EXPERIMENTAL EVIDENCE FOR PRESCHOOLERS' MASTERY OF 'TOPIC'

CÉCILE DE CAT

1. Introduction*

Since the 1970's, evidence has suggested that preschool children's encoding of information differs from that of adults in some respects. In their comprehension, preschoolers have been shown to allow target-deviant interpretation of indefinites, which Krämer (2000, 2003) argues is due to insufficient discourse integration. In their production, they have been reported to over-use definite noun phrases in contexts where adults would use indefinites (Maratsos 1974; Karmiloff-Smith 1979; Emslie & Stevenson 1981). This has recently been argued to be due to an immature pragmatic system (Schaeffer & Matthewson 2005). Delayed pragmatic competence has also been invoked to explain various phenomena, from binding errors (Chien & Wexler 1991) to the use of null subjects with finite verbs (Wexler 1998; Schaeffer et al. 2002).

On the other hand, research has also shown that from their first year of life, infants are already able to distinguish new from old information and that language production even at the one-word stage is constrained by a principle of informativeness (Baker & Greenfield 1988). The degree of informativeness of referents, along with other discourse factors, has since been shown to play a significant role in the realisation of arguments in early acquisition (see e.g. Allen 2001, 2007). It has recently been proposed that target-deviance in this area is not caused by a delay in discourse competence, but by the excessive processing demands arising from the *interaction* of discourse and narrow syntax (Serratrice, Sorace & Paoli 2004).

This paper focuses on the acquisition of topic by monolingual children. Topic is a key discourse/pragmatic notion, whose mastery has often been argued to be deficient in early acquisition (see e.g. Chien & Wexler 1991; Wexler 1998; Schaeffer et al. 2002). It is also the interface phenomenon *par excellence*, as it requires the integration of up to three components of the language faculty: discourse/pragmatics, syntax and phonology (De Cat 2007b). Given preschool children's processing limitations and their well-documented impact on language (Gathercole & Baddeley 1993; Reinhart 2004, 2006), one might expect that topic structures be problematic in early acquisition. However, the longitudinal study of early spontaneous production suggests otherwise: topic structures appear from the onset of expressive syntax and children do not make commission errors in the encoding of topics (De Cat 2003, 2007b). In this paper, I present experimental evidence confirming these findings, showing that, at least as early as 2;6, children correctly identify and encode topics.

2. The notion of topic and its acquisition

Topic is a core discourse notion,[1] and is traditionally regarded as a discourse primitive (see e.g. Reinhart 1981; Erteschik-Shir 1997, 2007). The topic of a sentence generally corresponds to what the sentence is about and it provides the referential frame with respect to which the predication is evaluated.

In many languages, topics are expressed as dislocated phrases (Gundel 1975; Reinhart 1981), i.e. they appear at the periphery of the sentence or clause and are typically resumed by a pronominal element

* This research was supported by an AHRC grant (D001099/1), which is gratefully acknowledged. I would like to thank the children who participated as well as their parents and the teachers of the E-J School (Nivelles, Belgium) for their collaboration and their interest.

[1] I will use the label *Discourse* to refer to the component in which topics are identified and interpreted. Depending on one's assumptions regarding the architecture of the language faculty, this could be the Pragmatic component, Information Structure, f-structure,... This has no impact on the argument developed here.

inside the clause, as illustrated in (1).[2] In spoken French, topics are obligatorily dislocated if they are expressed by more than a weak (pronominal) element (Lambrecht 1981; De Cat 2007b).

(1) **Les cochons**$_i$, ils$_i$ se sont enfuis. / Ils$_i$ se sont enfuis, **les cochons**$_i$.
the pigs they REFL are fled they REFL are fled the pigs
'The pigs have fled.'

Because of this, sentences with a "heavy" (i.e. non-pronominal) subject, as in (2), are only possible in contexts where the subject is not topic: either the subject is in narrow-focus, or the whole sentence is in focus. All-focus (or *thetic*) contexts report a state of affairs, and are typically triggered by questions such as *What happened?*

(2) Les cochons se sont enfuis.
the pigs REFL are fled
'The pigs have fled.'

