

## Discourse Anaphora

Francis Cornish

► **To cite this version:**

Francis Cornish. Discourse Anaphora. Keith Brown. Encyclopedia of Language and Linguistics (2e édition), Elsevier Ltd. (Oxford, RU), pp. 631-638, 2006. <hal-00963663>

**HAL Id: hal-00963663**

**<https://hal-univ-tlse2.archives-ouvertes.fr/hal-00963663>**

Submitted on 21 Mar 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Published in *Encyclopedia of Language and Linguistics* (2<sup>nd</sup> edition), ed. K. Brown, Oxford: Elsevier, 2006, pp. 631-638

## Text Analysis and Stylistics

### Discourse Anaphora

Francis Cornish

*CNRS ERSS (UMR 5610) and University of Toulouse-Le Mirail*

#### 1. Introduction

Discourse anaphora is a means of managing the memory representation of the discourse being constructed by the speech participants on the basis of a co-text as well as a relevant context (for further details of this view, see Cornish, 1999; 2003). But see also the articles “Anaphora, cataphora, exophora, logophoricity”, “Anaphora, Philosophical aspects”, and particularly “Deixis and Anaphora, Pragmatic approaches” in the present volume, for somewhat different perspectives. Where discourse is concerned, it is clear that not all referents will have been introduced via an explicit textual antecedent: it is also possible for them to be evoked “obliquely” in terms of an association or a (stereotypical) inference of some kind (see especially example (1) below).

This article takes what might be termed a “discourse-cognitive” view of anaphoric reference, rather than a textual-syntactic one. The use and interpretation of non-bound anaphors — that is, anaphoric expressions whose interpretation is not determined primarily by features of the clause in which they occur — require not only a relevant co-text as well as context, but also, crucially, a psychologically salient representation of the discourse evoked via what in previous work I call the **antecedent-trigger** (an utterance token, gesture or percept). See §3.1 below for a presentation of this term.

#### 2. Some useful concepts and distinctions in the study of indexical reference: “anaphora”, “deixis” and “textual/discourse deixis”

Let us start by drawing the more fundamental distinction between the dimensions of *text* and of *discourse* (see my other article in this edition of the *Encyclopedia* on “Understanding spoken discourse”, as well as the article “Macrostructure”). Very briefly, *text* refers to the ongoing physical, perceptible trace of the discourse partners’ communicative or expressive activity: this includes the verbal content of an utterance, but also prosody, pausing, semiotically-significant gestures (non-verbal signals: see the article “Nonverbal communication”), and of course punctuation, layout and other graphic devices in the written form of language. See also the article entitled “Text and Text Analysis”. The addressee or the reader exploits these textual features in order to infer the *discourse* being co-constructed by the participants. *Discourse* in this sense refers to the hierarchically-structured product of the constantly evolving sequence of utterance, illocutionary, propositional and indexical acts jointly performed by the discourse partners (see for

an illustration representation (7) below of the discourse corresponding to text (5)); this product is of course partly determined by the context invoked. See also the article on “Text World Theory” for a similar conception of discourse.

*Discourse anaphora*, then, constitutes a procedure (realized via the *text*) for the recall of some item of information previously placed in discourse memory and already bearing a minimal level of attention activation. It is essentially a procedure for the orientation of the interlocutor’s attention, which has as essential function the *maintenance* of the high level of activation which characterizes a discourse representation already assumed to be the subject of an attention focus by the interlocutor at the point of utterance. It is not only the anaphoric expression which is used (typically, a third person pronoun) which realizes (discourse) anaphora, but also the clause in which it occurs as a whole. This predicational context acts as a kind of “pointer”, orienting the addressee towards the part of the discourse representation already cognitively activated, and which will make it possible to extend in terms of an appropriate coherence relation (see Kleiber, 1994: Ch. 3, as well as the articles “Cohesion and coherence: linguistic approaches” and “Coherence: a psycholinguistic approach”).

