



Cercles 29 (2013)

PHRASING / REPHRASING

THE CHOICE OF LEXICAL INFORMATION IN ANAPHORIC DEFINITE DESCRIPTIONS

LAURE GARDELLE

École Normale Supérieure de Lyon

The aim of the present paper is to contribute to the study of the principles underlying lexical information selection in noun phrases (NPs) referring to an entity that has already been introduced via a textual element. The role of such NPs is not primarily to convey new descriptive information about the referent, as a (provisional) representation of it has already been given ahead of the anaphor.

The study looks more specifically into NPs that share a same thematic determiner,¹ *the*, and focuses on cases of co-referential textual anaphora. This restriction allows for a comparison between the lexical content of the anaphor and that of the antecedent. That antecedent, which is typically an NP but can also be clausal, might introduce the referent in discourse or might itself constitute a case of rephrasing.

The study proposes to test several hypotheses. The first one is whether the lexical content in the anaphor is determined by that of its textual antecedent: does a definite description in prototypical anaphora always take up the lexical information of the antecedent (whether all of it or just some of it), or might there be addition of information? A related question is whether a speaker always selects the minimal information required for referent retrieval. In order to address these questions, the first section establishes the types of rephrasing found in the anaphoric definite descriptions of the corpus. The study goes on to consider whether definite descriptions in coreferential textual anaphora show a preferred level of categorisation, using Rosch *et al.* (1976)'s concept of basic level. This criterion is found to be insufficient, and leads to examine the data in the light of Sperber & Wilson

¹ A thematic determiner is one that implies that the speaker considers the existence of the referent to be already established.

(1986)'s relevance principle. The conclusions drawn from the data finally lead to a few considerations on the influence of the textual antecedent on the anaphor.

Typology of rephrasing in anaphoric definite descriptions

The typology proposed here is based on a corpus of over 250 occurrences, which were taken across genres.² For spoken discourse, it was more difficult to find occurrences of rephrasing through definite descriptions, partly because speakers typically devote little time to a given referent (so that there is not always rephrasing), and also because they were found to use more personal pronouns than definite descriptions when they do rephrase. For those practical reasons, the 'spoken' occurrences are extracted from lectures (spoken corpus BASE³), interviews and forums (which are written documents, but with a more conversation-like style). The proportions mentioned in the article are not necessarily meant to be representative of overall language use as the corpus is too small for significant estimates, but they are given nonetheless to put the various kinds of rephrasing in perspective within the sample used.

The corpus shows great variety in the types of rephrasing, with no less than ten different processes. The most common is by far reduction: the information given in the antecedent is only partially taken up, as in *the rotting boat* rephrased as *the boat* [HOLDSTOCK : 23⁴]. This process is used in 94 occurrences of the corpus, that is, nearly 38%. The second most common type is identical repetition of a lexical content that is itself minimal: the lexical information is the same as in the antecedent NP, and the latter only contains a determiner, a basic level noun⁵ (or less commonly the pronoun *one*) and possibly a modifier or complement that is essential to referent retrieval. One example of this is *a tail* taken up as *the tail* [PACKARD : 53] or, about vacuum cleaners: *I use a standup one instead. The standup one is a Panasonic* [BRITISH EXPATS : 1]. This type of rephrasing is used in over 35% of occurrences in the corpus. All in all, therefore, the first two processes represent nearly two thirds of cases in the corpus and, more specifically, over 99% of spoken (or close to spoken) utterances. If one adds cases of rephrasing through a hyperonym⁶ (5.7%) and cases of synthesis via nominalisation (that is, cases of resumptive anaphora, in which the definite

² See the appendix at the end of the present article for references.

³ http://www.rdg.ac.uk/AcaDepts/ll/base_corpus (accessed 5 March 2013)

⁴ For full references, see the list of documents given in the appendix.

⁵ For a definition of this word, see the following section.

⁶ Such as *the most enormous hound* taken up as *the dog* [HOLDSTOCK : 34].

description rephrases a clausal antecedent,⁷ 2%), it can be concluded that over 80% of definite descriptions in prototypical anaphora do not add information. This is in keeping with the principle of co-referential anaphora: as a description of the referent has already been given ahead, lighter information is sufficient in the anaphor.⁸

This figure, however, also means that nearly 20% of anaphora in the study add information compared to the textual antecedent. In the corpus, the textual antecedent is then always an NP. This is achieved through six different modes:

