# Constraints on metalinguistic anaphora

## Philippe De Brabanter

▶ **To cite this version:**

Philippe De Brabanter. Constraints on metalinguistic anaphora. 2007. <ijn_00083956v2>

## HAL Id: ijn_00083956

## http://jeannicod.ccsd.cnrs.fr/ijn_00083956v2

Submitted on 6 Dec 2008

# Constraints on metalinguistic anaphora[1]

## Philippe De Brabanter

Institut Jean Nicod, Université Paris 4-Sorbonne

Just as it is possible to refer to any entity, concrete or abstract, in extralinguistic reality, it is also possible to refer to any linguistic entity, be that a phoneme, a word, a sentence, or any other linguistic object. Metalinguistic reference can be achieved in two ways, by means of 'autonymous' and 'heteronymous' mention (cf. Recanati 2000: 137). Autonymous mention is a matter of quotation: a token of a linguistic string can be produced in order to refer to another token of that string or to a type which it instantiates:

(1)　　"**Boston**" is disyllabic. (Quine 1940 :26)

(2)　　She said **"why don't you just drop dead?"**

Heteronymous mention concerns all the non-quotational means that can be resorted to in order to refer metalinguistically, namely descriptions and various types of pronouns.

(3)　　**The 4354th word of Chants Democratic** is disyllabic. (Quine 1940 :26)

The central difference between autonymous and heteronymous mention is that the former rests on an iconic relation between the mentioning expression and its linguistic referent, whereas the latter involves no such resemblance at all.

---

[1] I am grateful to the editors for allowing me to submit this paper with a year's delay. I also wish to thank two anonymous reviewers for their useful suggestions and for helping me improve the design of this paper.

The focus of this paper is on a subset of heteronymous mention, namely those cases in which the mentioning expression is, roughly speaking, anaphorically linked to the string it mentions. I will distinguish two subclasses. In the first one, the antecedent of the metalinguistic anaphor is a quotation. This means that both the antecedent and the anaphor refer to a linguistic entity (the same one, it turns out; these expressions are co-referential). In the second subclass, the antecedent is not a quotation; it is a string in ordinary use. Here we have no co-reference: whereas the anaphor refers metalinguistically, the antecedent either refers to an object in the world or does not refer at all. This second subclass is especially interesting because it instantiates a shift in the universe of discourse, from extralinguistic reality to language. Where such a shift occurs, I will speak of 'world-to-language' anaphora.

I will argue that metalinguistic anaphora is best described in terms of a theory that assumes that various anaphoric expressions encode various degrees of salience of referents. But I will also show that salience is built in the context of utterance. It is not necessarily an acquired feature of the referent by the time the anaphor is processed: there is adjustment between the anaphor and its immediate linguistic environment. Besides, we will see that other factors may also affect anaphora resolution, which suggests that the best account must, in essence, be pragmatic.

## 1. Formal varieties of metalinguistic anaphora

I start with a couple of examples of 'unshifted' metalinguistic anaphora:

(4)    "Harry said 'I didn't do that' but he said it in a funny way"

(5)    "'You are wrong'. That's exactly what she said"

Both examples are from Levinson, who writes about (4) that "*it* does not refer to the

proposition expressed but to Harry's utterance itself" (2004: 103). Note that *'I didn't do that'*, being a direct quotation, itself refers to Harry's utterance. Thus, *it* and its antecedent are co-referential and, in that respect (4) – like (5) – contains an ordinary case of anaphora.

World-to-language anaphora is less well-behaved. There is no co-reference between the antecedent and the anaphor, in which respect world-to-language anaphora belongs with bridging (cf. Clark 1977). There is, however, a major difference between bridging and metalinguistic anaphora: in all varieties of bridging that I am aware of, there are strong constraints on the linguistic form of the antecedent and of the anaphor. Thus, to mention two examples, Kleiber's (1999) 'associative anaphors' must be definite NPs with an NP-antecedent, and Erkü and Gundel's (1987) 'indirect anaphors' must be non-pronominal NPs whose antecedent quantifies over individuals, events, situations and facts. As we will see shortly, there are no such constraints on world-to-language anaphora: the antecedent merely needs to have been uttered (not too long before processing of the anaphor – this is a minimal temporal or spatial constraint) and the anaphor can, at first blush, be an indefinite NP, a definite NP, a demonstrative NP or pronoun, a relative or interrogative pronoun, and even an unstressed personal pronoun. Here are illustrations of each of these cases:

(6)   They genuinely tried to become, to use **a horrid word**, acculturated with the white invaders, even if they had no desire to be assimilated. (BNC AJV 758)

(7)   I still remember the day, before he was repatriated (Ray explained the meaning of **the word** to me very carefully) back to Paris by the French Government for treatment […].