The categorical realisation of subjects as dislocated if topic vs. not dislocated if not topic in spoken French provides the ideal testing ground to study the acquisition of this core discourse notion. So far, the longitudinal study of the spontaneous language productions of 4 children (between the ages of 1;10 and 3;6) has suggested that the linguistic knowledge required for target-like topic encoding is in place from the onset of expressive syntax (De Cat 2003, 2007b). The present study confirms these findings in controlled conditions, for a much larger group of children. It also tries to tease apart the linguistic from cognitive underpinnings of target-like usage.

3. Elicitation study

The aim of this study was to elicit sentences containing a "heavy" NP, in contexts requiring either a topic or a focus interpretation of that NP. If the required linguistic competence is in place, one expects categorical realisation of the former in a dislocated position and of the latter in a structural position associated with focus.

3.1. Design

45 monolingual children participated in the study. There were 3 groups of 15 children. The mean ages were: 2;11 for Group A (2;6.22 - 3;3.28), 3;12 for Group B (3;5.17 - 4;5.28) and 5;2 for group C (4;6.10 - 5;6.15). Children were filmed in their kindergarten.[3] Each child was presented in turn with 13 sets of pictures (7 sets for the topic condition, 6 sets for the focus condition).

In the topic condition, one of the challenges was to summon contexts in which a referent is obligatorily topic (which requires a certain degree of relevance and salience), but not *so* salient that encoding it with a mere clitic could be deemed sufficient. To achieve this, 3 target referents were presented simultaneously each time, so that the use of a clitic alone to identify any of them would be ambiguous. The first picture introduced a group of 3 agents (a family, a group of dogs,...) involved in the same activity. Each member was briefly described by the interviewer and child to individualise them.

[2] All dislocated elements appear in bold in the examples.

[3] Fifteen adults also participated in the study, but out of these, 11 were unfamiliar with the interviewer and this appears to have increased significantly adults' general tendency to use formal French in testing situations. Formal French follows the rules of written French and uses different structural encoding of information structure: it features hardly any dislocation (even of topics) and allows post-verbal subjects in thetic (i.e. all-focus) contexts – something which is normally ruled out in spoken French. The adult data were therefore not comparable with those of the children and had to be discarded. A new technique is currently being piloted to elicit more informal speech from adults and provide a more reliable benchmark for the performance of children in this experiment. For a corpus study of the use of dislocated structures in adult informal French, see De Cat (2007b).

The next page showed each of the agents involved in a different activity[4] and the child was asked: *Qu'est-ce qu'ils font maintenant ?* 'What are they doing now?'. This question aimed to prompt three sentences in which the target referents would be addressed in turn. Because the focus was on them in the previous picture and they were clearly the centre of attention in the test picture, they were relevant enough so that the child would want to predicate something *about* them (forcing a topic reading). And because each target referent was involved in a different activity, there would be a need to distinguish them (as a *contrastive* topic), hence requiring identification by more than just a clitic, i.e. a dislocated NP.

In the focus condition, the first picture introduced a scene (a beach, a room,...), briefly described by the interviewer and child. The second picture showed the same scene, but with a new character in it (a cow on a boat, a child on a bike,...). The child was asked: *Qu'est-ce qui se passe maintenant ?* 'What is happening now?'. This question prompted for a sentence in which all the information would be new, including the target referent (thus forcing a focus interpretation of it).

3.2. Identifying 'errors'

In this experimental setup, a structure is illicit either (i) if it compatible with a topic interpretation of the target referent but is uttered in a context requiring a focus interpretation, or (ii) if it is compatible with a focus interpretation of the target referent but is uttered in a context requiring a topic interpretation.

A topic interpretation is obligatory whenever the referent is encoded in a dislocated structure, as in (3). Dislocated elements were diagnosed using a combination of factors: presence of a resumptive element (a sufficient, but not necessary condition in child French), use of non-nominative pronouns expressing the subject, prosody, apparent disruption of the canonical ordering of constituents. For details and justification, see De Cat (2004, 2006, 2007a).

(3) **Le singe**$_i$, i(l)$_i$ (se) réveille.
the monkey he (REFL) wakes-up
'The monkey wakes up.'