- adding judgment, for instance *the Chief Inspector* rephrased as *the old sod* [DEXTER: 31]. The corpus only shows this occurrence of the process.
- rephrasing via a hyponym: again, there is a single occurrence of this, in a novel: *the dog* is rephrased as *the hound* [HOLDSTOCK : 35].
- re-expanding, i.e. giving information that was already available in a previous mention of the referent. For instance, in a lecture on the Parthenon, *the Parthenon sculptures* is rephrased as *the sculptures*, and later on as *the sculptures of the Parthenon* [BASE]. This mode only concerns 4 occurrences in the corpus.
- adding information that is part of the (assumed) shared knowledge of the speaker and addressee. In the corpus, this case was found only in essays, press articles and lectures, and makes up 7% of occurrences. The anaphor might then be a conventional periphrasis, as found in the press (1% of occurrences). For instance, in an article published during the American presidential election campaign of 2008, the antecedent *Hillary Clinton* is taken up as *the former first lady* [SCHOR]. The corpus also shows other alternations between proper and common names, such as *the temple* with *the Parthenon* as its antecedent [BASE], or conversely, *the country* taken up as *England* [PAXMAN : 6].
- adding information that was given in or inferred from the context, but outside the antecedent (6% of occurrences). For instance one reads:

⁷ As in the following: 'Was the "market meltdown" really caused by one medium-sized French bank's need to get a big risk off its books in a hurry? The idea was being debated in the City yesterday.' [CHRISAFIS *et al.*].

⁸ Cases of synthesis through nominalisation add categorisation; for instance, *the idea* categorises the whole preceding proposition as an idea. But this is not really addition of information.

(1) The horse began charging down the street with the milk-wagon careening wildly behind it. Too late, the milkman ran after the horse shouting. But the horse would not halt, not even for policemen who waved their night sticks at it. For half a mile the runaway horse raced.
[PACKARD : 84]

- in no less than 9 cases (that is, 3.6% of occurrences), adding information that is totally new to the addressee. For example, one reads:

(2) The church was in a prosperous white suburb, where the roads were lined with BMWs [...] The choir was made up of cleaning-ladies from the building where John had his office. They were poor, shoeless people (some literally), but when they sang 'Nkosi Sikeleli Afrika', the black anthem, the cavernous mock-gothic church rang with sweet passion. [PAXMAN : 10]

This type of rephrasing was found only in written discourse, more specifically in essays and novels.

This brief overview shows two major facts. On the one hand, even in anaphoric definite descriptions, information might be added to that contained in the antecedent. Two questions that will therefore have to be explored are under what conditions such additions can be made and why speakers choose to add information in anaphoric NPs. On the other hand, two modes of rephrasing stand out as much more common than the others – namely, reduction and identical repetition of a minimal lexical content. One question raised is how reduction is achieved (is the same noun used, for instance?). In particular, does the principle of a basic level of categorisation, as defined by Rosch *et al.* [1976], play a part, and if it does, to what extent? The following section addresses this question.

Assessment of the influence of the basic level of categorisation

Within the vertical dimension of categorisation, what differentiates a hyponym from its hyperonyms is its level of inclusiveness: the higher up the vertical axis a category is, the more inclusive it is [EVANS & GREEN : 256]. For example, *furniture* can apply to more elements than its hyponym *chair*, so the category FURNITURE is more inclusive than CHAIR. Based on this inclusiveness principle, the basic level is defined by Rosch *et al.* [1976 : 390 *ff.*] as the level of inclusiveness that provides optimum cognitive economy.

This level can be identified thanks to various tests. If one considers the categories CHAIR and FURNITURE, for instance, the former is regarded as a basic level category because it is the most inclusive category for which it is

possible to form a schematic image – in other words, for which the shapes of the objects are most similar. Another test is that it is the most inclusive level at which members of the category share the same function – that is, they are used in similar ways. Chairs of any kind are typically used for sitting, and can be moved; conversely, FURNITURE is not a basic level category, but a superordinate one, because its members – a chair and a chest of drawers, for instance – are not used in similar ways and do not allow a common schematic image. As for the hyponyms of CHAIR, they correspond to subordinate (rather than basic level) categories because although they allow a common schematic image and share the same function, they are not the most inclusive categories that meet these conditions (for other tests, see for instance ROSCH *et al.* : 390 *ff* and EVANS & GREEN : 259-260).

Various experiments have shown that the basic level was the preferred level of categorisation of informants. For example, when a foreigner asks ‘What is this?’ about a chair, the spontaneous reply will be ‘It’s a chair’ rather than ‘It is a piece of furniture / an artefact / a kitchen chair’ [TAYLOR : 49]. Similarly, in one experiment, a spaniel whose picture was shown to informants was more often categorised as DOG than as SPANIEL or ANIMAL [CROFT & CRUSE : 84]. Other evidence of the basic level being the preferred level of categorisation is that objects are more quickly categorised as members of basic level categories than as members of other categories. In one experiment, informants heard the name of a category and were then shown an image; they had to decide whether the image matched the word. Reaction times proved consistently shorter when the test involved a basic level term ROSCH *et al.* [1976], in EVANS & GREEN : 262]. Finally, basic level terms have been shown to be typically monolexemic, that is, comprised of a single word-unit, and shorter than subordinate terms [TAYLOR : 49], [EVANS & GREEN : 263].