   (www.brain.riken.go.jp/bsi-news/bsinews18/no18/networke.html)

(8)   A: I think of him as a family man.

B: Funny, I've always considered **that phrase** an oxymoron. (Julian Barnes 1998, *England, England*, Picador, p. 64)

(9) 'Yeah, you're all right. But you're not perfect, and you're certainly not happy. So what happens if you *get* happy, and yes I know **that**'s the title of an Elvis Costello album, I used the reference deliberately […]'. (Nick Hornby 1995, *High Fidelity*, Indigo, p. 223)

(10) Yes, everything went swimmingly, **which** is a very peculiar adverb to apply to a social event, considering how most human beings swim. (Julian Barnes, *Love, etc.*, Picador, pp. 70-71)

(11) It means nothing to you, I suppose, he said, it was just a, **what** do they call it, a one-night-stand. (David Lodge, *Nice Work*, Penguin, p. 297)

(12) After several hours of bouncing from one bureaucrat (notice **it**'s a French word) to another I was allowed into the hallowed chambers. (www.vt-fcgs.org/miscinfo.html)

In all these instances, there is a strong connection between a heteronymously mentioning expression ('heteronym', for short)[2] and some string occurring in the co-text. To that extent the strings in question can be regarded as antecedents. I wish to argue, however, that indefinite NPs are different and are not in fact anaphors. The metalinguistic NP in (6), *a horrid word*, occurs as part of an elliptical parenthetical clause. When completed, that clause is something like *I am going to use a horrid word*, and it has truth-conditions to the effect that the speaker has to use a horrid word. There is no constraint on *which* horrid word should be used: the word *acculturated* is not part of the truth-conditions of the parenthetical. In other words, the heteronym does not substitute anaphorically for its 'antecedent'. Actually, this is the conclusion one is led to every time an indefinite NP is used metalinguistically. Here are

---

[2] The boldtype for the heteronyms is an addition of mine.

two more examples:

(13)  The copper-haired woman, meanwhile, had almost canceled a Hawaii trip because of fear of terrorists (**a word** she pronounced with two syllables, like Laura Bush), […]. (starbulletin.com/2002/02/17/travel/story2.html)

(14)  And as the books about Peter Cook state, he was heavily influenced by the satirical nightclubs (always **an odd phrase**, to my mind) of France and Germany. (groups.yahoo.com/group/peter_cook/message/4052).

Here again the indefinite metalinguistic NPs occur in elliptical parentheticals. In the fleshed out clauses (*This is a word which she pronounced with two syllables* and *This is always an odd phrase, to my mind*), the indefinites do not refer to the strings *terrorists* and *satirical nightclubs*. This is all the clearer here because those clauses do include heteronymous anaphors to these expressions: the demonstrative pronouns that have been filled in. It is those demonstratives, not the indefinite metalinguistic NPs that contribute the strings *terrorists* and *satirical nightclubs* to the truth-conditions of the fleshed out clauses.

Things are significantly different in all of the other examples (7-12). There, the heteronym is necessarily interpreted in terms of the antecedent. For instance, what *that phrase* contributes to the truth-conditions in (8) is the expression *family man* and what *which* contributes to the truth-conditions in (10) is the word *swimmingly*. Therefore, in the rest of this paper, I will exclude examples like (6), (13) and (14) from the study of metalinguistic anaphora.

## 2.  Constraints on the antecedent and referent of the antecedent

We have just seen that the form of metalinguistic anaphors is relatively unconstrained. But are there perhaps stronger constraints on *the sort of antecedent* that these anaphors can take, or on *the referent* of the antecedent? First, what few constraints there are on the antecedents are not very severe: the antecedent must (i) be a linguistic expression and (ii) be close to the world-to-language anaphor.[3] As for the referent of the antecedent, there are no restrictions on it … because there is not even a requirement that the antecedent should *have* a referent. Take (7) and (9) again: neither *repatriated* nor *get happy*, which are the respective antecedents of *the word* and of *that*, have a referent. This does not prevent these expressions from functioning as antecedents of the anaphors.