However, if the child considered the referent to be sufficiently salient in the context, (s)he could equally omit the dislocated element and use just a clitic to express the target referent, as in (4). In such cases, children often used deixis to identify which referent they were talking about.

(4) Là, i(l) va dans l'eau.
there he goes in the water
'There, he goes in the water.'

Because many of the children in this sample were still at the null subject stage, some tokens contained a null element associated with the target referent, as in (5). The significance of opting for a (null) clitic to encode the target referent in the topic condition will be discussed in section 4.

(5) 0 fait dodo.
0 makes sleep
'S/he is sleeping.'

A focus interpretation is required in three cases: if the target referent occupies the canonical subject position, as in (6); if it appears on its own in the utterance, as a *fragment* (see Merchant 2004; De Cat & Tsoulas 2006), as in (7); or if it appears in the *il y a* existential construction, as in (8). An existential reading is obligatory in (7) and (8) because the indefinite is mentioned for the first time in the context.

(6) Maintenant la petite fille joue aux balles.
now the little girl plays at-the balls
'Now the little girl is playing with balls.'

[4] Following a suggestion from Peter de Villiers, the picture was cut into three images on the same page, to draw children's attention to each of the three agents in turn.

(7) Une vache!
a cow
'(There's) a cow!'

(8) Il y a une voiture qui arrive.
there is a car that arrives
'A car arrives.'

To sum up, following the rules of adult spoken French (Lambrecht 1994; De Cat 2007b), utterances like (3), (4) and (5) are licit in the topic condition, but illicit in the focus condition, while utterances like (6), (7) and (8) are licit in the focus condition, but illicit in the topic condition.

In addition to these clear-cut cases, there is one type of structure that is hybrid between topic and focus and has to be discarded. The first image of the topic condition introduced three agents involved in the same activity. If the child (or adult) focused on the set rather than the three members (i.e. the family, the group of friends,...) in spite of the interviewer's attempt to individualise them, then it would be the set itself that qualified as a topic in the next picture, and the speaker would need to introduce each individual within that set. Each target referent would therefore be encoded as a d-linked indefinite, as in (9). This informational structure is known as *subordinate update* (Erteschik-Shir 1997): a referent is identified (via focus) among a d(iscourse)-linked set, and is thereby promoted as the sentence topic.

(9) Et i(l) y en a un qui fait dodo.
and there is one who makes sleep
'One of them is asleep.'

Because of the hybrid status of the target referent in such structures, they had to be discarded from the analysis.

3.3. Results

Table 1 presents the results in detail. For each age group, the total number of tokens is given for each of the experimental conditions (topic vs. focus). Tokens are classified (vertically) according to the structural position and nature of the target referent. The numbers in parentheses correspond to illustrative examples in the text. The bottom row represents the number of items for which there was no data: either the child failed to mention one of the target referents (in the topic condition), (s)he did not respond to the question, or the interviewer made a mistake. There is more data in the topic condition because each experimental item included 3 target referents (vs. 1 in the focus condition), and because there were 7 sets of pictures (vs. 6 in the focus condition).

Age group	**Group A**		**Group B**		**Group C**	
Condition	**Topic**	**Focus**	**Topic**	**Focus**	**Topic**	**Focus**
1. Dislocated NP[5] (3)	94	0	120	1	173	0
2. Subject clitic (4)	94	3	108	1	33	0
3. Null subject (5)	29	1	32	1	0	0
4. Heavy subject (6)	0	10	3	6	3	9
5. Fragment (7)	2	42	0	40	0	24
6. *Il y a* / cleft + existential indefinite (8)	11	19	0	34	0	54
7. *Il y a* + definite	5	3	3	1	8	0
8. *Il y a* + d-linked indefinite (9)	0	0	15	0	77	0
9. *Il y a* + unclear	8	5	1	0	1	0
10. Unclear	2	4	2	0	0	0
Total	245	87	284	84	295	87
+ Void items	68	3	31	6	20	3

Table 1: Encoding of the target referent according to structural position and nature

[5] All the dislocated NPs produced by the children were definite.