As shown by the few examples cited here, the experiments were mainly carried out in metalinguistic contexts. The question is whether the notion of basic level is still relevant in discourse. If it is the preferred level of categorisation, it can be expected to correspond to unmarked rephrasing.

The basic level is indeed the preferred level of categorisation for anaphoric definite descriptions in the corpus. There are nouns of this level in all cases of reduction, in most cases of identical repetition of a minimal lexical content, in half the cases of use of a hyperonym and in all cases of synthesis via nominalisation. Basic level categorisation, however, is not the *only* choice criterion for lexical information in the corpus. 10 utterances show either a superordinate word (4 occurrences), such as *the beast*, which rephrases *the dog* [HOLDSTOCK : 34], or a subordinate word (6 occurrences),

for instance *the uninitiated stranger* after *a young man making his first visit to France* [PAXMAN : 24], or *the cut-glass tumbler* after *the glass* [DEXTER : 23]. What is also important in these three examples is that the antecedent NP does display a basic level noun; yet it is not retained in the anaphor. It must be concluded that contextual factors are also at play.

The data in the corpus also shows that the very notion of basic level is not always an operating tool. The various metalinguistic experiments cited in Rosch *et al.* [1976], Taylor [1995], Croft & Cruse [2004] and Evans & Green [2006] present categories with an easily identifiable basic level. Besides, they are all based on classes of concrete items of everyday life: furniture, food, artefacts, vehicles, natural objects, ... Outside that area, it sometimes proves difficult to decide whether a category is a basic level one. It is the case, in particular, with human beings, due to the possible alternation between a proper name and *the + common noun*. For instance, when *Hillary Clinton* is rephrased as *the former first lady*, should it be considered that the occupational title is by nature a superordinate term, with the basic level instantiated by the proper name? Or do both phrasings correspond to the basic level? Another case in which the basic level is difficult to identify is with some common nouns, as illustrated by (3):

- (3) (DCI to his assistant) 'Something for you, Lewis. Remember the Lower Swinstead murder ?'
 'Well, vaguely, yes. And I've seen the bits in the paper, you know, about the calls. I was never really on the case myself though.' [DEXTER : 31]

The criteria set by Rosch *et al.* [1976 : 382] would lead to consider *murder* as a basic level term; for instance, the answer to the question 'What happened?' or 'What is this?' could be 'There has been a murder', or 'This is a case of murder', but not 'There has been / There is a case'. *Case*, which is a hyperonym of *murder* since it is more inclusive, should therefore be regarded as a superordinate word. However, in the context of utterance (3), it is difficult to regard *case* as a superordinate word. Indeed, *be on the case* forms a collocation, which makes *case* the preferred, 'unmarked' level of categorisation at this point.

Finally, the concept of basic level proves problematic when confronted to the corpus because it applies solely to nouns. This raises the question of how NPs should be treated when the nouns are used along with modifiers or complements. For instance, in (2), the NP *the cavernous, mock-gothic church* shows the basic level noun *church*, but it cannot be concluded that there is unmarked rephrasing with this NP, as the qualities added with the adjectives convey new information to the addressee. While this example

confirms that basic level categorisation is indeed a specific level in vertical categorisation (the speaker retains *church* rather than use a hyperonym such as *building*), it also shows that the sole notion of basic level categorisation is not sufficient to account for lexical choices in anaphoric definite descriptions.

This conclusion leads us to reconsider the theoretical definition of basic level given by Rosch et al. (390 ff., in EVANS & GREEN : 256): it is identified as the level that enables optimum cognitive economy. Could it be that even though other contextual factors than the criteria given to determine basic level in metalinguistic contexts are at work, this more general principle of optimum cognitive economy might still hold in context? In order to answer this question, the study now turns to the criteria for lexical choices in the corpus in the light of a theoretical framework that makes use of the idea of cognitive economy in context: Relevance Theory.

