object in the target location (gazes that could have been interpreted as an indication that she was ready for the next instruction were thus not included). Verbal comments were fully transcribed.

Children’s reactions to under-informative instructions were coded as *clarification requests* either when the child explicitly asked ‘‘Which one?’’ or when she expressed her inability to respond to the instruction by looking intently at the experimenter without picking-up any object. Other cases where children looked (briefly) at the experimenter were coded as *gaze checks*.

Reaction times to pick up the target object were measured for all of the 16 experimental items (one over-informative, one under-informative and the two corresponding controls for each of the four grids) from the offset of the instruction to the moment the child’s hand left the grid holding the object. In the Non-modified condition, trials with a clarification request from the child were not included in the reaction time analysis. In the Modified condition, children sometimes reacted before the end of the sentence (‘‘false starts’’) and we removed these trials from the reaction time data. 53 trials (out of 210) for the 3-year-old group and 25 for the 5-year-old group were discarded on this basis.

Finally, children’s descriptions in the tidy-up phase were transcribed and coded as either *modified* or *non modified*. Data for the item ‘‘the clock with hands’’ were removed because five of the 3-year-old subjects experienced difficulties in understanding the description and finding the corresponding target object.

To test for the reliability of the first author’s coding, 10% of the data (90 trials) were randomly selected for re-analysis by a trained assistant. Intercoder percentage agreement was 100% for verbal responses and gaze checks and Cronbach's a reliability coefficient reached .99 for reaction times.

3. Results

The children were motivated to play the game and generally responded attentively during all trials.

We consider the children's reactions first to under-informative trials and then to over-informative trials using three measures: (1) requests for clarification (in the Non-modified condition) or comments (in the Modified condition); (2) gaze checks to the Experimenter and (3) reaction times. Reaction times elicited by 3- and 5-year-olds were not comparable because the distances between the grid and the children were not the same for the two groups (the position of the grid was adapted to the size of the children). For ease of interpretation, then, reaction time data for the two groups were treated separately. Reaction times were log-transformed to reduce skew in their distribution. We also looked at the spontaneous production of adjectives in the tidy-up phase. Finally, we explored whether individual differences in children's responses to the two trial types were associated with each other.

Because they were binary variables, we analyzed clarification requests, gaze checks and adjective production data using mixed-effect logistic regression, which treated the experimental factor of condition (Contrast vs. No Contrast) as a fixed effect and subjects and items as random effects (Jaeger, 2008). Reaction time data were analyzed using linear mixed-effect regression (Baayen et al., 2008).

3.1. Effects of under-informativity

When instructions were optimal, no clarification requests were observed for either age group. Fig. 1 presents the distribution of clarification requests for each age group in the under-informative condition. Both tended to elicit clarification verbally (70% of 3-year-olds' and 78% of 5-year-olds' clarification requests were verbal, with the remainder being achieved by persistently looking at the experimenter until she clarified herself). To assess whether children of each age group were significantly more likely to request clarification in under-informative compared to optimal trials, we fitted a mixed-effect logistic regression model to the clarification request data. Table 1 shows that the effect of age on the number of clarification requests was significant, with 5-year-olds asking for clarification for a mean of 2.4 out of the 4 under-informative trials (in this age group, 83% asked for clarification at least once) compared to a mean of 1.1 request by the 3-year-olds (in this age group, 53% asked for clarification at least once).

A more subtle index of the child's sensitivity to under-informative instructions was the number of gaze checks to the experimenter in trials with no clarification request. Fig. 3 presents the proportions of trials where the child gaze checked the experimenter as a function of Modification, Contrast and Age group. Focusing on the non-modified instructions, 3-year-olds gaze checked the experimenter in 16% of under-informative trials with no clarification request and in 10% of control trials (respectively 6% and 0% for the 5-year-olds). The results of the mixed-effect logistic regression (see Table 2) showed that children were significantly more likely to gaze check their interlocutor when the instruction was under-informative rather than optimal, and that 3-year-olds relied significantly more on gaze checking overall.

The last index of sensitivity to infelicity that we considered was reaction times. We tested whether, on trials with no clarification request, children nonetheless showed any signs of slowing in response to infelicity. Means of log-transformed reaction times to instructions in the optimal and under-informative conditions were compared (there were199 observations for the 3-year-olds and 129 observations for the 5-year-olds). As noted in the method, only trials with no clarification request were included in the analysis. Thus, some children (those who asked for clarification in all under-informative trials) do not appear in the analysis (2 in the 3-year-old group and 10 in the 5-year-old group). Fig. 4 shows that 5-year-olds took more time in responding to non-optimal instructions for which they did not ask for clarification (mean log-transformed RTs = 1.134 compared to 0.799 in the optimal condition). Table 3 summarizes the model results and shows that this effect of condition was significant for the 5-year-old group. No such difference was observed among the 3-year-old group, suggesting that if children did not ask for clarification, they probably did not notice the contrast.