## 3.  A top-down approach

The above observations mean that there is little sense in working out a bottom-up account based on the sorts of forms that enter into metalinguistic anaphora. It makes better sense to approach it top-down, starting from a general cognitive or pragmatic principle. There are at least two theories in the literature that attempt to ground anaphora resolution on a cognitive principle. One is Mira Ariel's 'Accessibility Theory' (1988, 1991), the other Gundel et al.'s 'Givenness Hierarchy' (1993, 2003; Borthen et al. 1997). Both share the view that different types of grammatical forms or constructions (personal or demonstrative pronouns, definite descriptions, etc.) encode different degrees of salience[4] of referents. In other words, the

---

[3] How close is an important question that I cannot answer now. Some work in psycholinguistics provides pointers, however. Levelt (1989: 122) writes that it is likely that "in conversation, literal recall not supported by salient content or pragmatic significance is short-lived, probably going back only as far as the last clause uttered". Note too that *what* in (11) precedes its 'antecedent', indicating at least that the antecedent need not always come immediately before the anaphor (which some would therefore call a *cataphor*. Still, I will follow a widely accepted tradition (see e.g. Huddleston & Pullum 2002: 1455) and use *anaphor* regardless of the position of the antecedent).

[4] Other possible terms here include *accessibility* and *activation*. However, for the sake of convenience, I will use *salience* throughout.

required degree of salience of a referent is part and parcel of the lexical meaning of referring expressions in general and anaphors in particular.

In the following, I will refer mainly to the Givenness Hierarchy, though the discussion could be extended to Ariel's Accessibility Theory.[5] The Givenness Hierarchy relies on the notion that the choice of an anaphoric form reflects the speaker's assumptions about how salient the referent of the antecedent is to the hearer (how easily recoverable it is). Gundel et al. distinguish six levels pairing a 'cognitive status' with linguistic forms. I illustrate levels 3 to 6, the other two being irrelevant to my present purposes as they are not available in metalinguistic anaphora.

As one moves up the hierarchy, one encounters forms that place increasing constraints on the cognitive status of the referent:

(15) [lev 3] Steve's car let him down yesterday. **The battery** was dead.

(16) [lev 4] Steve's car let him down yesterday. **That battery** was dead.

(17) [lev 5] ??Steve's car let him down yesterday. **That** was dead.

(18) [lev 6] ??Steve's car let him down yesterday. **It** was dead.

(15) is a basic instance of bridging. All that is required for the referent of *the battery* to be accessed is that it be 'uniquely identifiable' by the hearer. It must be recoverable as a distinct object in the context of utterance, but need not have been represented in the hearer's mind to begin with. The only condition is possession of a mental frame specifying that cars have batteries. For (16) to be acceptable, we need something more: the referent must be 'familiar' to the hearer, i.e. represented in his long-term memory. The only way to interpret *that battery* here is as *that battery we've already talked about* or some such phrase. As for (17) and (18),

---

[5] Ariel views accessibility as a property of the mental representations of referents, rather than of referents themselves. But that difference has no impact that I can see on the account given here.

it is very unlikely that they license the intended interpretation, i.e. that on which the anaphor refers to the battery in Steve's car. According to Gundel et al., that would be because the previous co-text is insufficient to enable the battery to be 'activated' (i.e. placed in short-term memory) at level 5 or 'in focus' at level 6 (i.e. "at the current center of attention", among the "entities which are likely to be continued as topics of subsequent utterances" (Gundel et al., 1993: 279)).

The cognitive constraints on the highest level in the hierarchy are quite severe. Thus, Gundel et al. have shown that even explicit mention does not guarantee in-focus status. Consider the next pair of examples (assume that they are uttered with the same unmarked intonation):

(19)   I've just bought a parrot from Peru. **It**'s a wonderful bird.

(20)   ?? I've just bought a parrot from Peru. **It**'s a wonderful country.