Lexical information and the Relevance principle

The idea of optimum cognitive economy is at the core of Relevance Theory, as laid out by Sperber & Wilson [1986], where it is better phrased as 'maximum cognitive efficiency'. The notion of cognitive efficiency is based on the tenet that 'all human beings automatically aim at the most efficient information processing possible' [*ibid.* 49]. Sperber & Wilson show that interpreting an utterance does not simply involve identifying the assumption expressed by the speaker: the hearer also works out the consequences of adding this assumption to a set of assumptions that have already been processed [*ibid.* 118]. This working out of a connection between a given assumption and previous assumptions conveys 'contextual effects'. These may consist in the triggering of implications, in a modification of the strength of some old assumptions, or in erasure of old assumptions [*ibid.* 117]. For instance, in the following dialogue:

- (4) Mary: 'What I would like to eat tonight is an osso-bucco.'
Peter: 'I had a long day. I'm tired.'

Peter's remark creates the assumption *Peter is tired*. From this and the premise that *if Peter is tired, he wishes Mary would make dinner*, Mary forms the assumption *Peter wishes Mary to make dinner* [145]. In this case, the contextual effect is the triggering of an implication (the last assumption).

From this, maximum cognitive efficiency is defined as the optimum balance between contextual effect and processing effort: 'Human cognitive processes, we argue, are geared to achieving the greatest possible cognitive effect for the smallest possible processing effort.' [*ibid.* vii]. The notion of maximum cognitive efficiency is directly related to that of relevance.

Relevance is defined as the fact that an item of information has a contextual effect in the context in which it is used; an item of information will be regarded as more relevant than another if its contextual effect in a given context is larger and the processing effort required smaller [*ibid.* 125]. This applies to speech as well as to perception. For instance, Mary and Peter are sitting on a bench, and as Peter leans back, he reveals three people: an ice-cream vendor that they had already noticed, an ordinary stroller and their acquaintance William, whom they find boring. The stroller is new to Mary, but he brings no contextual effects in the interpretation of her environment; so what she can infer is said to be of little relevance to her. On the contrary, on seeing William, she can draw many conclusions, from which many other conclusions will follow. Consequently, this is the most relevant change in her cognitive environment [*ibid.* 49].

A key idea of Relevance Theory for the analysis of the present corpus is that addressees always assume that 'relevance has been attempted' [*ibid.* 159]. In other words, there is a presumption of relevance in any speech act. Besides, a speaker is typically expected to maximise the relevance of the information to be processed [*ibid.* 49], so that no information, including the lexical information, is deemed gratuitous. This echoes Grice [1975]'s cooperative principle, more specifically the maxims of quantity and relevance:

(5) Maxims of quantity:

1. Make your contribution as informative as required.
2. Do not make your contribution more informative than is required.

Maxim of relevance:

Be relevant.

But presumption of relevance is shown to be more than norms contributing to successful communication, as Grice's maxims suggest. They are a generalisation about ostensive-inferential communication, which therefore applies without exception [SPERBER & WILSON : 162].

The question to be addressed now is whether the Relevance principle, which Sperber & Wilson [1986] apply to whole propositions, also applies within the definite descriptions under study. If relevance is the criterion for information selection, maximum cognitive efficiency will be achieved either if the information given is the minimal required for referent retrieval, or if the additional information has contextual effect, that is, if its processing triggers inferences.

Let us start with the cases of rephrasing in which the speaker adds information to that contained in the antecedent, more specifically those in

which that information is already familiar to the addressee thanks to elements of the context. A first criterion for this addition is illustrated by (6):

(6) At Yale an experimenter put two white rats in identical cages side by side. If either would learn to press a little iron bar, a food pellet would drop into its cage. *One of the rats* was very hungry. *The other* had just had a full meal.

The hungry rat, eagerly exploring his cage for possible food, soon accidentally hit the bar. [...]

Meanwhile the full-bellied rat in the cage a few inches away was lying quietly in its corner digesting its food [...] [PACKARD: 19]

The adjectives *hungry* and *full-bellied* take up the information given in the predicates *was very hungry* and *had just had a full meal*. They are essential because they directly play a part in singling out the referents: the sole noun *rat* would be ambiguous because it would select two potential entities. Relevance Theory applies here: any more information on the rats would not be relevant in the context, because the focus is on the influence of a variable (having had food or not) and not on the individuality of the rats beyond that variable. *The hungry white rat*, for instance, would be inappropriate because *white* would convey no contextual effect.

This first choice criterion – viz. giving the minimal information required for referent selection – is the only one found in the spoken occurrences of the corpus (6 occurrences), whereas in the written documents, it only concerns 3 utterances out of 10. Extract (7) presents another criterion:

(7) (first lines of a newspaper article) A brilliant young rogue trader, who spun an elaborate web of fake transactions from his desk, has cost France's second-biggest bank €4.9bn (£3.7bn) in what appears to be the largest-ever fraud by a single trader.

Société Générale was last night struggling to shore up confidence in the banking system after the huge fraud was conducted undetected by the junior trader at its headquarters. [CHRISAFIS et al.]