Table 1

Results of mixed-effect logistic regression model for number of clarifications.

Non-modified condition (nb of obs: 240)			
Random effects		Variance	Std dev
Subject (intercept)		4.146	2.036
Item (intercept)		0.000	0.000
Fixed effects	Coefficient	Odds ratio	*z*-Value
Intercept	0.652	1.919	1.466
Age (three)	-2.310	0.099	-3.589***

Null model likelihood ratio test: $x^2(1) = 11.877$, $p < 0.001$.

* $p < 0.05$.

** $p < 0.01$.

*** $p < 0.001$.

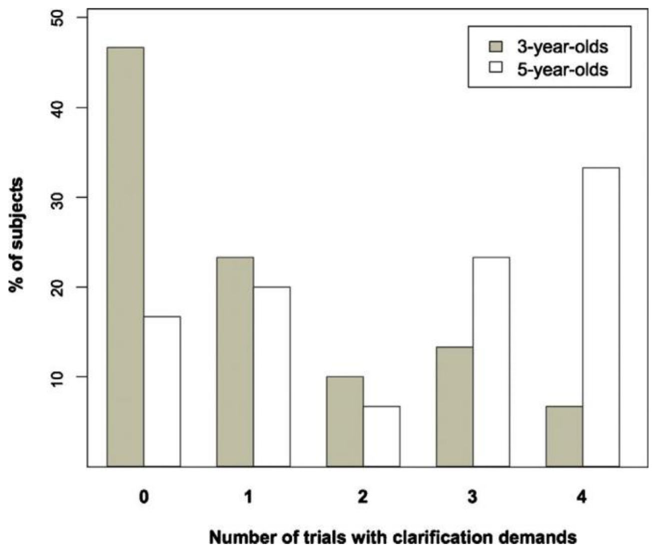

Fig. 2. Distribution of clarification demands in response to under-informative instructions.

3.2. Effects of over-informativity

Five children (all 5-year-olds) verbally noted the oddness of an instruction at least once when it was over-informative, for example by saying ‘‘Il y a un autre X?’’ (‘‘Is there another X?’’). Since they always did so after having selected the target object, we did not exclude those trials from analyses of reaction times.

The logistic regression results for Modified instructions shown in Table 2 indicate that mean numbers of looks to the experimenter did not differ according to the felicity of the instruction across both groups. The 3-year-olds gaze checked significantly more often than the older group across conditions (16% of all trials, compared to 7% for the 5-year-olds. See Fig. 2). The interaction between the two variables was not significant.

Mean log-transformed reaction times in the modified condition are presented in Fig. 5. Again, separate mixed effects regression models (with Condition as a fixed effect) were tested for each age group because reaction times elicited by the two groups were not comparable. As can be seen in Table 3, 5-year-olds were significantly slower to respond when the instruction was over-informative than when it was optimal (mean log-transformed RTs to over-informative trials = 1.095 compared to 0.872 in the optimal condition). However, no such model fitted the data satisfactorily for the 3-year-old group. It is possible that the presence of a contrast set made it more difficult for the children to resolve the referring expression.

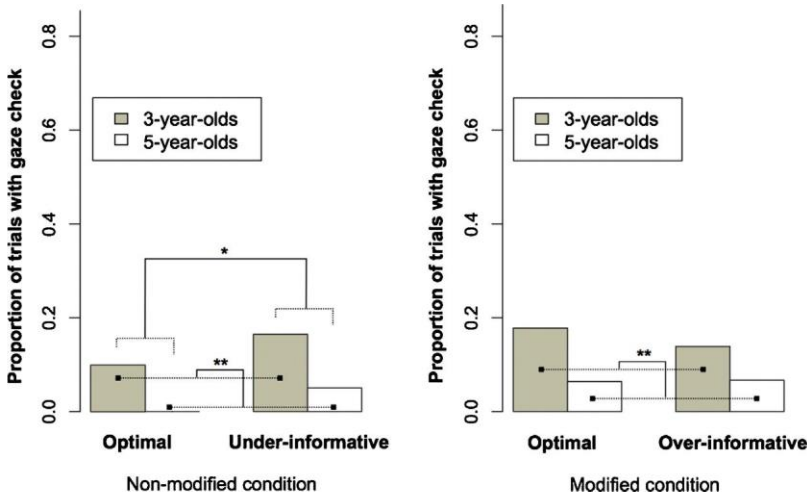

Fig. 3. Proportions of trials where the child gaze checked the Experimenter as a function of Age, Instruction and Presence/Absence of contrast set. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Table 2

Results of mixed-effect logistic regression models for number of gaze checks.