The light shaded cells indicate adequate encoding of information structure: the target referent was encoded as a topic in the topic condition (rows 1-3), or as a focus in the focus condition (rows 4-6). The dark shaded cells indicate the 'errors': either the target referent was encoded as a topic in the focus condition (as in the first 3 rows), or it was encoded as a focus in the topic condition (as in rows 4-6). Note that the 11 tokens in the first cell of row 6 were all produced by one child (age 2;7.14), whose performance was otherwise similar to that of her age group. I come back to these 11 tokens in section 3.4. All the shaded cells correspond to the utterances in which the information structural status of the target referent is unambiguously topic or focus.

Notice the striking parallel between the use of null subjects and subject clitics to encode the target referent. Null subjects disappear in Group C, at the same time as the use of subject clitics in the topic condition drops dramatically. The use of clitics and null subjects is almost exclusive to the topic condition, similarly to dislocated NPs. I will discuss the implications of using a clitic instead of a dislocated NP in section 4.

The d-linked indefinites (row 8) have to be discarded due to their hybrid status, as explained in section 3.2. Note that all the data in this row are of the type *Il y en a un qui...* 'There is one (of them) that...', where the d-linked set is implied by the use of the partitive pronoun *en* 'of them'. This is one of the latest acquired clitics (Hamann, Rizzi & Frauenfelder 1996), which might explain the absence of such tokens from Group A's data. Also discarded were *Il y a* structures introducing a noun phrase with a missing or unintelligible determiner (as it was impossible to decide whether it was definite or not). There were also a few cases where most of the utterance was unintelligible. Hence the data in rows 8-10 have to be excluded from the counts.

What is less clear is whether the structures of type '*Il y a* + definite NP' (row 7) should be treated on a par with row 6 or row 8. According to Lambrecht (1994), such structures are thetic (i.e. all the information they convey is new) and the difference between these and sentences with a heavy subject is that the referent introduced by the *il y a* clause is "treated as an identifiable entity in the discourse" (Lambrecht 1994:158). However, because of the thetic reading, that referent is also new information. A topic interpretation is therefore ruled out. Only d-linked indefinites in that position can trigger the creation of a topic via subordinate update. Therefore the use of this structure in the topic condition should count as an 'error'. The use of a definite in this structure is also unexpected in the focus condition, but turns out to be licit: the 4 instances of *il y a* + definite NP in that condition are all licensed via bridging. Bridging allows the use of a definite on first-mention via inference from a combination of knowledge available from the context and knowledge of the world (see e.g. Hawkins 1978; Avrutin & Coopmans 2000). For instance, in (10), the child assumed that the squirrel that appears on the test picture belongs to the child introduced in the previous picture. This licenses its encoding as a (possessive) definite even on first mention.

(10) I(l) y a son écureuil qui est là.
there is his squirrel who is there
'His squirrel is there.'

Table 2 summarises the results for the subset of data including all clear cases (corresponding to rows 1-6 in Table 1) plus the *il y a* + definite NP structures (row 7).

Age group	**Group A**		**Group B**		**Group C**	
Condition	**Topic**	**Focus**	**Topic**	**Focus**	**Topic**	**Focus**
Target encoded as topic (cf. rows 1-3)	217 (92%)	4 (5%)	260 (98%)	3 (4%)	206 (95%)	0 (0%)
Target encoded as focus (cf. rows 4-7)	18 (8%)	74 (95%)	6 (2%)	81 (96%)	11 (5%)	87 (100%)
Total	235	78	266	84	217	87

Table 2: Encoding of target referent as topic or focus in each condition

The diagonal effect is so clear in Table 2 that statistics are hardly needed to show that, for each age group, the difference between the use of a topic vs. a focus construction in each condition is most probably not due to chance. Chi-square tests returned the following values: Group A: $\chi^2 = 210.44$, $p < .0001$; Group B: $\chi^2 = 298.09$, $p < .0001$; Group C: $\chi^2 = 251.87$, $p < .0001$.