- 1) [*Fragment of dialogue in film:*] *Woman*: “Why didn’t you write to me?” *Man*: “I did..., started to, but I always tore ’em up”. (Extract from the film *Summer Holiday*. Example (5.5), p. 157, in Cornish, 1999, reproduced as (6d) in Cornish, 2005, p. 204)

In (1), an instance of “indirect anaphora”, it is the illocutionary point of the woman’s initial question, which bears on the non-existence of a letter or letters which she had expected the man to write to her, together with the lexical-semantic structure of the verbal predicate *write* (in the sense “engage in correspondence”), which provide an interpretation for the unstressed pronoun *’em* in the third conjunct of the man’s reply. The example clearly shows the extent to which inferences based on existing discourse representations, lexical and general knowledge are mobilized in the operation of discourse anaphora, which clearly does not **require** the co-presence of an explicit textual antecedent, under the traditional co-textual account of anaphora, in order to exist (see also Blackwell, 2003, in connection with a study of Spanish conversations and spoken narratives, and Ziv, 1996).

Here is an example involving different possible continuations of the antecedent-trigger predication in terms of distinct anaphoric predications:

- (2) Jason<sub>i</sub> witnessed a terrible accident<sub>j</sub> yesterday<sub>k</sub> at the Dunton crossroads<sub>l</sub>. He<sub>i</sub> was very shaken/ It<sub>j</sub> resulted in two deaths/ #It<sub>k</sub> was dull and overcast/ ?#It<sub>l</sub>/The place<sub>l</sub> is a known danger-spot.

(N.B. Subscripted letters indicate identity or otherwise of the intended referents of the expressions so-marked). In (2), the first two argument referents introduced (‘Jason’ and ‘the terrible accident Jason witnessed the day before the utterance of (2)’) may be naturally continued via unaccented pronouns – but not the scenic referent ‘the day before utterance time’, nor (or at least, not as easily as with the first two entity referents evoked) ‘the Dunton crossroads’, which is expressed by an adjunct, and which serves as a locative frame of reference for the situation evoked as a whole (see also the point made in regard to certain natural Spanish conversational data by Blackwell, 2003: 118, 122-3). The slashes here are meant to indicate alternative continuations of the initial sentence. The crosshatch preceding an example is intended to signal

that, as a potential utterance, it is unnatural in the context at hand. Example (2) is intended to be discourse-initial, and not part of an earlier, ongoing discourse.

*Deixis*, on the other hand (see also “Deixis and anaphora, Pragmatic approaches” and “Deixis: stylistic approaches”), is a procedure which relies on the utterance context to re-direct the interlocutor’s attention towards something associated with this context (hence which is potentially familiar to him or her), but to which s/he is assumed not already to be attending. As Kleiber and other pragma-semanticists have observed, *deixis* causes a break in the continuity of the discourse at the point where the deictic procedure is used, so that in effect the interlocutor is invited to “step out” of this discourse context to grasp a new referent in terms of the current situation of utterance – or alternatively another aspect of a same referent, which has already been focused upon. So *deixis* serves to introduce a new referent into the discourse, on the basis of certain features of the context of utterance.

Now, *textual* as well as *discourse deixis* provide a transition between the notions of *deixis* and *anaphora*, since they consist in using the deictic procedure to point to part of a pre- or post-existing textual or memory representation, but which is not necessarily highly activated. The interlocutor will therefore need to exert a certain cognitive effort in order to retrieve it. This interpretative effort will involve **constructing** an ‘entity’, on the basis of the discourse representation in question, in order for it to be the subject of a predication, an anchor for the introduction of new information. Where there is a difference in topic-worthiness between the representation introduced by a trigger and the intended referent, the *discourse-deictic* and not *anaphoric* procedure must be used, as in (3), an attested utterance:

- (3) [End of the words of welcome uttered by the director of the Language Centre, at the start of a conference, University of Edinburgh, 19 September 1991]“...We intend to record the guest speakers, so *these* will be available to participants at the end of the Conference...” (Example (20) in Cornish, 2005, p. 212)