Non-modified condition (nb of obs: 376)			
Random effects		Variance	Std dev
Subject (intercept)		1.123	1.060
Item (intercept)		0.508	0.713
Fixed effects	Coefficient	Odds ratio	*z*-Value
Intercept	-5.183	0.006	-6.394***
Condition (under-inf.)	0.999	2.715	2.222*
Age (three)	2.243	9.421	2.831**

Null model likelihood ratio test: $x^2(2) = 18.727$, $p < 0.001$.

Modified condition (nb of obs: 420)

Random effects		Variance		Std dev
Subject (intercept)		0.745		0.863
Item (intercept)		0.050		0.223
Fixed effects	Coefficient		Odds ratio	z-Value
Intercept	-2.895		0.055	-7.494***
Condition (over-inf.)	-0.204		0.815	-0.618
Age (three)	1.109		3.031	2.583**

Null model likelihood ratio test: $x^2(2) = 7.413$, $p = 0.024$.

*$p < 0.05$.

**$p < 0.01$.

***$p < 0.001$.

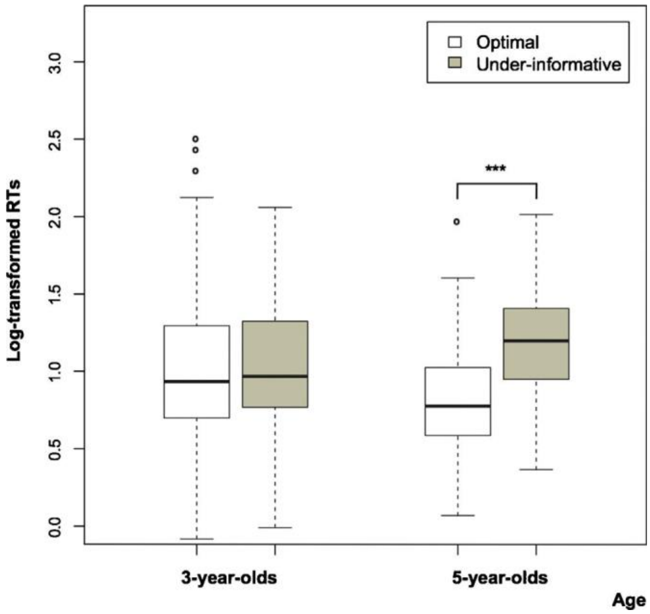

Fig. 4. Log-transformed mean reaction times as a function of Age and Presence/Absence of contrast set for Non-modified instructions. NB: data only include trials where the child did not ask for clarification. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Table 3

Results of mixed-effect linear models for 5-year-olds' log-transformed reaction times.

Non-modified condition (nb of obs: 129)

Random effects	Variance	Std dev
Subject (intercept)	0.021	1.146
Item (intercept)	0.000	0.000
Residual	0.137	0.371

Fixed effects	Coefficient	Odds ratio	*t*-Value
Intercept	0.799	2.223	15.16***
Condition (under-inf.)	0.328	1.388	4.807***

Null model likelihood ratio test: $x^2(1) = 21.431$, $p < 0.001$.

Modified condition (nb of obs: 185)

Random effects	Variance	Std dev
Subject (intercept)	0.034	0.186
Item (intercept)	0.018	0.134
Residual	0.153	0.391

Fixed effects	Coefficient	Odds ratio	*t*-Value
Intercept	0.873	2.394	12.067***
Condition (over-inf.)	0.215	1.240	3.663***

Null model likelihood ratio test: $x^2(1) = 12.977$, $p < 0.001$.

*$p < 0.05$.