3.4. Summary and preliminary conclusions

In the topic condition, children used almost exclusively topic-compatible encoding of target referents. The error margin was of 5% (35/718 tokens). 'Errors' included 2 target referents encoded as fragments, 6 encoded as heavy subjects, 16 encoded as *il y a* + definite NP and 11 encoded as *il y a* + existential indefinite (all from the same child from Group A). This could be an attempt by that child to use a d-linked indefinite structure (as is relatively frequently used by the older age groups) when the clitic *en* has not yet been acquired. (Recall that all d-linked indefinites produced by children in Groups B and C are as in (9), with *en* summoning the d-linked set.) If this was the case, these 11 tokens should be discarded from the counts, on a par with other d-linked indefinites (see section 3.2). The error margin would come down to 3% (7/235) for that age group. The slightly higher number of errors in Group C's topic condition is due to an increase in the use of *il y a* + definites.

In the focus condition, children used almost exclusively focus-compatible encoding of target referents. The error margin was of 3% (7/249 tokens). 'Errors' included 1 target referent encoded as dislocated, 4 encoded as subject clitics and 2 as null subjects.

All the children, even the youngest, used dislocated NPs to encode the target referent in the topic condition. There was only 1 instance where a dislocated NP was erroneously used in the focus condition. Children never used existential indefinites (nor indeed any indefinite) in a dislocated position, confirming that they are 'aware' of the topic status of dislocated NPs and of the fact that existential indefinites can never be topics (Reinhart 1981).

This experiment thus brings overwhelming evidence for the mastery of the discourse/pragmatic notion of topic by preschool children, even in the youngest age group.

4. Children's evaluation of salience

A closer look at Table 1 reveals that the children in Groups A and B were relatively reluctant to use full NPs to encode the target referent. A first interesting observation is that children only avoided full NPs when it was informationally licit to do so: apart from a tiny error margin (2%, i.e. 6/258 tokens), they never used a mere clitic or a null subject to encode the target referent in the focus condition. This demonstrates their awareness of the information status of referents: they (almost) always encoded new, focused referents as full NPs. Within the topic condition, the younger children relied extensively on the 'reduced' option, expressing the target referent with either a clitic or a null subject more than half of the time. In Group A, children encoded topics with a full NP (dislocated) only 43% of the time (94/217). In Group B, they did so 46% of the time (120/260). In Group C, the proportion climbed to 84% (173/206).

In the adult language, topics can be encoded with just a clitic when their referent is sufficiently salient in the context. A referent is sufficiently salient when the hearer can identify it effortlessly in the discourse or situational context. Salience is evaluated by the speaker on the basis of their assumptions regarding the hearer's state of knowledge and beliefs. In the present experimental conditions, there were 3 target referents in each case, and the use of a simple clitic to identify them would in principle have led to ambiguity for the hearer. Children may have been aware of this ambiguity to an extent, because they often used deixis (i.e. pointing, often accompanied by a locative adverb) to identify each of the referents in turn, as they described the activity they were involved in. Only a relatively small proportion of clitics or null subjects lacked any form of disambiguation: 12% (26/217) in Group A, 23% (60/260) in Group B and 9% (19/206) in Group C. However we have no proof that the *intention* behind the pointing was indeed disambiguation for the sake of the hearer. To be able to assess their listener's needs (and therefore evaluate referents' salience), children have to be able to appreciate that (s)he might have a different perspective from their own. This ability is acquired as part of Theory of Mind development.

5. Salience evaluation and Theory of Mind

A Theory of Mind test based on O'Neill (1996) was carried out on the same children immediately after the topic experiment. The test required children to answer 'yes' or 'no' to some questions, which were kept as simple as possible (i.e. without multiple or d-linked *wh*-elements).

The child was invited to choose two animal puppets, and asked to identify them (e.g. as 'the lion' and 'the monkey'). One of them (e.g. the lion) felt tired and went to bed, where it stayed, asleep, for the

duration of the test. A box was introduced, and the active puppet looked inside. In the first condition, the child was also allowed to look inside the box. In the second condition, the child could not see what was in the box. Then (in both conditions) the child was asked the following questions in turn: Did the lion see what is inside the box? Did the monkey see what is inside the box? Does the lion know what is in the box? Does the monkey know what is in the box?