Huge and *junior* are not necessary for referent retrieval. Reminding of those qualities, though, is a way of emphasizing the contrast between the extent of the fraud and the low rank of the frauder, so as to foreground the extraordinariness of the event, in combination with *undetected*. One can also see in this contrast an explanation for the main predicate (*struggling*). In other words, these adjectives appear to have been chosen because the information they bring serve the overall propositional content: without them, the addressee would not have included the contrast in the representation of the event. The network formed by the three words therefore triggers contextual effect, in the form of a foregrounding of a

previous assumption. In that sense, the information is relevant. Similarly, in (8):

(8) the 'possum's reputation is undeserved. He is the dim-wit of the pair. (...) He is too dumb to be an actor. / In a sense the opossum is a living fossil. He is the only surviving marsupial in North America. (...) How the inefficient, disorderly-looking opossum with its small brain in a long, pointed skull has survived in America is an intriguing question. [PACKARD: 50]

The information denoted by the two underlined adjectives is taken up from the leftward co-text: *inefficient* was suggested by *dim-wit* and *dumb*, while *disorderly-looking* is a synthesis of a description given ahead of the passage. As in (7), *inefficient* and *disorderly-looking* do not serve to single out the referent (there is just one generic opossum), but act as a reminder to justify why the question of its survival should be 'intriguing'.

Relevance also applies to most cases in which the anaphor conveys information that is totally new to the addressee, as in (9):

(9) Memory of the tranquil mill-pond haunted me, and I moved away from the south track, through waist-high nettles and a tangle of ash and hawthorn scrub. I came out close to the bank of the wide, shadowy pool, its full extent hidden by the gloom of the dense stand of oak woodland that began on its far side. Almost hidden among the rushes that crowded the nearer edge of the pond was the shallow boat from which Chris and I had fished, years before. [HOLDSTOCK : 18]

The first sentence suggests the narrator's fascination for the pond, which prompts his actions. The pond is also foregrounded to some extent, as it is part of the subject (*memory of the tranquil mill-pond*). In that context, giving the information *wide* and *shadowy* is interpreted as relevant for the overall message: the narrator seems to be retrieving the features he experienced as a child. In the last sentence, on the other hand, the focus is now on the boat, which triggers fishing memories. At that point, therefore, it is the reference to the boat that shows additional information (*the shallow boat*), while the anaphor referring to the pond is reduced to *the + basic level noun (the pond)*.

Cases such as (9), in which new information is added, also raise the question of how coreference between the anaphor and the antecedent is established by the addressee. The occurrences of the corpus show at least two conditions. On the one hand, the head noun in the anaphor must be compatible with that of the antecedent; in (9), for instance, *pool* is a hyperonym of *pond*. But other elements in the context, too, orient towards the referent. *And*, which here conveys consequence, leads the addressee to

assume that the narrator's journey will take him to the pond; *bank* then introduces the semantic field of rivers and lakes, so that all it takes is for the NP that follows *the bank of* to be definite for the reader to expect coreference. This could explain why in the corpus, 7 NPs out of the 9 that display totally new information are in post-verbal position (direct object or prepositional complements), as in (9), with elements earlier in the same sentence orienting towards coreference. For instance, in (10):

- (10) Strange's tone was semi-peremptory as he thrust a folded sheet of ruled A4 across at Morse, in the process knocking his glass on to the parquet flooring. Where it broke into many pieces.
 'Ah! Sorry about that!'
 Morse rose reluctantly to fetch brush and pan from the kitchen.
 'Could have been worse, though,' continued Strange. 'Could have been full, eh?'
 As Morse carefully picked up the slivers of the cut-glass tumbler... he experienced an irrational anger [...] [DEXTER : 23]

Although *cut-glass* and *tumbler* convey new information, *the* and the leftward co-text impose a coreferential reading of *the cut-glass tumbler* and the antecedent *his glass*: the predicate *fetch brush and pan from the kitchen* conveys the assumption that the broken glass is about to be picked up, so that *picked up the slivers of* necessarily implies that the complement of *of* will be an NP that refers to that same glass. As a consequence, the noun in the anaphor does not even need to denote a glass, as it does here; if the context had been relevant, one might have found, for instance, *picked up the slivers of his mother's wedding gift to him* and still understood the referent to be the tumbler. The only restriction is that the information be compatible with the representation of a broken tumbler.

When the anaphor that contains new information is the syntactic subject (which is the case in only 2 occurrences in the corpus, both in press articles), and is therefore placed before the predicate, this type of pre-orienting cannot occur. Consequently, these are extreme cases, and if the addressee reads too quickly, the coreference might even be missed, as in (11):

- (11) Leeson, who single-handedly brought down one of the world's oldest banks, said the shocking scale of the SocGen fraud proved financial institutions had not learned their lessons.
The now chief executive of Galway United football club told the BBC:
 '[...]' [CHRISAFIS *et al.*]

The adverb *now* alone orients towards coreference, in that it implies to look for a referent that used to have another occupation.