**$p < 0.01$.

***$p < 0.001$.

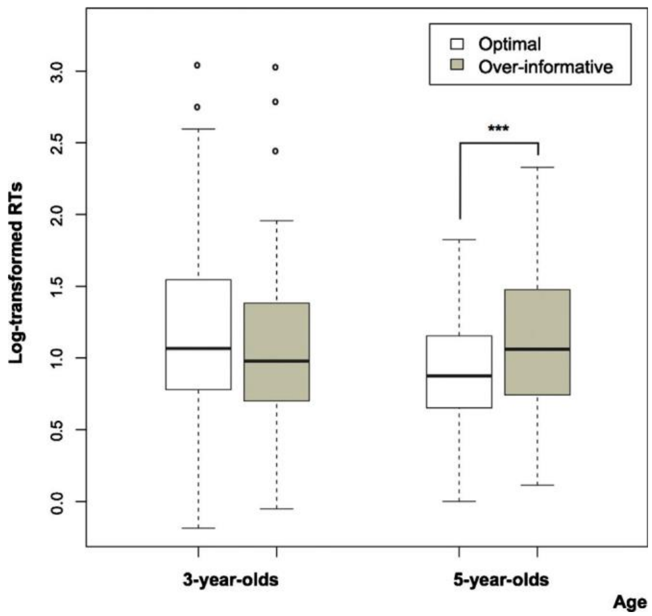

Fig. 5. Log-transformed mean reaction times as a function of Age and Presence/Absence of contrast set for Modified instructions. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Table 4

Results of mixed-effect logistic regression models for adjective production in the tidy-up phase.

Non-modified condition (nb of obs: 480)

Random effects		Variance	Std dev
Subject (intercept)		0.000	0.000
Item (intercept)		0.077	0.277

Fixed effects	Coefficient	Odds ratio	*z*-Value
Intercept	-5.253	0.005	-5.081***
Condition (under-inf.)	3.146	23.242	3.019**
Age (three)	-0.613	0.542	-1.314

Null model likelihood ratio test: $x^2(2) = 25.103$, $p < 0.001$

Modified condition (nb of obs: 420)

Random effects	Variance	Std dev
Subject (intercept)	0.000	0.000
Item (intercept)	0.054	0.232

Fixed effects	Coefficient	Odds ratio	*z*-Value
Intercept	-1.938	0.144	-6.612 ***
Condition (over-inf.)	-0.385	0.680	-0.968
Age (three)	-1.438	0.237	-3.036**

Null model likelihood ratio test: $x^2(2) = 12.363$, $p = 0.002$.

*$p < 0.05$.

**$p < 0.01$.

***$p < 0.001$.

3.3. Children's production of referring expressions during the tidy up phase

Overall, children rarely used modification in this phase of the study, when asked to name the objects (see Table 4).[1] In the Non-modified condition, only the presence of a contrast led children to specify the adjective: children never spontaneously used a modified description when it would have been over-informative. No significant difference between age groups was observed (7 and 14 trials with adjective production respectively in the 3- and 5-year-old groups, out of 240 trials). In the Modified condition, the number of adjective productions differed by age group (6 and 23 trials with adjective production respectively in the 3- and 5-year-old groups, out of 210 trials) but no significant difference was observed between cases where it was optimal and cases where it was not.

3.4. Individual differences on under- and over-informative trials

It was interesting to consider whether children who were sensitive to one type of infelicity tended to be sensitive to the other. An index of individual sensitivity to infelicity in both modified and non-modified conditions was calculated, using the difference between numbers of either gaze checks or verbal reactions to critical instructions (under- and over-informative) and their respective controls. No significant correlation was observed in either group: Spearman's $r(3835) = 0.09$, $p = .64$ for the 3-year-olds and $r(4889) = 0.177$, $p = .43$ for the 5-year-olds.

4. Discussion

The goal of this developmental study was to investigate how preschoolers' expectations of informativity affect language processing. We were interested in whether children would take action when confronted with messages that do not meet their expectations. To summarize, we found that both 3- and 5-year-

[1] NB: because there was no trial with adjective production in the non-modified condition, we dealt with the problem of testing zero variance components by adding one positive trial for the purpose of the analysis.

olds were sensitive to under-informative utterances and looked to the experimenter to resolve the ambiguity. The 3-year-olds sought clarification (either verbally or by looking intently at the experimenter until she provided a clarification) on just over a quarter of trials, whereas the 5-year-olds did so roughly twice as often. Also, both age groups gaze checked the experimenter significantly more often when her instructions were under-informative compared to when they were optimal. Even on trials where they did not request clarification, 5-year-olds slowed down significantly in response to ambiguous instructions. They were also significantly slower to respond to over-informative utterances than optimal ones of the same length, whereas the younger children showed no sensitivity to over-informative utterances. In line with our assumption that 5-year-olds were slowed down because they had made a contrastive inference, five children in this age group went so far as to comment on the infelicity of over-informative instructions. When it came to producing their own descriptions in the tidy-up up phase, children rarely produced modified expressions. When they did so with items that had been in the non-modified condition, it was always because there was a contrast object. Finally, children's performance on under- and over-informative trials did not appear to be related.