The detail of the results is given in Tables 3 and 4. Some children had to be left out of one or both of the experimental conditions because of unclear responses (this explains the small differences in numbers between the groups across conditions).

Errors on...	2;6 – 2;11		3;0 – 3;5		3;6 – 3;11		4;0 –	
1. Active puppet has seen	0/8		0/8		0/5		1/21	(5%)
2. Active puppet knows	0/8		1/8	(13%)	0/5		1/21	(5%)
3. Sleeping puppet has seen	4/8	(50%)	3/8	(38%)	0/5		0/21	
4. Sleeping puppet knows	6/8	(75%)	4/8	(50%)	5/5	(100%)	1/21	(5%)
Either 3 or 4	8/8	(100%)	6/8	(75%)	5/5	(100%)	1/21	(5%)

Table 3: Number of children who gave the wrong answer on a ToM question, by age group, in the first condition: the child has seen inside the box

Errors on...	2;6 – 2;11		3;0 – 3;5		3;6 – 3;11		4;0 –	
1. Active puppet has seen	1/8	(13%)	0/7		0/6		1/21	(5%)
2. Active puppet knows	1/8	(13%)	0/7		2/6	(33%)	1/21	(5%)
3. Sleeping puppet has seen	4/8	(50%)	2/7	(29%)	0/6		2/21	(10%)
4. Sleeping puppet knows	6/8	(75%)	3/7	(43%)	5/6	(83%)	1/21	(5%)
Either 3 or 4	8/8	(100%)	3/7	(43%)	6/6	(100%)	2/21	(10%)

Table 4: Number of children who gave the wrong answer on a ToM question, by age group, in the second condition: the child has *not* seen inside the box

All the children but one[6] could correctly judge whether the active puppet had looked inside the box. From 3;6, all children could correctly make that judgement about the sleeping puppet.[7] The lateness of this threshold is likely to be due to the verbal nature of the test. Non-verbal tests have shown that 2-3 year olds monitor whether another person sees something happen or not O'Neill 1996. After the age of 4, 18 (out of 21) children passed the seeing-leads-to-knowing test in both conditions, indicating that they can (but most of their age group cannot) make an accurate judgement regarding the knowledge state of both puppets, irrespective of their own knowledge state.[8] Before 4;0, only 2 children (out of 20) passed the test in both conditions. The threshold of 4;0 corresponds to the age at which children are generally reported to master false belief, the definite test of mental-state understanding (Wellman, Cross & Watson 2001).

The data was reorganised in 2 groups (Group 1 under 4;0 and Group 2 over 4;0 – see Table 5), so that Group 2 would include children who have the prerequisite cognitive abilities to evaluate the knowledge state of their listener and therefore can appreciate that (s)he may need help to identify the referents they are talking about. In other words, children in Group 2 should be able to evaluate the salience of discourse referents, based on their Theory of Mind development. And indeed, they do use full NPs to encode topics much more often than children in Group 1. Yet we find that Group 2 still do use just clitics to identify the target referents in the topic condition 29% of the time (96/329), without deixis to help the listener in 17% of cases (57/329). This suggests that they fail to appreciate that using a clitic is not sufficient to let their listener know which of the three target referents they are talking about. Theory of Mind thus cannot be the determining factor underlying the use full forms (vs. clitics) to meet the salience requirement.

[6] That child was 5;4.12. He acted very confidently during the test but systematically gave the wrong answer.

[7] There are two exceptions: the child above (5;4.12) and another child (5;5.25), although she did answer correctly in one of the conditions.

[8] It is not exactly the same group of children who pass the test in condition 1 and in condition 2, which is why only 18 and not 19 children are mentioned here (contrary to what the tables seem to suggest).