In order to access the referent targeted via the proximal demonstrative pronoun *these* (namely, ‘the recordings of the guest speakers’ papers’), the hearer will have to draw an inference of the type: “If the guest speakers’ papers are recorded at time  $t_0$ , then at time  $t_n$  ( $t_n > t_0$ ), there will be recordings of these papers”. Unlike the indirect referent in (1) and the first two more ‘direct’ ones in (2), here the implicit referent has not attained the status of a potential *topic* (see the articles “Thetic-categorical distinction” and “Topic and comment”) by the time the initial clause is processed; for it is ‘the guest speakers’ which enjoys this status at this point. So it is predictable that the elaborative *so*-clause which immediately follows will continue to be about these entities. The demonstrative pronoun *these* in (3) directs the hearer’s attention towards a referent which s/he must create on the basis of the representation introduced via the initial conjunct, as well as in terms of his or her knowledge of the world. So it is an instance of *discourse deixis* rather than of *anaphora*. Indeed, the (anaphoric) personal pronoun *they* in its place would have maintained the situation evoked via the initial conjunct, resulting in the retrieval of the only salient topic-worthy entity within it, ‘the guest speakers’ – an interpretation leading to quite severe incoherence here.

### 3. Three essential ingredients of the operation of discourse anaphora: “antecedent-trigger”, “antecedent” and “anaphor”

#### 3.1 The antecedent-trigger

This is not necessarily an explicit, textual expression (a phrase of some kind). It may also be a percept or a non-verbal signal (see Cornish, 1996 and 1999: ch.4). In (1) it is the illocutionary point of the woman's initial question, in conjunction with the use of the verb *write* which triggers the discourse representation in terms of which the pronoun 'em refers, while in (2) it is the use of the descriptive noun phrases *Jason, a terrible accident* and *the Dunton crossroads*. The broader notion of "antecedent-trigger", in relation to the traditional, canonical textual "antecedent" of normative written prose, which is required to be morpho-syntactically and semantically parallel to the anaphor, is useful in that it enables us to include both "exophora" and "indirect anaphora" (see example (1)) within the purview of anaphora *per se* — of which both these phenomena are instances. See also Cornish (1999: 41-3).

### 3.2 The antecedent

This is a psychologically salient discourse representation in terms of which the anaphor refers or denotes. As this characterization suggests, it is a unit of *discourse*, not of *text* (see the distinction drawn at the beginning of section 2 above), and may be constructed via direct interpretation of the co-text in terms of a relevant context, or in terms of the context alone in conjunction with relevant aspects of mutual knowledge, or in terms of inferences from either of these. See as an example the informal description of the antecedent of the unstressed pronoun 'em in (1): 'the set of unfinished, torn and unsent letters which the man had begun writing to the woman'. See also Dahl & Hellman (1995), Langacker (1996), Van Hoek (1997) and Cornish (1999: 44-7). A given antecedent-trigger may give rise to several distinct 'antecedents' (in this sense), as a function of the possible drawing of inferences, of what is predicated of the former's referent, or of the functioning of the type of anaphor chosen to target it. (4a,b) provide examples, where the antecedent-trigger *one of the new Toyota models* in the first sentence of (4a) gives rise to different 'antecedents' targeted by the pronouns *they*, the *it* in the second anaphoric continuation, and *one*; while the entire initial sentence acts as antecedent-trigger for the antecedents created via the anaphors *that* and the *it* in the final anaphoric continuation in this example:

- (4) a John bought one of the new Toyota models yesterday. *They* are really snazzy cars./ *It* is standing outside his front door./ Mary bought *one* too./ At least, *that's* what he told me./ *It* took only half an hour to complete.
- b "The grouse season begins today, and *they're* being shot in large numbers"  
(*Today Programme*, BBC Radio 4, 10.12.04)

### 3.3 The anaphor

This is a referentially dependent indexical expression. The relation is not exclusively between antecedent-trigger and anaphor (except in the case of metalinguistic occurrences, as in this example: A: *Psephism was much in vogue in those times*. B: *What does that mean?*: but these in any case, as Lyons, 1977 points out, are instances of *textual deixis*). First, then, the anaphor refers, not to its antecedent(-trigger), but in terms of whatever its antecedent(-trigger) refers to (see Lyons, 1977, vol. 2: 660). Second, the discourse referent evoked via the antecedent-trigger is not necessarily the same at the point of retrieval via the anaphor as it was at the point of introduction: minimally, what will have been predicated of the referent concerned within the antecedent-trigger predication (and potentially within subsequent predications) will have altered