The findings for under-informative utterances build on those showing that 4-year-olds are able to detect ambiguity albeit implicitly (cf. Nilsen et al., 2008). Studies from the language learning literature suggest that sensitivity to ambiguity is present well before age 3, and in some simple cases by 13 months (Vaish et al., 2011). In these word-learning tasks, though, the detection of ambiguity relies on a more obvious gap between the message (a novel word) and the speaker's meaning. Our results show that this ability can also apply earlier than previously thought in a referential task that requires the hearer to detect the ambiguity even when a possible candidate fitting the description is available from the child's point of view. Thus, as they get older, children become increasingly aware that messages, whatever their source, can be problematic, and they acquire a growing knowledge of how to deal with them, including the fact that sometimes seeking clarification is necessary.

Sensitivity to the over-informative instructions in this task was even more demanding because the infelicity did not prevent the child from responding accurately to the instruction. Contrary to other experimental set-ups used in the literature on contrastive inferences, the modification in our task was post-nominal. By using pre-nominal adjectives, researchers have been able to demonstrate children's ability to make use of the presence of a contrast set in order to resolve a referring expression more rapidly (e.g., Huang and Snedeker, 2008). Here, children could accurately identify the referent with the information conveyed by the head noun. Thus, any contrastive inference processed afterwards was purely gratuitous. Yet, 5-year-olds responded significantly more slowly to over-informative instructions than to the exact same message when given under optimal conditions. In doing so, they demonstrate a solid grasp of how much information they expect to be used to describe a target, and deviations from these expectations are strong enough to generate slower reactions. In several cases this violation even warranted a verbal comment, clearly demonstrating that they understood that modification implied contrast and that they wished to understand why the speaker had used such an expression (see also

Shwe and Markman, 1997, for an example where the child corrects a listener's misunderstanding of what he himself has said).

How can we explain these behaviors? In our everyday environment, some features are more relevant to mention than others because they are unusual (or ''unpredictable'') as opposed to predictable features, like the yellow of a banana (Sedivy, 2003). For instance, mentioning an unpredictable feature can be relevant because the speaker wants to help the hearer identify the target or merely to stress the fact that the object possesses an unusual feature. In contrast, no contextual feature would normally justify the use of a highly predictable adjective other than the presence of a contrast object (an atypical one that did not possess this typical feature of the regular object). Sedivy (2003) demonstrated that people, including children, spontaneously make a contrastive inference only when the adjective is highly predictable. The instructions used in our over-informative trials contained modifications that described a typical feature of the referent exclusively (thus highly predictable for children who have gained a minimal level of conceptual knowledge about the target objects). Since the target objects were all familiar to young children, the presence of a contrast object in the grid was the only good reason that could justify the speaker's choice to use such predictable modifier. In the control condition, this contrast object was next to the target object and thus readily accessible. We therefore interpreted the longer reaction times observed among 5-year-olds in the over-informative condition as evidence of children's attempt to make sense of the over-informative description. Most probably, this result would disappear if we used less predictable adjectives.

In our introduction, we assumed that by 3 years of age, children were generally quite capable of drawing contrastive inferences in simple settings. In Clark's terms, they know that ''every two forms contrast in meaning'' (Clark, 1987). Therefore, the question of why 3-year-olds did not slow down in response to over-informative utterances needs addressing. Of course, we cannot exclude the possibility that the source of the 3-year-olds' troubles may have been methodological. For example, a lack of effect might be due to the number of relevant objects in the over-informative trials (a unique object of its type) compared to the matched control trials (two objects of the same type). We might expect young children to take longer in the control trials with two objects because they need to discriminate between two potential targets. Thus, the 3-year-olds may simply have spent more time on this discrimination task in the control condition. Given these methodological limitations, the performance of the 5-year-olds is all the more striking.

Another potential explanation for the 3-year-olds' performance is that, even though they spontaneously used non-modified expressions to describe unique objects, they did not have strong expectations about how the object was supposed to be referred to and so they could not contrast their expectations with what was said. 5-year-olds may have stronger expectations to this regard, such that if they are not met they will infer something to the effect of ''if she said that, she must have a reason'' (see Gergely et al., 2002, for an example of such reasoning in a non-linguistic test). It is further reasonable to assume that the strength of expectation is related to the degree of pragmatic tolerance a child demonstrates (Katsos and Bishop, 2011). Another possibility, consistent with the reaction time findings, is that the younger children lacked sufficient cognitive resources to rapidly process the complex description. Children who did not detect the infelicity may have processed the head noun (which sufficed) and located the target before they processed the redundant information (either at all or in any depth). In this sense they would not even have noticed the infelicity. Such an explanation might also explain inverse frequency effects found in tests of children's anaphora resolution (Matthews et al., 2009).