Age group	Group 1 (under 4;0 – 23 children)		Group 2 (4;0 and above – 22 children)	
Condition	**Topic**	**Focus**	**Topic**	**Focus**
Dislocated NP	154 (44%)	1	233 (71%)	0
Subject clitic	145 (31%)	4	90 (27%)	0
Null subject	55 (16%)	2	6 (2%)	0
Total of topics	354	7	329	0

Table 5: Target referents encoded as topics in each condition, by different age grouping

Note that what is at stakes here is quite subtle. All the children in this study *are* able to evaluate the salience requirement on topics to an extent, as they only encode as topics those referents that have been introduced in the previous picture (i.e. they never encode new referents as topics). What they fail to appreciate, especially before 4;6, is that a heavy NP is needed to clearly distinguish the three target referents in each case. Instead of relying on the (linguistic) discourse to resolve the ambiguity between the three referents, children tend to rely on the physical context. By exploiting joint attention, they bank on what is visible to them and their addressee to reduce the amount of information encoded linguistically. This is evident in their frequent use of deixis to single out the referent in question. Initially, the physical context seems to be their preferred (sub-)domain of reference (an idea proposed by Krämer 2005), and they only express linguistically what is essential from an information structural point of view.

6. Conclusions

This study clearly shows that preschool children as young as 2;6 already possess the discourse/ pragmatic competence necessary to encode topics. This requires them to be able to evaluate the information status of discourse referents (are they new or given), their relevance (does it make sense to predicate something about the chosen topic), and, up to a point, their salience (are they identifiable in the context).

To the extent that the Theory of Mind test used in this research accurately identifies when children become able to appreciate differences between their own perspective and that of their addressee, it can be postulated that the notion of topic is in place ahead of the cognitive milestones required to be aware of the point of view of others. However, an *implicit* understanding of differences in perspective might already have been in place even in the youngest children, so one has to remain speculative here.

By the age of 4, children's awareness of other minds is established, and yet children older than 4 were still found to often use clitics rather than full NPs to encode topics that were not quite salient enough in the context to avoid ambiguity. Theory of Mind can therefore not be the determining factor in the evaluation of subtle salience distinctions. It was proposed that children rely maximally on joint attention to minimise what to express with overt syntax. But crucially, this (almost) never violates the linguistic rules of topic encoding.

The early mastery of the discourse notion of topic, at an age of well-documented processing limitations, implies that the integration of information from syntax and discourse is not *per se* too demanding for young children.

7. References

Allen, S. (2001). The importance of discourse-pragmatics in acquisition. *Bilingualism: Language and Cognition,* 4, 23-25.

Allen, S. (2007). Interacting pragmatic influences on children's argument realisation. In M. Bowerman & P. Brown (eds.), *Crosslinguistic Perspectives on Argument Structure: Implications for learnability.,* pp. 191-210. Mahwah, NJ: Erlbaum.

Avrutin, S. & Coopmans, P. (2000). Children Who Build Bridges. *Proceedings of the Annual Boston University Conference on Language Development,* 24 (1), 80-91.

Baker, N. & Greenfield, P. (1988). The development of new and old information in young children's early language. *Language Sciences,* 10 (1), 3-34.

Chien, Y.-C. & Wexler, K. (1991). Children's knowledge of locality conditions in Binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition,* 1, 225-295.

De Cat, C. (2003). Syntactic manifestations of very early pragmatic competence. In B. Beachley, A. Brown & F. Conlin (eds.), *Proceedings of BUCLD27*, pp. 209-219. Somerville, MA: Cascadilla Press.

De Cat, C. (2004). Apparent non-Nominative subjects in L1 French. In J. Paradis & P. Prévost (eds.), *The Acquisition of French in Different Contexts: Focus on Functional Categories*, pp. 60-115. Amsterdam: John Benjamins.

De Cat, C. (2006). Do Root Infinitives ever have overt subjects in child French? In T. Parodi, A. Sorace, S. Unsworth & M. Young-Scholten (eds.), *Paths of development in L1 and L2 acquisition. In honor of Bonnie D. Schwartz*, pp. 55-76. Amsterdam: John Benjamins.

De Cat, C. (2007a). French dislocation without movement. *Natural Language & Linguistic Theory*, 25 (3), 485-534.

De Cat, C. (2007b). *French Dislocation: Interpretation, Syntax, Acquisition.* Oxford: OUP.