In this case, can one still talk of maximum cognitive efficiency? The answer is no if one considers the huge effort required from the addressee – which might explain why this type of phrasing was not found in spoken discourse. Less effort would be required, for instance, if a relative clause had been used: *Leeson, who is now chief executive...* But there is maximum cognitive efficiency if one considers the resulting phrasing and the genre in which the NP is used. First, this choice of phrasing enables the information to be compressed. With a single NP, Leeson is referred to and new, relevant information is given about him: the addressee infers that despite the fraud, Leeson has made it. Besides, as this type of re-phrasing is typical of the press, readers might expect this sort of processing, and so require less effort to retrieve the referent than in other speech situations.

Finally, the principle of relevance of the lexical information at the point when the anaphor is uttered holds for nearly all occurrences of the other types of phrasing in the anaphoric definite descriptions of the corpus. But it should be noted that in a handful of utterances, other reasons than relevance are the primary force for the choice of lexical information. The corpus shows two. The first is avoidance of repetition, through a hyperonym or a hyponym, as in (12):

(12) I took the meat, half a granary loaf, and a jug of bottled beer out into the yard. The stranger was crouched now, the hound lying beside him, reluctantly, it seemed to me. As I tried to approach them, the dog growled, then barked in a way that set my heart racing and nearly made me drop my gifts. The man shouted at the beast, and said something to me. [HOLDSTOCK : 35]

The alternation between *dog*, *hound* and *beast* seems to be motivated mostly or solely by the wish to avoid repetition. The choice of *dog* does not appear to be more relevant than, for instance, *beast* or *hound*. In the last sentence, *beast* does not seem more relevant than the other two nouns: it usually connotes more distance than *dog*, for instance, but no element in the context of the last sentence orients towards added distance. Similarly with *the stranger*, which is used many times by the author in the novel to alternate with *the man*: nothing in the context justifies the change of noun other than a wish to avoid repetition.

The other criterion that is not dictated by relevance is what RANGER (2002 : 87) terms cases of 'disposable anaphora' (*anaphore jetable*), and which we will call 'decorative effect'. These are typical of the press, and involve phrases such as *the former first lady* for *Hillary Clinton*. In the occurrences of the corpus, the choice of one phrase over the other sometimes appears to

correspond more to conventional naming than to an influence of the immediate context, as in (13):

(13) 'Iran's nuclear program remains a serious concern but it is clear from the [intelligence estimate] that vigorous and coordinated diplomacy is the right way to approach it,' Feingold said.
The spectre of war with Tehran has trailed Democratic presidential candidates throughout the campaign, particularly New York senator Hillary Clinton. The former first lady's foes have criticised her for backing a September resolution that linked the US military presence in Iraq to the behaviour of Iran, creating what liberals considered a new justification for war. [SCHOR]

Such choices, however, remain compatible with the principle of maximum cognitive efficiency: in the name of variation, the speaker considers that a different set of lexical information is required, and thus that taking this constraint into account, the processing effort has been kept to a minimum.

As the ultimate choice criterion for lexical information in definite descriptions is maximum cognitive efficiency, and as this is typically achieved through their relevance *at the point when the anaphor is mentioned*, a final question is, what is the true influence of the antecedent NP on the anaphor?

Influence of the antecedent NP on the anaphor

In its narrow, textual sense, anaphor is defined as a 'relation between two linguistic expressions in discourse'⁹ [ZRIBI-HERTZ : 603, *in* GIRARD : 6]. The anaphor is traditionally viewed as referentially non autonomous [CORNISH : 41] because it is interpreted only via the antecedent, which saturates it referentially by transferring its lexico-semantic properties [MILNER : 20].

But evidence from the corpus shows that, as proposed by Cornish (*ibid.*), the antecedent NP has a more limited influence. The most obvious evidence is that information can be added to that contained in the antecedent phrase, whether that information is new or has been introduced ahead. Besides, example (9) suggested that the referent's status in discourse could evolve from the mention of the antecedent to that of the anaphor; the lexical information given in the antecedent might therefore no longer be necessary or relevant as the anaphor is uttered. It is therefore inappropriate, for instance, to say that textual anaphora consists in taking up a given segment 1 by an (anaphoric) segment 2; in rephrasing, there is more at stake than the

⁹ « [U]ne relation entre deux expressions linguistiques au sein du discours ».

