

Syntactic manifestations of very early pragmatic competence

Cécile De Cat
University of York

1 Introduction

This paper demonstrates that very young children have the required knowledge/competence necessary to encode topics. Evidence to that effect is provided from the spontaneous speech of five children acquiring French, a language where topics (when they are expressed by a non-weak element) are straightforwardly identifiable from overt syntax and prosodic cues. A detailed description of the basic rules of topic encoding in the adult language is used as a benchmark.

2 Background

2.1 On topics

In the vast literature on topics, a confusing amount of definitions are proposed. The one that will be adopted for the purpose of this article is as general and wide-ranging as possible, to reflect the fact that the phenomenon under study is most probably universal, although different languages resort to different means of encoding it. Topics will be taken to express *what the sentence is about* (following Reinhart 1981), hence the label *aboutness topic*.¹

Cross-linguistically, aboutness topics strongly tend to be encoded as dislocated elements — as bolded in (1) (Gundel 1975; Reinhart 1981; Rizzi 1997).

* This research was supported by the Economic and Social Research Council (grants R00429834373 and 041R00433), which is gratefully acknowledged. I wish to thank David Adger, Aafke Hulk and Bernadette Plunkett for comments on an earlier version of the manuscript as well as the Acquisition groups at Groningen and Utrecht for discussion.

¹This is intended to avoid confusion with another kind of topic which needs not concern us here. The reader is referred to De Cat (2002) for a definition of *stage topics* (based on Erteschik-Shir 1997) and its manifestations in spoken French. In De Cat (2002), I argue in favour of an overarching definition of topic-hood in terms of *frame within which the predication is evaluated* (following Erteschik-Shir 1997). The point of the present paper, i.e. that children appear to have the required competence to encode topics from the onset of the multiword stage, holds for topics in general, though it is only explicitly made here with respect to aboutness topics.

- (1) **Les frites**_{*i*}, *c'*_{*i*} est bon.
 the fries it is good
 'Fries are nice.'

Aboutness topics correspond in most cases to the subject of the sentence/clause (Gundel 1975; Li & Thompson 1976; Givón 1976; Reinhart 1981; Lambrecht 1994). A native speaker of English, if presented with a sentence like (2), will by default interpret the subject (here *your barrels*) as the topic. In other words, the context in which a sentence like (2) is felicitous is by default one in which the barrels in question are salient (and these barrels are what is being talked about).

- (2) Your barrels are in the shed.

Thetic sentences are the only sentence type that is incompatible with an aboutness topic. Such sentences express a state of affairs and are typically uttered out-of-the-blue or after a question like *What happened?*. In thetic sentences, a topic interpretation of e.g. the subject or the object is impossible — hence the dislocation of these elements is not allowed in spoken French. This will be illustrated in the following section.

2.2 Topic encoding in (adult) spoken French

In spoken French, the association between dislocated elements and topic interpretation is extremely robust. Topics are obligatorily dislocated (except when they are expressed by a weak pronoun only), and dislocated elements are obligatorily interpreted as topics.

2.2.1 Heavy topics are obligatorily dislocated

In spoken French, topics are obligatorily (left- or right-)dislocated when they are not expressed by a weak pronoun only (De Cat 2002). The presence of a resumptive clitic such as *c'* 'it' in (1) can be used as a (sufficient, though not necessary) diagnostic for dislocation in that language. The DP *tes fûts* 'your barrels' can be interpreted as a topic in (3) but not in (4).

- (3) Ils sont dans la remise, **tes fûts**.
 they are in the shed your barrels
 'Your barrels are in the shed.'

Heavy elements (i.e. elements other than weak pronouns) in the canonical subject position are obligatorily in focus in spoken French.² This is the case either if the sentence is thetic (as in (4)) or if there is a narrow focus on the subject (as in (5), where the main stress falls obligatorily on *aussi* 'too').

²I take the focus to correspond to the most informative part of the sentence, identified by the most prominent stress (Cinque 1993; Reinhart 1996).