that referent's representation – perhaps even radically. Third, it is not simply the anaphor on its own which retrieves the (updated) discourse referent at the point where it occurs in the co-text, but the anaphoric (or “host”) predication as a whole: compare the anaphoric continuations in examples (2) and (4a) above in particular in this respect. So what is predicated of the referent of the anaphor acts as a filter, ruling out theoretically possible referents or denotata, and as a pointer, targeting and selecting a salient discourse representation which is compatible with what is predicated of the anaphor's referent (see also Yule, 1981 and Dahl & Hellman, 1995).

As we shall see in analyzing text (5) in the next section, there is a variety of types of anaphor — zero forms, ordinary pronouns (see the article “Pronouns”), demonstrative pronouns (see the article “Demonstratives”), reduced proper names, demonstrative, definite and possessive full NPs, ellipses of various kinds (see the article “Ellipsis”), and so on — which each have distinct indexical properties; as such, they each function to establish different kinds of discourse anaphoric structures, and are each sensitive to specific types of discourse context and function. See Cornish (1999: 51-68) for some discussion. On the use of demonstratives in narrative discourse, see in particular Himmelmann's (1996) typological study.

#### 4. The text- as well as discourse-sensitivity of discourse anaphora

The text we are going to analyze for illustration is taken from a British newspaper, *The Guardian* (1 July 1998, p. 3), reproduced under (5) (for convenience in the analysis which follows, the paragraphs are each numbered in the left-hand margin).

##### (5) *Monet waterlilies set £20m record*

###### *Luke Harding*

1. *A painting of the most famous garden in the history of art last night sold for £19,801,500, shattering all records for a work by Claude Monet.*
2. *Two frenzied telephone bidders pushed the price for Monet's Waterlily Pond and Path by Water to almost £20 million at Sotheby's, suggesting that good times are back again for the fickle art market.*
3. *The price, reached after six minutes of bidding, comfortably shatters the previous £13 million record for a painting by the artist. Waterlily Pond is now the most expensive Impressionist work sold by a European auction house since 1990. Sotheby's had estimated the sale price more modestly at £4-£6 million.*
4. *The oil painting, executed in 1900, was acquired by a private British collector in 1954 and has not been shown in public since then.*
5. *The identity of the buyer is a mystery. “We are still totting up the figures for the total auction,” a jubilant Sotheby's spokeswoman said last night. “It's been a very very good night.”*
6. *Monet was passionate about flowers and intrigued by landscape architecture. In 1893 he purchased a plot of land which adjoined the rural house in Giverny, near Paris, where he had moved 10 years earlier. A small stream ran through the plot, and Monet turned the garden into a horticultural paradise.*
7. *Monet worked tirelessly during the summer months, producing 12 pictures in 1899 and six in 1900. The oil sold last night shows the left section of his water garden, with the Japanese-style footbridge and path gently curving through patches of purple irises and tall grass.*
8. *“It took me some time to understand my waterlilies,” Monet said in a conversation with the author Marc Elder in 1924. “All of a sudden I had the revelation of how enchanting my pond was. Since then I have had hardly any other subject.” Waterlily Pond and Path by the Water is now the 11th most expensive ever painting sold at auction. Its sale price is easily eclipsed, though, by another work completed just nine years earlier — Portrait du Dr. Gachet — by a then little-known artist, Vincent Van Gogh, which went for \$82,500,000 (£55 million) in 1900.*
9. *Last night's sale follows a gradual recovery in the art market — unlike the overheated boom of the late 1980s, where it was focused in just one or two areas. Recent sales of Impressionist and Old Master*

*works have been encouraging — despite allegations that many of Van Gogh's best-known works are fakes.*  
(Example (8) in Cornish, 1998: 30-31)