Finally, the fact that 3-year-olds relied more heavily on gaze checking (they used it more often than 5-year-olds across conditions) is of interest because it signals a difference in children's relation to verbal instructions, with 3-year-old children seeming less confident in their responses[2]. Looks to the experimenter in such situation can indeed be interpreted as an attempt to monitor one's action under the experimenter's guidance. The issue of whether the speaker's status played a role in the way children reacted to the instructions (in both under- and over-informative cases) should be considered here. Following the hypothesis of Sonnenschein (1986), it is possible that child comprehenders attribute a certain amount of intelligence and language ability to the experimenter, relative to their own. In Sonnenschein's study, the age of the speaker (child vs. adult) had an effect on children's evaluations of non-informative messages. When the child's procedural rules for this evaluation were either lacking or unstable, adults' messages were evaluated more positively than peers'. It might thus be the case that to some extent, the hierarchical situation prompted children to act automatically without questioning relevance in the context (cf. Grodner and Sedivy, 2011). As they get older, children's confidence in the norms of language use allows them to be more demanding of their communicative partners, regardless of their status or presumed knowledge.

In sum, the study shows that between 3 and 5 years of age, children increase their sensitivity to infelicities of both types, making particular advances in their reactions to over-informative utterances. However, while most 5-year-olds clearly found both over- and under-informative utterances non-optimal, there were substantial individual differences in each condition. It is interesting to note that it was not the case that children who tended to respond to under-informative utterances by a gaze check or a verbal response were the same ones who tended to respond in the same way to over-informative utterances. This suggests that sensitivity to pragmatic infelicity can be piecemeal in nature. A larger sample size would be required to say anything definitive about individual differences, however. We assume it is probably the case that common features underlie these two cognitive abilities. In the adult psycholinguistic literature, it has been found that individuals with lower self-reported pragmatic abilities had a lower standard of relevance, to the extent that they were less likely to draw pragmatic inferences based on linguistic stimuli (Nieuwland et al., 2010) and were less likely to do well on the communication

[2] NB: although it is difficult to draw conclusions from the comparison of 3- and 5-year-olds' eye-gaze data due to multiple factors involved, the difference in the number of gaze checks between conditions within each age group remains relevant for our purpose.

subset of the ASQ questionnaire (Baron-Cohen et al., 2001). Further studies are needed to determine whether individual differences in terms of social competence are linked with performance in such referential tasks and, in the case of too much information, what makes them go as far as trying to explain the infelicity by looking for a contrast object.

In any case, by 5 years of age, most of the children in this study had developed clear expectations about the amount of information that a referring expression should include, and, if these expectations were not met, they sought information that would explain why not. It is precisely this curiosity that will allow them to make ever more fine-grained inferences as they become adult-like language users.

Acknowledgements

The work reported on in this publication has been financially supported by the European Science Foundation (ESF), in the framework of the Research Networking Programme EURO-XPRAG.

Appendix 1. Instructions

Grid 1

L'avion gris (the gray airplane); La fleur avec des pétales (the flower with petals)**; La jupe (the skirt); La tortue (the turtle); Le vélo (the bike)*; L'avion rouge (the red airplane); Le ballon (the balloon)*; Les gants (the gloves); Le cheval avec des oreilles (the horse with ears)**.

Grid 2

La moto orange (the orange motorbike); L'horloge avec des aiguilles (the clock with hands)**; La chemise (the shirt); Le mouton (the sheep); L'oiseau (the bird)*; La moto multicolore (the multicolor motorbike); La pomme (the apple)*; Les lunettes (the glasses); La voiture avec des roues (the car with wheels)**.

Grid 3

La framboise (the raspberry); Le chat avec une queue (the cat with a tail)**; Le pull (the sweater); Les bottes violettes (the purple boots); La banane (the banana)*; Le stylo (the pen); Le train (the train)*; Les bottes roses (the pink boots); La maison avec des fenêtres (the house with windows)**.

Grid 4

La fraise (the strawberry); Le bateau avec une voile (the boat with a sail)**; Les chaussettes (the socks); Les ciseaux bleus (the blue scissors); La poule (the hen)*; Le manteau (the coat); La chaise (the chair)*; Les ciseaux noirs (the black scissors); L'arbre avec des feuilles (the tree with leaves)**.

The items marked with * are experimental items for the under-informative condition for half subjects, and control items for the under-informative condition for the other half. Subjects who were administered the item ‘‘bike’’ as an experimental item, were also administered the items ‘‘apple’’, ‘‘train’’ and ‘‘hen’’ as experimental items. Subjects who were administered the item ‘‘balloon’’ as an experimental item, were also administered the items ‘‘bird’’, ‘‘banana’’ and ‘‘chair’’ as experimental items.