- (4) [F Tes fûts sont dans la remise]. (out-of-the-blue only)
 your barrels are in the shed
 ‘Your barrels are in the shed.’
- (5) [F Les autres aussi] veulent venir.
 the others too want to-come
 ‘The others want to come too.’

Evidence to that effect can be found in Individual Level Predicates. I will adopt as a working definition that such predicates (ILPs) express permanent properties (Milsark 1974).³ ILPs cannot appear inthetic sentences (De Cat 2002). Consequently, a DP expressing the subject of an ILP is obligatorily dislocated in spoken French, unless there is a narrow focus on the subject. In (6), both A’s statement and B’s rectification contain an ILP. In B’s utterance, there is a narrow (contrastive) focus on the DP expressing the subject. This DP is thus forced to appear in the canonical subject position (where the contrastive focus reading is allowed). Yet the property reading is allowed in both utterances. In A’s statement, the topic is *mon voisin* ‘my neighbour’ and in B’s it is a covert aboutness topic corresponding to something like (*horrible*) *people in A’s neighbourhood*.

- (6) A: **Mon voisin**, c’ est une crapule.
 my neighbour_M it is a ruffian
 ‘My neighbour is a ruffian.’
- B: Non, TA VOISINE est une crapule.
 no your neighbour_F is a ruffian
 ‘My neighbour is a ruffian.’

2.2.2 Dislocated elements are obligatorily topics

Dislocated elements are obligatorily interpreted as topics in spoken French. As a consequence, existential indefinites are banned from the dislocated position in that language (as illustrated in (7-a)). The existential reading is incompatible with a topic interpretation because topic referents must by definition be identifiable in the (discourse) context. That being said, nothing prevents non-existential indefinites from being dislocated. In (7-b), for instance, the dislocated indefinite is interpreted generically, which makes it acceptable as a topic. This sentence is interpreted as transcending particular facts: it is not about a particular child but about a typical behaviour of children. In (7-a), by contrast, the generic reading is not available (due to the tense specification which forces a specific reading) and the sentence is unacceptable as a result.

- (7) a. #**Un enfant**, il est arrivé pi il t’ a posé une question.
 a child he is arrived then he to-you has asked a question

³Things are more complex than I am suggesting in the text. For an in-depth discussion of the properties of ILPs, see Jäger (2001).

- b. Tsé, **un enfant**, il arrive pi il te pose une question
 y'know a child he arrives then he to-you asks a question
 'You know, a child comes and asks you something' (Auger 1994)

3 Signs of early pragmatic competence

I will now endeavour to demonstrate that very young children show clear signs of having the required competence for the encoding of topics. The data used for this purpose comes from two longitudinal corpora following the acquisition of French by five children from different dialectal areas: Belgium (Léa, Tom, Chloé), Canada (Max) and France (Anne). Data from Léa, Max and Anne are part of the York corpus and data from Tom and Chloé are part of the Cat corpus.⁴ The data was collected fortnightly (York corpus) or monthly (Cat corpus) in an environment familiar to the child, over a period of 18 months in average. Transcription and checking were done by native speakers. Data from 12 adults interacting with the children were also used for the present analysis.

Dislocated elements in the child data were diagnosed on the basis of the presence of resumptive elements, the prosody and apparent disruption of the canonical ordering of constituents (see De Cat 2002 for details, including a preliminary prosodic analysis of the dislocation data in adult and child French).

A first observation of the data reveals that the children's dislocated elements are always compatible with a topic interpretation, even when the utterance is not fully target-like. Even when the dislocated elements fail to be resumed by a weak pronoun (as in (8-a)) or occur in verbless utterances (as in (8-b)), they still correspond to the topic of the sentence/utterance.⁵

⁴The York corpus was collected under the direction of Bernadette Plunkett (ESRC grant #R000 22 1972), who has kindly allowed me to use it. The Cat corpus was collected for my doctoral research, which was funded by ESRC grant R00429834373. Details on transcription and coding procedures can be found in De Cat (2002) and De Cat & Plunkett (2002).