In this text, there are several “topic chains” (see Cornish, 1998; 2003 for further details). A topic chain is a sequence of mainly anaphoric (referentially dependent) expressions within a text which retrieve the same referent, which is thus the subject of several predications for a segment of the text. This referent may have been introduced explicitly via a *referentially autonomous* expression, such as a full proper name, an indefinite NP or a full definite NP. This is the “head” of the chain, the anaphoric expressions retrieving its referent then being the “links”. We will adopt Dik’s (1997: 218) definition of topic chains (what he calls “anaphorical chains”) in recognizing three theoretical discourse-functional positions within them: 1) the head of the chain, which introduces the topic referent into the discourse; 2) a second-link position (only exploited in “macro”-topic chains), whose function is to “reconfirm” the installation of the topic referent in question — i.e. it has an essentially addressee-oriented function —; and 3) a third “position” which may be multiply filled, consisting of purely anaphoric retrievals of the topic referent whose function is to maintain the high attention focus now accorded (or assumed to be so accorded) to that referent by the addressor. By the third link, then, the referent retrieved is taken as enjoying full topic status in the discourse.

The four most important topic chains in text (5) are the following: 1) the one dealing with the article’s overall topic, the painting by Monet which had just been sold by auction for a record price; 2) the one bearing on the price fetched by the sale; 3) the one having to do with the artist himself; and finally 4) the one dealing with the plot of land which he had bought at Giverny in 1893, of which the stream which flowed through it served as a model for his painting. These chains are made up of the following successions of expressions:

1. *A painting of the most famous garden in the history of art...Monet's Waterlily Pond and Path by Water...Waterlily Pond...The oil painting... $\phi$ ...The oil sold last night...Waterlily Pond and Path by the (sic) Water...Its...;*
2. *£19,801,500...the price for Monet's Waterlily Pond and Path by Water...The price... the sale price...Its sale price...;*
3. *Claude Monet...Monet...the artist...Monet... $\phi$ ...he...he...Monet...Monet... $\phi$ ...his...Monet;*
4. *A plot of land which adjoined the rural house in Giverny, near Paris, where he had moved 10 years earlier...the plot...the garden...his water garden...*

Let us represent these four topic chains schematically, using the abbreviations ‘R-A’ for “referentially autonomous expression” and ‘R-NA’ for “referentially non-autonomous expression”, as follows (‘H’ = “Head of chain”, ‘L2’ = “Link-2”, and ‘L3’ = “Link-3”):

(6) *Schematic representation of the 4 topic chains in (5)*

**Topic Chain 1:** H: R-A; L2: R-A; L3: R-NA, R-NA, R-NA, R-NA, R-A, R-NA.  
(‘the painting by Monet’)

**Topic Chain 2:** H: R-A; L2: R-A; L3: R-NA, R-NA, R-NA.  
(‘the sale price reached by the painting’)

**Topic Chain 3:** H: R-A; L2:  $\emptyset$ ; L3: R-NA, R-NA, R-NA, R-NA, R-NA, R-NA,  
(‘Claude Monet’) R-NA, R-NA, R-NA, R-NA, R-NA.

**Topic Chain 4:** H: R-A; L2: Ø; L3: R-NA, R-NA, R-NA.  
 ('Monet's garden')  
 (Item (9) in Cornish, 1998: 32, slightly adapted)

This representation points up the fact that referentially autonomous and anaphoric expressions do not occur indiscriminately in any position within a chain. For apart from the autonomous expression which occurs in fifth position within the link L3 in chain 1 (*Waterlily Pond and Path by the (sic) Water*), autonomous referring expressions always occur in the central positions within chains (positions H and L2); while anaphoric expressions appear only within link-position L3. See Ariel (1996) on the distinction between referentially autonomous and non-autonomous indexical expressions (particularly as far as the distinction between full and reduced proper names is concerned).

Interestingly, it is precisely in the two topic chains which are intuitively the most central to text (5) as a whole (namely, chains 1 and 2), that we find link L2 realized by an autonomous expression. The other two chains (3 and 4), where this same link is by hypothesis unfilled, evoke referents which are subsidiary within this discourse in relation to the referents developed by chains 1 and 2: the article deals, after all, with the particular work by Monet as well as with the record price it fetched in auction, rather than with the artist or his garden as such. Furthermore, the representation in (7) (below) of the discourse structure associated with (5) shows that, while chains 1 and 2 are set up within central discourse segments (paragraphs 1-3, 5, 8b and 9), chains 3 and 4 are restricted to background, subsidiary segments (the segments corresponding to paragraphs 4 and 6-8a). So it is not surprising that the last two topics should not have required an L2 link for their installation within the discourse.