The items marked with ** are experimental items for the over-informative condition for half subjects, and control items for the over-informative condition for the other half. Subjects who were administered the item ‘‘flower’’ as an experimental item, were also administered the items ‘‘car’’, ‘‘cat’’ and ‘‘tree’’ as experimental items. Subjects who were administered the item ‘‘horse’’ as an experimental item, were also administered the items ‘‘clock’’, ‘‘house’’ and ‘‘boat’’ as experimental items.

References

Ariel, Mira, 1988. Referring and accessibility. Journal of Linguistics 24, 65-87.

Baayen, R. Harald, Davidson, Doug, Bates, Douglas, 2008. Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language 59, 390-412.

Bannard, Colin, Klinger, Jorn, Tomasello, Michael, 2013. How selective are 3-year-olds in imitating novel linguistic material? Developmental Psychology.

Baron-Cohen, Simon, Wheelwright, Sally, Skinner, Richard, Martin, Joanne, Clubley, Emma, 2001. The autism spectrum quotient (AQ): Evidence from Asperger syndrome/high functioning autism, males and females, scientists and mathematicians. Journal of Autism and Developmental Disorders 31, 5-17.

Bishop, Dorothy, Adams, Caroline, 1991. What do referential communication tasks measure? A study of children with specific language impairment. Applied Psycholinguistics 12, 199-215.

Clark, E.V., 1987. The principle of contrast: a constraint on language acquisition. In: Macwhinney, B. (Ed.), Mechanisms of Language Acquisition. Erlbaum, Hillsdale, NJ, pp. 1-33.

Davies, Catherine, Katsos, Napoleon, 2010. Over-informative children: production/comprehension asymmetry or tolerance to pragmatic violations? Lingua (Special Issue on Asymmetries in Child Language) 120 (8) 1956-1972.

Diesendruck, Gil, 2005. The principles of conventionality and contrast in word learning: an empirical examination. Developmental Psychology 41, 451-463.

Fernald, Anne, Thorpe, Kirsten, Marchman, Virginia, 2010. Blue car, red car: developing efficiency in online interpretation of adjective-noun phrases. Cognitive Psychology 60, 190-217.

Flavell, John H., Speer, James Ramsey, Green, Frances L., August, Diane L., Whitehurst, Grover J., 1981. The development of comprehension monitoring and knowledge about communication. Monographs of the Society for Research in Child Development 46 (5) 1-65.

Gergely, Gyorgy, Bekkering, Harold, Kiraly, Istvan, 2002. Rational imitation in preverbal infants. Nature 415, 755.

Grassmann, Susanne, Stracke, Marén, Tomasello, Michael, 2009. Two-year-olds exclude novel objects as potential referents of novel words based on pragmatics. Cognition 112, 488-493.

Grodner, Daniel, Sedivy, Julie, 2011. The effects of speaker-specific information on pragmatic inferences. In: Pearlmutter, N., Gibson, E. (Eds.), The Processing and Acquisition of Reference. MIT Press, Cambridge, MA.

Huang, Yi Ting, Snedeker, Jesse, 2008. The use of referential context in children's online interpretation of adjectives. In: Proceedings of the 30th Annual Conference of the Cognitive Science Society. Cognitive Science Society, Austin, TX, pp. 1380-1385.

Jaeger, Florian, 2008. Categorical data analysis: away from ANOVAs (transformation or not) and towards Logit Mixed Models. Journal of Memory and Language 59, 434-446.

Karmiloff-Smith, Annette, 1979, reprinted 1981. A Functional Approach to Child Language. Cambridge University Press, Cambridge. Katsos, Napoleon, Bishop, Dorothy, 2011. Pragmatic tolerance: implications for the acquisition of informativity and implicature. Cognition 120 (1) 67-81.

Lausberg, Hedda, Sloetjes, Han, 2009. Coding gestural behavior with the NEUROGES-ELAN system. Behavior Research Methods 41 (3) 841- 849.

Lloyd, Peter, Mann, Sandi, Peers, Ian, 1998. The growth of speaker and listener skills from five to eleven years. First Language 18, 81-103. Markman, Ellen, 1977. Realizing that you don't understand: a preliminary investigation. Child Development 43, 986-992.

Markman, Ellen, 1981. Comprehension monitoring. In: Dickson, W.P. (Ed.), Children's Oral Communication Skills. Academic Press, New York, p. 1981.