⁵The transcription conventions used in the examples reported in this paper are as follows: (i) Missing elements are indicated either by 0 (in cases allowing multiple possibilities) or by bracketing (where the missing element was clearly identifiable). (ii) *Là* with a capital is deaccented and is interpreted as a situational deictic rather than as a true locative. (iii) The hash # sign indicates a short pause. (iv) E stands for a verb ending ambiguous between infinitive (-er) and participial (-é) morphology. (v) False starts with or without correction are indicated by [//] and [/] respectively. (vi) xx correspond to an unintelligible word. (vii) The comment in square brackets introduced by =! describes what the child does to enact the (omitted) object of her sentence. (viii) Interrupted words are marked by an ampersand &. (ix) Double commas indicate a tag. (x) A *e* in the transcription stands for an embryonic element (e.g. a proto-(clitic) pronoun). (xi) Alternative

- (8) a. 0 est pas belle, **le bébé**. (Anne 2;0.27)
 is not pretty the baby
 ‘The baby isn’t pretty.’
 b. (le) minou (à) Yolande, **ça**. (Max 1;11.0)
 the cat to Yolande that
 ‘(That’s) Yolande(’s) cat.’

In sections 3.1-3.3, I point to clear indications that very young children have the competence required for the target-like encoding of topics.

3.1 Children comply by the ‘ILP requirement’

Testing whether children abide by the requirement that the heavy element expressing the subject of ILPs be dislocated except if it is in focus can be done by applying the following reasoning:

Only heavy elements expressing the subject are targeted by the ‘ILP requirement’. So only sentences containing such a heavy element (whether dislocated or in the canonical subject position) need concern us here, not those where the subject is expressed by a weak pronoun only or is missing (on the uncontroversial assumption that missing subjects are missing pronouns). What needs to be shown is that

1. children show clear signs of compliance with the ‘ILP requirement’, i.e. they do dislocate heavy elements expressing the subject of ILPs, and
2. children do not violate the ‘ILP requirement’, i.e. they do not utter sentences containing an ILP with a heavy subject unless there is a clear narrow focus on that subject.

The sentences in (9) prove the first point. In all cases, a property reading of the predicate was clearly intended.

- (9) a. **les sucettes**, ça finit pas. (Anne 3;1.15)
 the lollipops it finishes not
 ‘Lollipops don’t end.’
 b. **le poisson Là, ça**, c’ est un poisson. (Max 2;5.1)
 the fish there that it is a fish
 ‘That fish is a fish.’
 c. **ça**, c’ est un dinosaure. (Max 2;8.9)
 that it is a dinosaur
 ‘That one is a dinosaur.’
 d. **le coca**, ça soigne le hoquet aussi. (Léa 2;9.5)
 the coke it cures the hiccough too
 ‘Coke cures hiccough too.’

transcriptions are given in square brackets, after a =? sign. (xii) Speakers are designated by a three letter code. E.g. *TAT is Anne’s child minder.

- e. et **les vaches**, # elles mangE # de l' herbe. (Tom 2;1.13)
 and the cows they eat some the grass
 'And cows eat grass.' (Clearly generic context)

The corpora were then searched for heavy subjects. A total of 186 cases was found in the child data. Only one case (given in (10)) involves an ILP, and it is target-like: the subject is in narrow (contrastive) focus. The other cases are clearly not ILPs, and typically contain eventive predicates, as in (11).

- (10) mon papa aussi est gentil. (Anne 2;8.3)
 my dad too is nice
 'My dad's nice too.'
- (11) a. ah et tout [/] tout le monde est tombé. (Anne 2;7.1)
 ah and all all the people is fallen
 'Ah everybody's fallen over.'
- b. (la) vache # mange # (des) ca(r)ottes. (Max 2;5.29)
 (the) cow eats (some) carrots
 'The cow's eating carrots.' (Clear ongoing interpretation)
- c. Parrain, Luc s' est fait mal. (Léa 2;8.22)
 grandad Luc REFL is done pain
 'Grandad, Luc's hurt himself.'
- d. et pourquoi # les lunettes s' étaient xx envolées? (Tom 2;6.12)
 and why the glasses REFL were ? flown
 'And why did the glasses fly away?'