Let us look now at the relationship between the occurrence of an expression realizing a given link in a chain and the discourse function of the unit in which it occurs, in terms of the structure of the discourse as a whole. Schema (7) represents the structure of text (5) as discourse (indentations indicate subsidiary segments):

(7) *Discourse structure corresponding to text (5)*

1. [Para 1: **Introduction of the global discourse topic**, the painting by Monet and its record price reached at an auction in London]
  2. [Para 2: **Continuation** of the sequence of events surrounding 1]
  3. [Para 3: **Development on the record price** reached by the sale of the work]
    4. [Para 4: **Background segment** on the history of the painting, from its inception to the present]
  5. [Para 5: **Return pop to the central topic**. Introduction of two local topics (not developed in the remainder of the text): the buyer's identity, and the calculation of the total price of the sale]
    6. [Para 6: **Flashback** to the subject matter of the painting and its origin: the purchase by Monet of a plot of land near his country house in Giverny - the inspiration behind the work. No reference to the painting as such]
    7. [Para 7: **Development of this background topic**: what the painting shows of the garden]
    - 8a. [1<sup>st</sup> half of Para 8: **Continuation of the topic of Monet's inspiration drawn from his garden at Giverny**]
  - 8b. [2<sup>nd</sup> half of Para 8: **Return to the central topic of the record price of the painting**, and comparison with the astronomical price reached by another painting of the same period]
  9. [Para 9. **Conclusion: Extrapolation to the art market in general** - the recovery of the art sale market precipitated by this auction. No reference to Monet's painting]
- (Item (10) in Cornish, 1998: 34)

The structure of the first two of the four topic chains in (5) in this respect is as follows. The title of the article already sets up the global theme, the sale of Monet's painting *Waterlily Pond and Path by Water* and the record price fetched thereby. In the introductory paragraph, this dual aspect of the global theme is made explicit in a complex sentence. As an introductory sentence-paragraph, it has a "thetic" character, where the information it presents to the reader is entirely new (see the article "Thetic-Categorical Distinction").

The second paragraph is an elaboration of the situation established by the first, dealing more specifically with the price of this sale at auction; but it also serves to *identify* the painting by Monet which is the topic of the first paragraph. This cannot be an instance of *cataphora*, where the antecedent-trigger follows the "cataphor", for all that, given the referentially autonomous character of *both* nominal expressions used here (full indefinite NP and full proper name); so the referential dependency of anaphor on "antecedent-trigger" does not obtain — that is, we have *coreference* without *anaphora* in the strict sense here. The use of two indexically strong (referentially autonomous) expressions at this point — *Monet's Waterlily Pond and Path by Water* and *the price for Monet's Waterlily Pond and Path by Water* — is no doubt motivated by the concern to promote their referents to global topics within the article, following their brief introduction in the initial paragraph.

The third paragraph continues this theme of the price fetched by the sale of the painting. Note that the two references to the dual global topics of the article are made via lexically explicit NPs (a reduced proper name for the painting (*Waterlily Pond*) and a definite, also reduced NP for its price (*the price*)), and not via pronouns. (The fact that these are reduced NPs means that they are not referentially autonomous, but are potentially anaphoric, like pronouns). See Geluykens (1994) on the question of anaphoric "repairs" in spoken interactions, where the speaker mis-assesses their addressee's current attention state, and uses a reduced indexical form type (a pronoun of some kind), which s/he immediately corrects to a fuller form (a definite or demonstrative NP or a proper name). There are two reasons behind the use of *the price* as subject of the initial sentence of this third segment of the discourse: first, this reference is followed immediately by a non-restrictive relative clause in apposition, a position from which unaccented pronouns (here *it*) are excluded (this is also the case with the lexical NP *the oil painting* in paragraph 4); and second, the repetition of the definite article and of the lexical head of the complex NP which were used in the previous paragraph to "topicalize" the referent at issue, signals at the start of this new segment that it will continue to be about it. In other words, the referent in question, though topical, is nonetheless re-evoked at the very beginning of a *new* discourse segment, and no longer within the one in which it was originally topicalized: see Fox (1987) in this respect, who argues that repeated proper nouns in English spoken and written texts may have this function, and also Blackwell (2003) in relation to her spoken Spanish data.