De Cat, C. & Tsoulas, G. (2006). The structure of fragments in (child) French. In A. Belleti, E. Bennati, C. Chesi, E. D. Domenico & I. Ferrari (eds.), *Language Acquisition and Development. Proceedings of GALA 2005*, pp. 142-147. Cambridge: Cambridge Scholar Press.

Emslie, H. & Stevenson, R. (1981). Pre-school children's use of the articles in definite and indefinite referring expressions. *Journal of Child Language*, 8 (2), 313-328.

Erteschik-Shir, N. (1997). *The Dynamics of Focus Structure.* Cambridge: Cambridge University Press.

Erteschik-Shir, N. (2007). *Information Structure. The Syntax-Discourse interface.* Oxford: OUP.

Gathercole, S. & Baddeley, A. (1993). *Working Memory and Language.* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Gundel, J. (1975). Left dislocation and the role of topic-comment structures in linguistic theory. *OSU WPL*, 18, 72-131.

Hamann, C., Rizzi, L. & Frauenfelder, U. H. (1996). On the Acquisition of Subject and Object Clitics in French. In H. Clahsen (ed.), *Generative Perspectives on Language Acquisition: Empirical Findings, Theoretical Considerations and Crosslinguistic Comparisons*, pp. 309-34. Amsterdam: Benjamins.

Hawkins, J. (1978). *Definiteness and indefiniteness.* London: Croom Helm.

Karmiloff-Smith, A. (1979). *A functional approach to child language. A study of determiners and reference.* Cambridge: CUP.

Krämer, I. (2000). Interpreting Indefinites. An experimental study of children's language comprehension. Doctoral Dissertation. University of Utrecht.

Krämer, I. (2003). Reference of indefinite and pronominal noun phrases in a story context: English children's comprehension. *Proceedings of BUCLD27*, pp. 449-460. Somerville, MA: Cascadilla Press.

Krämer, I. (2005). Egocentrism, deixis or conflict of domains? Lack of anaphoricity in child language. Talk presented at the workshop "Language Acquisition Between Sentence and Discourse", May 2005, Radboud Universiteit Nijmegen.

Lambrecht, K. (1981). *Topic, antitopic and verb agreement in non-standard French.* Amsterdam: John Benjamins.

Lambrecht, K. (1994). *Information Structure and Sentence Form. Topic, focus, and the mental representation of discourse referents.* Cambridge: CUP.

Maratsos, M. (1974). Preschool Children's Use of Definite and Indefinite Articles. *Child Development*, 45, 446-455.

Merchant, J. (2004). Fragments and ellipsis. *Linguistics and Philosophy*, 6, 661-738.

O'Neill, D. (1996). Two-year-old children's sensitivity to a parent's knowledge state when making requests. *Child Development*, 67 (2), 659-677.

Reinhart, T. (1981). Pragmatics and linguistics: An analysis of sentence topics. *Philosophica*, 27 (Special issue on Pragmatic Theory), 53-94.

Reinhart, T. (2004). The Processing Cost of Reference Set Computation: Acquisition of Stress Shift and Focus. *Language Acquisition*, 12 (2), 109-155.

Reinhart, T. (2006). *Interface Strategies. Optimal and Costly Computations.* Boston, MA: MIT Press.

Schaeffer, J., Gordishevsky, G., Hadar, G. & Hacohen, A. (2002). Subjects in English-speaking Children with Specific Language Impairment. In J. Costa & M. J. Freitas (eds.), *Proceedings of the GALA 2001 Conference*, pp. 223-231. Lisboa: Associação Portuguesa de Linguística.

Schaeffer, J. & Matthewson, L. (2005). Grammar and Pragmatics in the Acquisition of Article Systems. *Natural language and Linguistic Theory*, 23, 53-101.

Serratrice, L., Sorace, A. & Paoli, S. (2004). Transfer at the syntax-pragmatics interface: Subjects and objects in Italian-English bilingual and monolingual acquisition. *Bilingualism: Language and Cognition,* 7, 183-206.

Wellman, H., Cross, D. & Watson, J. (2001). Meta-analysis of Theory of Mind development: The truth about false belief. *Child Development*, 72 (3), 655-684.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua,* 106, 23-79.