Children's spontaneous production thus clearly suggests that they know that ILPs cannot appear inthetic sentences, which requires for them to be able to identify and encode topics in a target-like fashion.

3.2 Children do not dislocate existential indefinites

Another indication that children have the required competence to encode topics adequately comes from the absence from their spontaneous production of dislocated indefinites requiring an existential interpretation.

The overwhelming majority of dislocated indefinites I have found in these five children's spontaneous production clearly required a generic interpretation (as in (12)) or a D-linked one (as in (14)). D-linked indefinites refer to a set previously established in the discourse, which makes them possible topic candidates. Note that the example in (12-b) is not target-like: a plural would be expected, with a dummy definite determiner (*les gâteaux aux fraises* 'the cakes with strawberries'). Yet, the indefinite is clearly used with an intended generic reading and therefore meets the requirement for encoding as a topic. This is confirmed in the immediately following context, where the mother asks for clarification and is answered by the child uttering the clearly generic sentence in (13).

- (12) a. parce que c' est blanc, **un ours**. (Max 2;9.27)
 because it is white a bear
 'Because bears are white.'
- b. et moi, j' aime bien, **un gâteau aux fraises**. (Chloé 2;11.24)
 and me I love well a cake to-the strawberries
 'I like strawberry cake(s).'
- (13) j' ai dit: moi, j' aime bien les # aux fraises.
 I have said me I love well the to-the strawberries
 'I said I liked the strawberry ones.'

In (14-a), the child is talking about a set of cars that has just been fixed by the interviewer; in (14-b), the child is referring to a group of animals wanting to enter the zoo enclosure; in (14-c), she is talking about the members of a family she has put to sleep under a leaf.

- (14) a. **toutes les autos**, elles sont réparées. (Max 2;4.18)
 all the cars they are fixed
 'All the cars are fixed.'
- b. **les copains, tous les copains**, i(ls) veulent entrer. (Anne 2;7.1)
 the friends all the friends they want to-enter
 'All the friends want to come in.'
- c. **tout le monde**, i(l) dort. (Anne 2;8.20)
 all the people he sleeps
 'Everybody's sleeping.'

The only dislocated indefinites not compatible with a generic or a D-linked interpretation are given in (15), (17) and (18). However, there are good reasons to believe that the use of indefinites here is due to performance errors or the use of unanalysed chunks by the child.

The context immediately preceding (15) is given in (16), where *CAT is the interviewer. Given this context, and that the child is looking at the little baby in question as he utters (15), it is clear that a specific reading is intended, not an existential one.

- (15) **un petit bébé**, je sais pas c' est quoi. (Max 2;7.25)
 a little baby I know not it is what
 'I don't know what the little baby's called.'
- (16) *CAT: tu as guéri le petit bébé?
 'You cured the little baby?'
- *CAT: comment il s' appelle, **le petit bébé**?
 'What is the little baby called?'

The sentence in (17) is clearly a performance error, as it is immediately followed by a self-correction from the child.

- (17) **des cochons**, ils font 0 [=! she swirls]. (Léa 3;0.5)
 some pigs they go
 ‘Pigs go...’ (She demonstrates what pigs do)

As for the dislocated indefinites in (18), they were prompted by adult speech. In (18-a), the child does not seem to be able to identify what *de l’herbe* ‘grass’ refers to on the picture she’s been looking at. The context in which (18-b) is uttered is one in which the child has been asked to place face attributes (eyes, mouth, moustache etc) back onto a sheet, and is highly confused as to what the adults expect from her. In both cases, the intended interpretation is clearly not existential.