Similarly, the use of a proper noun, albeit reduced (*Waterlily Pond*), at the point in this segment where this reference occurs, is made necessary by the evident need to distinguish this referent from the other central referent which has already been evoked in this paragraph ('the painting's sale price'), but which enjoys an advantage over it in terms of topic-worthiness at the point where the reference is made. The pronoun *it* used in its place would certainly have retrieved this latter referent, and not 'the painting' as such. This fits in well with what is stated by Levinson's (1995) "M-principle" (see also Huang, 2000: 208), to the effect that the use by a speaker of a phonologically and lexically more substantial expression where a more attenuated one could have been used in its place, is normally intended and interpreted as **not** meaning the same as if the more unmarked expression had been used (see the article "Deixis and Anaphora,

Pragmatic approaches” for further details of this account).

As for the two references to these two macro-topics throughout paragraph 4, where the focus switches to background considerations relating to the central theme, the first is made via a definite lexical NP (*The oil painting*, in initial subject position of the segment) and the second by means of a null form, the ellipsed subject of the second conjunct of the clause which realizes this segment. The motivation behind the use of the former expression type at the beginning of this segment is exactly the same as that of its counterpart *the price* in the same position at the start of the previous paragraph. Since it is followed by a non-restrictive relative, a pronoun could not have occurred in its place; but even if one could, it would be excluded for reasons of anaphoric ambiguity: for the pronoun *it* here — leaving aside what is predicated of the referent of this expression in this context — would have retrieved the referent ‘the sale price’ evoked by the immediately preceding clause.

Paragraph 7, which falls together with paragraphs 6 and the first half of 8 in a background discourse segment, includes a reference to the painting, a reference which re-evokes at the same time the event of the sale on 30<sup>th</sup> June (*the oil sold last night*). Once again, we have to do here with a definite (elliptical) NP; and just as in the previous cases, the reason for it is the existence of comparable referents which are in competition in terms of topicality. The first half of paragraph 8 (8a) continues the theme of the subject of the painting (‘Monet’s garden’), a theme which is abruptly interrupted in the middle of this paragraph by the opening of a segment returning to the macro-topic of the record sale price reached by this painting at auction at Sotheby’s.

Now, it is precisely by means of a full proper name that this transition to a segment dealing with the circumstances of the sale of the painting is carried out. As such, we can hypothesize that it corresponds to what Dik calls a *resumed topic*. It is this particular marked discourse function realized via this referentially autonomous expression which motivates its exceptional filling of the “anaphoric” L3 link in this macro-topical chain (recall the structure of Chain 1 given in schema (6) above). Once its topic status has been re-established within this new segment — which does not correspond this time to the start of a new paragraph in terms of textualization: see the *text/discourse* distinction drawn in §2 above —, the next (anaphoric) reference may be realized by means of an unaccented pronoun: in this case, by the pronoun contained in the possessive determiner *Its*.

## 5. Conclusion

As we have seen in connection with text (5) in particular, the occurrence of different types of non-autonomous, potentially anaphoric expressions in a text is in large part determined by the discourse function of the unit of discourse corresponding to the textual segment in which the expression appears, as well as by its position within that segment. In the case of written newspaper articles of the kind seen in (5) at least, it is clear that initial position within a unit is reserved for lexically-based NPs (reduced proper nouns and definite NPs, as well as demonstrative ones), whatever the degree of topicality and accessibility their intended referent may enjoy at that point; unaccented pronouns and of course null pronouns are virtually excluded from such positions, since they serve to mark the continuity of the attention focus established prior to their occurrence. We have also seen how the co-presence of competing referents in the immediately prior co-text may favor the use of an indexically stronger form type than a pronoun or a null anaphor, in order to avoid unintended anaphoric continuities.