information given in the antecedent itself. Cornish therefore reduces the role of the antecedent NP to that of an antecedent trigger, following Hawkins (*trigger*) and Reichler-Béguelin (*introduceur textuel*) [CORNISH : 43]. This role is defined as follows:

(14) The antecedent trigger introduces an entity into the discourse via its predicational and utterance context, and an anaphor of a particular type and form accesses that mentally represented discourse entity at a later point in the discourse, adding to this representation further properties resulting from the processing of the anaphoric clause as a whole. [*ibid.* 44]

In other words, the indirect relation mentioned in the anaphoric mechanism – the fact that the anaphor is interpreted via its antecedent – mainly concerns the process through which the referent becomes part of the representation: it is the antecedent trigger (if it is a first mention) which establishes its existence and makes the referent a given when the anaphor is uttered. In anaphora processing proper, however, there is not necessarily indirect access to the referent in the mental representation.

Saying that the antecedent trigger's influence is limited, however, should not lead to deny any influence. In particular, as the antecedent trigger denotes a set of properties of the referent, any rephrasing must be compatible with those properties, as shown above. Only, the antecedent NP is just *one* of the elements that might influence the phrasing of the lexical information in the anaphor: others include the context, the perspective of the message or even the decorative effects the speaker is looking for. In spoken discourse, the role of the antecedent trigger might be more significant. In the transcriptions from lectures collected in the corpus, for instance, an element referred to as *the sheet* is always referred to as *the sheet* (rather than *the hand-out*, for instance). This may be simply because the noun *sheet* is a basic level term; but an initial phrasing might also favour a repetition of the same noun, which requires less cognitive effort than having to find a synonym. A broader-scale study of spoken uses would lead to a better understanding of this aspect.

Conclusion

In coreferential textual anaphora, the criterion for lexical selection in definite descriptions appears to be maximum cognitive efficiency, which is typically achieved through relevance for the overall message at the point when the anaphor is uttered. THEISSEN [1997] reaches a similar conclusion in her study of anaphora in French when she proposes the criterion of adequate information (*apport adéquat d'informations*), but the notions of cognitive efficiency and relevance seem to us to better account for what is at stake

than 'adequate'. What the study also confirms is that the fact that an anaphor is related to a textual antecedent should not minimise the autonomy of the anaphor: the phrasing of the antecedent trigger is in fact of limited influence. The notion of basic level of categorisation is not sufficient either, and does not seem to exist for all categories, although the existence of such a level is obvious for some of them. Finally, the study suggests an influence of genre on the use of lexical information, the extent of which can only be established with more extensive data.

Appendix: documents used for the corpus

NB. All links to websites were checked on 5 March 2013

The occurrences were collected from samples of the following documents:

- * (essay) PAXMAN, Jeremy. *The English: A Portrait of a People*. London: Penguin, 1999.
- * (fiction, novel) HOLDSTOCK, Robert. *Mythago Wood*. London: HarperCollins, 1998.
- * (fiction, detective story) DEXTER, Colin. *The Remorseful Day*. London: Macmillan, 2000.
- * (forum) British expat discussion forum – vacuum cleaners section. <<http://britishexpats.com/forum>>
- * (forum) Car forum. <<http://car-forums.com>> (accessed 10/09/2008)
- * (lectures) http://www.rdg.ac.uk/AcaDepts/ll/base_corpus (accessed 07/01/2009)
- * (press) BBC News, 'Fire-rescue nanny goat has kids'. *BBC News*, 23/05/2006. <http://news.bbc.co.uk/2/hi/uk_news/england/oxfordshire/5007512.stm> (accessed 12/09/2008)
- * (press) CHRISAFIS, Angelique, Jill TREANOR and Katie ALLEN. 'He lost his bank £3.7bn. So was it his fault the markets crashed?' *The Guardian*, 25/01/2008 <<http://www.guardian.co.uk/business/2008/jan/25/banking.france>> (accessed 10/09/2008)
- * (press) JA Alok, 'Biologist claims significant step towards artificial life'. *The Guardian*, 25/01/2008

< <http://www.guardian.co.uk/science/2008/jan/25/genetics.science>>

(accessed 10/09/2008)

* (press) NASAW, Daniel, 'Hillary's speech by numbers'. *The Guardian*, 27/08/2008

<<http://www.guardian.co.uk/world/2008/aug/27/usa.uselections2008?INTCMP=SRCH>> (accessed 10/09/2008)

* (press) SCHOR, Elana, 'Democrats welcome "check" on White House Iran policy'. *The Guardian*, 3/12/2007

<<http://www.guardian.co.uk/world/2007/dec/03/usa.iran?INTCMP=SRCH>> (accessed 10/09/2008)* (science for the general public)

PACKARD, Vance. *The Human Side of Animals : Animal IQ*. New York: Pocket, 1950.