Matthews, Danielle, Lieven, Elena, Anna, Theakston, Tomasello, Michael, 2006. The effect of perceptual availability and prior discourse on young children's use of referring expressions. Applied Psycholinguistics 27, 403-422.

Matthews, Danielle, Lieven, Elena, Tomasello, Michael, 2007. How toddlers and preschoolers learn to uniquely identify referents for others: a training study. Child Development 76 (6) 1744-1759.

Matthews, Danielle, Lieven, Elena, Theakston, Anna, Tomasello, Michael, 2009. Pronoun co-referencing errors: challenges for generativist and usage-based accounts. Cognitive Linguistics 20 (3) 599-626.

Matthews, Danielle, Butcher, Jessica, Lieven, Elena, Tomasello, Michael, 2012. Two- and four-year-olds learn to adapt referring expressions to context: effects of distracters and feedback on referential communication. Topics 4 (2) 184-210.

Nadig, Aparna, Sedivy, Julie, 2002. Evidence of perspective-taking constraints in children's on-line reference resolution. Psychological Science 13 (4) 329-336.

Nieuwland, Mante, Ditman, Tali, Kuperberg, Gina R., 2010. On the incrementality of pragmatic processing: an ERP investigation of informativity and pragmatic abilities. Journal of Memory and Language 63, 324-346.

Nilsen, Elizabeth, Fecica, Agnieszka, 2011. A model of communicative perspective-taking for typical and

atypical populations of children. Developmental Review 31 (1) 55-78.
Nilsen, Elizabeth, Graham, Susan, 2012. The development of preschoolers' appreciation of communicative ambiguity. Child Development 83 (4) 1400-1415.
Nilsen, Elizabeth, Graham, Susan, Smith, S., Chambers, Craig, 2008. Preschoolers' sensitivity to referential ambiguity: evidence for a dissociation between implicit understanding and explicit behavior. Developmental Science 11 (4) 556-562.
Norbury, C., Atypical pragmatic development. In: Matthews, D. (Ed.), Pragmatic Development in First Language Acquisition. John Benjamins, Amsterdam, in press.
Norbury, Courtenay Frazier, Nash, Marysia, Baird, Gillian, Bishop, Dorothy, 2004. Using a parental checklist to identify diagnostic groups in children with communication impairment: a validation of the Children's Communication Checklist-2. International Journal of Language and Communication Disorders 39 (3) 345-364.
Nurmsoo, Erika, Bloom, Paul, 2008. Preschoolers' perspective-taking in word learning: do they blindly follow eye gaze? Psychological Science 19, 211-215.
Ramscar, Michael, Yarlett, Daniel, Dye, Melody, Denny, Katie, Thorpe, Kirsten, 2010. Feature-label-order effects and their implications for symbolic learning. Cognitive Science 34 (6) 909-957.
Revelle, Glenda L., Wellman, Henry M., Karabenick, Julie D., 1985. Comprehension monitoring in preschool children. Child Development 56, 654-663.
Robinson, Elizabeth J., Peter Robinson, W., 1977. Children's explanations of communication failure and the inadequacy of the understood message. Developmental Psychology 13, 156-161.
Sedivy, Julie, 2003. Pragmatic versus form-based accounts of referential contrast: evidence for effects of informativity expectations. Journal of Psycholinguistic Research 32 (1) 3-23.
Sedivy, Julie, Tanenhaus, Michael, Chambers, Craig, Carlson, Greg, 1999. Achieving incremental semantic interpretation through contextual representation. Cognition 71, 109-147.
Shwe, Helen, Markman, Ellen, 1997. Young children's appreciation of the mental impact of their communicative signals. Developmental Psychology 33 (4) 630-636.
Singer, Joyce B., Flavell, John H., 1981. Development of knowledge about communication: children's evaluations of explicitly ambiguous messages. Child Development 52, 1211-1215.
Sonnenschein, S., 1982. The effects of redundant communication on listeners - when more is less. Child Development 53, 717-729. Sonnenschein, Susan, 1986. Development of referential communication skills: how familiarity with a listener affects a speaker's production of redundant messages. Developmental Psychology 22 (4) 549-552.
Sperber, Dan, Wilson, Deirdre, 1986/1995. Relevance: Communication and Cognition, 2nd ed. Blackwell, Oxford.
Vaish, Amrisha, Demir, Özlem Ece, Baldwin, Dare, 2011. Thirteen- and eighteen-month-old Infants recognize when they need referential informations. Social Development 20 (3) 431-449.
Viethen, Jette, Dale, Robert, 2007. Evaluation in natural language generation: lessons from referring expression generation. Traitement Automatique des Langues 48 (1) 141-160.