- (18) a. c’ est où, **le l’ &her** [/] **le l’ herbe** [/] **de l’ herbe**?
 it is where the the gra- the the grass PART the grass
 ‘Where’s “the grass”?’ (Anne 2;6.2)
 b. oui, **un** [/] **un oeil**, (ç)a va où?
 yes a an eye it goes where (Max 2;4.4)
 ‘Where does (that) eye go?’

I conclude that children’s spontaneous production gives no indication that they ever dislocate existential indefinites. This does not appear to be due to a sampling artifact, given that dislocated indefinites do appear in their speech (though in accordance with target usage).

3.3 Target-like use of left- vs. right-dislocation

Left- and right-dislocated elements both express the topic of the sentence, but they do so in a slightly different manner. Essentially, these differences stem from a greater (prosodic and structural) saliency of the left periphery (De Cat 2002). As a result, some nuances are available to left-hand topics but not to their right-hand counterparts.

Wh-questions are an environment in which the left- vs. right-dislocation contrast is most visible: the majority of dislocated elements coreferential with the subject (henceforth dislocated subjects) appears in the right-periphery in such questions. Significantly, the very same proportion of right-dislocated subjects is observed in adult speech (78%)⁶ and in child speech: In the speech of Max, Anne and Tom between the ages of 1;9 and 2;6, dislocated subjects in *wh*-questions appear in the right-periphery in 76% of cases in average. This is shown in Table 1. Very similar results are obtained even if this period of development is broken into two (with Period 1 corresponding to the core of the null subject stage, when subjects are omitted in at least 15% of obligatory contexts).

⁶This is the proportion observed in a random sample of 5613 utterances from the adult speakers of the York and the Cat corpora, out of which 479 are root *wh*-questions.

	Left-dislocation		Right-dislocation	
Period 1	28%	(54/190)	72%	(136/190)
Period 2	17%	(24/140)	83%	(116/140)
All	24%	(78/330)	76%	(252/330)

Table 1: *Wh*-questions involving a subject dislocation in child French

Other indications that children differentiate between left- and right-dislocation in a target-like fashion can be found in their exclusive use of left-dislocation to express contrastive topics (as illustrated in (19)) or whenever there is a need for emphasis on the topic (i.e. whenever structural/prosodic saliency is required — as for instance in the repair strategy illustrated in (20)).

- (19) **maman**, e fait xx # **moi**, e fais (le) drapeau. (Anne 2.0.27)
mum e makes ?? me e make (the) flag
‘Mum makes the ?? and I make the flag.’
- (20) *ANN: i(l) est où, **mon cass(er)ole**? (2;4.2)
it is where my saucepan
*TAT: qu’ est ce que tu cherches?
‘What are you looking for?’
*ANN: **mon cass(er)ole**, (e)lle [= ? (i)l] est où?
my saucepan it is where

I conclude that children not only know that heavy topics must be dislocated, they are also aware from an early age of the subtle differences between left- and right-dislocation of topics.

4 Implications for the study of null subjects

It has been proposed in the acquisition literature that subject omissions in finite contexts were due to a lack of ability to license topics adequately (e.g. Bromberg & Wexler 1995; Wexler 1998; Schaeffer, Gordishevsky, Hadar, & Hachohen 2002). On the assumption that subjects can be dropped in adult English under certain pragmatic conditions, these researchers postulate that children lack the ability to evaluate such pragmatic conditions and that as a result, they omit more subjects than the target grammar would allow.

- (21) a. *Young children are less discriminating than adults as to the variety of pragmatic contexts in which they will allow null topics.* (Bromberg & Wexler 1995:244)
b. *The null-subjects of finite verbs represent a kind of pragmatic error.* (Wexler 1998:33)

- c. [...] young children lack the pragmatic rule that governs subject drop in English. [As a result they] produce non-adult-like null subjects. (Schaeffer et al. 2002:225)

What the relevant pragmatic conditions for subject drop are is however left rather in the vague, as no precise account of the workings of subject drop in adult English is provided. The closest one comes to a definition is that dropped subjects have to be *very strong topics* (Wexler 1998:35). I suppose *strength* here is to be understood in terms of saliency in the context.