* (science handbook) MAITLAND, Peter S. & R. Niall CAMPBELL. *Freshwater Fishes of the British Isles*. London: HarperCollins, 1992.

* (science handbook) WILLSON, Mary. *Vertebrate Natural History*. Philadelphia: Saunders College, 1984.

* (transcript of interview) 'Bill Hall', *Scope*, 2006. <http://www.speakingforourselves.org/uk/index.php/interviews/bill_hall> (accessed 10/09/2008)

* (website of St Paul's Cathedral) Dean and chapter of St Paul's, 'About St Paul's', 2008. <<http://www.stpauls.co.uk>> (accessed 12/09/2008)

Works Cited

- BEAVER, David I. « The optimization of discourse anaphora ». *Linguistics and Philosophy* 27-1 (2004) : 3-56.
- CHARAUDEAU, Patrick & MAINGUENEAU, Dominique. *Dictionnaire d'analyse du discours*. Paris : Seuil, 2002.
- CHARREYRE, Claude. « Perspective énonciative sur l'anaphore nominale ». *Cycnos* 18.2, *Anaphores nominale et verbale*. Nice : CRELA, 2001.
<http://revel.unice.fr/cycnos/index.html?id=100> (accessed 5 March 2013)
- CORBLIN, Francis. *Les formes de reprise dans le discours : Anaphores et chaînes de référence*. Coll. *Langues et discours*. Rennes: PU de Rennes, 1995.
- CORNISH, Francis. *Anaphora, Discourse and Understanding : Evidence from English and French*. Oxford: Clarendon, 1999.
- CROFT, William & CRUSE, Alan. *Cognitive Linguistics*. Cambridge : University Press, 2004.
- DUCROT, Oswald & SCHAEFFER Jean-Marie. *Nouveau dictionnaire encyclopédique des sciences du langage*. Paris : Seuil, 1995.
- EVANS, Vyvyan & GREEN, Melanie. *Cognitive Linguistics : An Introduction*. Edinburgh: University Press, 2006.
- FOX, Barbara. *Discourse Structure and Anaphora : Written and Conversational English*. Cambridge : University Press, 1987.
- FUCHS, Catherine, dir. *La linguistique cognitive*. Paris : Ophrys, 2004.
- GIRARD, Geneviève. « Les paramètres à la construction des anaphores ». *Cycnos* 18.2, *Anaphores nominale et verbale*. Nice : CRELA, 2001.
<http://revel.unice.fr/cycnos/index.html?id=100> (accessed 5 March 2013)
- KLEIBER, Georges. *L'anaphore associative*. Paris : PUF, 2001.
- MAES, Alfons A. *Nominal Anaphors : Markedness and the Coherence of Discourse*. Leuven: Peeters, 1996.
- MILNER, Jean-Claude. *Ordres et raisons de langue*. Paris : Seuil, 1982.
- MOESCHLER, Jacques & REBOUL Anne. *Dictionnaire encyclopédique de pragmatique*. Paris : Seuil, 1994.

- RANGER, Graham. « Notes sur la traduction de l'anaphore lexicale du français vers l'anglais ». *Langues et cultures en contact : Traduire e(s)t commenter*, dir. Catherine PAULIN, collab. Philippe RAPATEL. Besançon : PU franc-comtoises, 2002.
- ROBERTS, Craige. « Uniqueness in definite NPs ». *Linguistics and Philosophy* 26-3 (2003) : 287-350.
- ROSCH, Eleanor *et al.* « Basic objects in natural categories ». *Cognitive Psychology* 8 (1976) : 382-439.
- ROTGE, Wilfrid & LAPAIRE, Jean-Rémi. *Réussir le commentaire grammatical de textes*. Paris : Ellipses, 2004.
- SPERBER, Dan & WILSON, Deirdre. *Relevance : Communication and Cognition*. Cambridge, Massachusetts: Harvard University Press, 1986.
- SZABÓ, Zóltan Gendler. « Descriptions and uniqueness ». *Philosophical Studies : An International Journal for Philosophy in the Analytic Tradition* 101-1 (2000) : 29-57.
- TAYLOR, John R. *Linguistic Categorisation*, 2nd edition. Oxford : University Press, 1995.
- THEISSEN, Anne. *Le choix du nom en discours*. Genève : Droz, 1997.
- UNGERER, Friedrich & SCHMID, Hans-Jorg. *An Introduction to Cognitive Linguistics*. London: Longman, 1996.
- ZRIBI-HERTZ, Anne. « De la deixis à l'anaphore : quelques jalons ». In MOREL and DANON-BOILEAU (eds.), *La deixis*. Paris : PUF, 1992.