Discourse anaphors, in sum, are sensitive to the hierarchical structure of the discourse

which may be assigned to a given text, in conjunction with an appropriate context, and their choice by a speaker or writer is clearly a function of his or her ongoing assessment of the conditions under which their addressee or reader will be operating at the point of use. In the case of discourse anaphors, it is clear, in Dahl & Hellman's (1995: 84) colorful words, that their 'antecedents' "aren't just sitting there, waiting to be referred to, but rather ha[ve] to be created by some kind of operation": this is a reflex of the fact that anaphora operates within the dynamic, ongoing construction of *discourse*, rather than exclusively in terms of the more static dimension of *text*.

## References

- Ariel, M. (1996). 'Referring expressions and the +/- coreference distinction'. In Fretheim, T. & Gundel, J.K. (eds.), *Reference and referent accessibility*. Amsterdam & Philadelphia: John Benjamins, 13-25.
- Blackwell, S.E. (2003). *Implicatures in discourse. The case of Spanish NP anaphora*. Amsterdam & Philadelphia: John Benjamins.
- Cornish, F. (1996). "'Antecedentless" anaphors: deixis, anaphora, or what? Some evidence from English and French'. *Journal of Linguistics* 32, 19-41.
- Cornish, F. (1998). 'Les "chaînes topicales": leur rôle dans la gestion et la structuration du discours'. *Cahiers de Grammaire* 23, 19-40.
- Cornish, F. (1999). *Anaphora, discourse and understanding. Evidence from English and French*. Oxford: Clarendon Press.
- Cornish, F. (2003). 'The roles of (written) text and anaphor-type distribution in the construction of discourse'. *Text* 23, 1-26.
- Cornish, F. (2005). 'Degrees of indirectness: two types of implicit referents and their retrieval via unaccented pronouns'. In Branco, A., McEnery, T. & Mitkov, R. (eds.) *Anaphora processing: linguistic, cognitive and computational modelling*. Amsterdam & Philadelphia: John Benjamins, 199-220.
- Dahl, Ö. & Hellman, C. (1995). 'What happens when we use an anaphor?' In Moen, I., Simonsen, H. G. & Lødrup, H. (eds.), *Papers from the XVth Scandinavian Conference on Linguistics*, Oslo, January 13-15, 1995. Department of Linguistics, University of Oslo, 79-86.
- Dik, S.C. (1997). 'Anaphora'. Ch. 10 in *The Theory of Functional Grammar Part 2: Complex and derived constructions*. Mouton de Gruyter: Berlin & New York, 215-228.
- Fox, B.A. (1987). *Discourse structure and anaphora*. Cambridge: Cambridge University Press.
- Fox, B.A. (ed.). (1996). *Studies in anaphora*. Amsterdam/Philadelphia: John Benjamins.
- Geluykens, R. (1994). *The pragmatics of discourse anaphora in English: evidence from conversational repair*. Berlin: Mouton de Gruyter.
- Himmelmann, N.P. (1996). 'Demonstratives in narrative discourse: a taxonomy of universal uses'. In Fox (ed.), 206-254.
- Huang, Y. (2000). *Anaphora. A cross-linguistic study*. Oxford: Oxford University Press.
- Kleiber, G. (1994). *Anaphores et pronoms*. Louvain-la-Neuve: Duculot.
- Langacker, R. W. (1996). 'Conceptual grouping and pronominal anaphora'. In Fox (ed.), 333-378.
- Levinson, S. C. (1995). 'Three levels of meaning'. In Palmer, F. (ed.), *Grammar and meaning. Essays in honour of Sir John Lyons*. Cambridge: Cambridge University Press, 90-115.
- Lyons, J. (1977). Ch. 15: 'Deixis, space and time'. In *Semantics*, Vol.2, Cambridge: Cambridge University Press, 636-724.
- Van Hoek, K. (1997). *Anaphora and Conceptual Structure*. Chicago: University of Chicago

Press.

Yule, G. (1981). 'New, current and displaced entity reference'. *Lingua* 55, 41-52.

Ziv, Y. (1996). 'Pronominal reference to inferred antecedents'. *Belgian Journal of Linguistics* 10, 55-67.