While I do not dispute that missing subjects are (interpreted as) topics, I believe to have presented here clear indications that throughout the null subject stage, children have the ability to identify and encode topics adequately. It is therefore highly unlikely that they would omit certain subjects erroneously *because* they are (salient) topics.

5 Conclusion

The competence required to encode topics has been shown to be available to children from the earliest attested stages of language production. Dislocated elements expressing the topic appear at the onset of word combination in child French, and children's use of left- and right-dislocation not only resembles target usage but also complies by its requirements: the present corpora suggest that children only dislocate indefinites when a topic interpretation is allowed, and that they abide by the requirement that the heavy subject of ILPs be dislocated except when it is in focus, which requires them to be able to identify and encode topics in a target-like fashion.

I have left unanswered the question of the locus of the relevant competence: Is it a pragmatic module, responsible for all discourse-related phenomena and possibly UG-external? Is it a module dedicated to information structure in particular and integrated in UG? Further research will be necessary to shed light on this issue.

References

- Auger, J. (1994). *Pronominal Clitics in Québec Colloquial French: A Morphological Analysis*. PhD thesis. University of Pennsylvania.
- Bromberg, H. & Wexler, K. (1995). Null subjects in child Wh-questions. *MIT Working Papers in Linguistics* 26. 221–247.
- Cinque, G. (1993). A null theory of phrase and compound stress. *Linguistic Inquiry* 24. 239–298.
- De Cat, C. (2002). *French Dislocation*. PhD thesis. University of York.
- De Cat, C. & Plunkett, B. (2002). Qu' est ce qu' i(l) dit, celui-là ? Notes méthodologiques sur la transcription d'un corpus francophone. In Pusch, C. D. & Raible, W. (eds.), *Ro-*

- manistische Korpuslinguistik: *Korpora und gesprochene Sprache / Romance Corpus Linguistics: Corpora and Spoken Language*. CD-ROM edition. Tübingen: Narr.
- Erteschik-Shir, N. (1997). *The Dynamics of Focus Structure*. Cambridge: Cambridge University Press.
- Givón, T. (1976). Topic, pronoun and grammatical agreement. In Li, C. (ed.), *Subject and Topic*. New-York: Academic Press. 149–211.
- Gundel, J. (1975). Left dislocation and the role of topic-comment structures in linguistic theory. *OSU WPL* 18. 72–131.
- Jäger, G. (2001). Topic-comment structure and the contrast between Stage Level and Individual Level Predicates. *Journal of Semantics* 18. 83–126.
- Lambrecht, K. (1994). *Information Structure and Sentence Form. Topic, focus, and the mental representation of discourse referents*. Cambridge: CUP.
- Li, C. N. & Thompson, S. A. (1976). Subject and topic: a new typology of language. In Li, C. (ed.), *Subject and Topic*. New-York: Academic Press. 457–489.
- Milsark, G. (1974). *Existential Sentences in English*. PhD thesis. MIT.
- Reinhart, T. (1981). Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27(Special issue on Pragmatic Theory). 53–94.
- Reinhart, T. (1996). Interface economy: Focus and markedness. In Wilder, Gaertner, & Bierwisch (eds.), *The Role of Economy Principles in Linguistic Theory*. Berlin: Akademik Verlag.
- Rizzi, L. (1997). The fine structure of the left periphery. In Haegeman, L. (ed.), *Elements of grammar. A handbook of generative syntax*. Dordrecht: Kluwer. 281–337.
- Schaeffer, J., Gordishevsky, G., Hadar, G., & Hacoheh, A. (2002). Subjects in English-speaking children with Specific Language Impairment. In Costa, J. & Freitas, M. J. (eds.), *Proceedings of the GALA 2001 Conference*. Lisboa: Associação Portuguesa de Linguística. 223–231.
- Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the Optional Infinitive stage. *Lingua* 106. 23–79.