After the first sentence in (19) has been processed, the parrot from Peru is in focus because the phrase mentioning it occurs as a direct object of *bought*, a position which, like the subject position, is capable of bringing a referent into focus. By contrast, (20) is less clearly acceptable (or requires an enriched context, in which, for example, Peru was the topic before (20) was uttered, or in which *Peru* is given prosodic prominence). Reduced acceptability stems from the fact that *Peru* occurs in a prepositional phrase modifying the head of the direct-object NP, a syntactic position that is not in itself sufficient to bring a referent into focus.

## 4.   What endows linguistic referents with the required level of salience?

In this section, we need a strict distinction between world-to-language and unshifted

metalinguistic anaphora. In the latter case, it may be assumed that the linguistic referent derives its salience from being *mentioned*. Take:

(21)  The term "berber," while still used by some, is problematic. **The term** is of Greek derivation, meaning "foreigner" or "non-Greek speaker."

(www.amazighworld.org/communication/who/abouttheportal.php)

The occurrence of *The term* in bold face is anaphoric: it is to be interpreted in terms of the previous occurrence of (*The term*) *"berber"*. Here anaphora resolution seems a straightforward affair since the antecedent has metalinguistic reference and already mentions the referent of the anaphor.

World-to-language anaphors are a different kettle of fish, precisely because the domain of reference shifts between the antecedent and the anaphor. In (7) to (12), the metalinguistic anaphors have antecedents that are not mentioned but *in ordinary use*. How come they are available at all for anaphora?

The question becomes all the more pressing when one realises that even unstressed personal pronouns can occur as world-to-language anaphors. As Ariel states, the speaker who uses an anaphoric personal pronoun assumes that the mental representation to be retrieved is *highly accessible* (1988: 77, 1991: 449). We saw above that Gundel et al. make a similar claim. In a recent contribution (2003: 284), Gundel, with other collaborators, goes so far as to say that "[…] unstressed personal pronouns, including *it*, are appropriately used only when the referent can be assumed to be in focus for the addressee *prior to processing of the referring form*" (italics mine).[6]

---

[6] A very similar position is expressed in Cornish (1999: 6).

If these claims are correct, the prediction is that any unstressed personal pronoun occurring as a successfully interpretable world-to-language anaphor signals that its referent (some linguistic entity) *was already in focus*. This must be the case in sentence (12), repeated below, as well as in (22):

(12)   After several hours of bouncing from one bureaucrat (notice **it**'s a French word) to another I was allowed into the hallowed chambers.

(22)   This grinder uses a dead man switch to activate the grinder. **It**'s what we call it because as soon as you release pressure the switch turns off. (www.wholelattelove.com/buyingguide.cfm?buyingguideID=4)

I gather that these utterances are not especially difficult to interpret, that anaphora resolution takes place quite smoothly and that ordinary hearers/readers would not notice anything special going on. None the less, it is a fact that, in (12) and (22), *it* heteronymously refers to a linguistic entity about which nothing has been *said* in the previous context. Nor, for that matter, have the expressions *dead man switch* and *bureaucrat* been brought into special prominence, one way or another (e.g. by scare quotes in writing, by prosody in speech).

It may seem then that all it takes for *dead man switch* and *bureaucrat* to be in focus (i.e. available for reference by means of *it*) is for these words to have been uttered, i.e. to have been made perceptible in the preceding co-text. This in itself is an intriguing proposal. But there is more: the entities that are made manifest in the co-text are not the actual referents of the two occurrences of *it*: these anaphors refer to expression-*types*, not to the particular tokens occurring in an utterance of (12) or (22). It is to these expressions as types (here, as lexical items) that the predicates *call* and *French word* apply.

To sum up, we are faced with two main questions. First, there is a question that concerns all instances of world-to-language anaphora: How come those linguistic entities are available for reference at all? This issue needs to be addressed regardless of which anaphoric expression is used. Second, we need to ask how those linguistic entities can sometimes be so highly activated that they are felicitously picked up by cognitively demanding anaphors, notably by unstressed personal pronouns. While trying to answer these questions, especially the second one, I will have to say something about how salience is built in the context of a discourse.

*4.1 What enables those linguistic entities to function as referents?*

A plausible answer to that question can be found in a 1998 paper on quotation by Paul Saka. There, the author puts forward a 'multiple ostension' hypothesis according to which any expression used in a spoken utterance 'directly ostends' a phonetic token (say /dʒɒn/) and 'de d ferringly ostends' a number of other features or entities: when I utter the word *John*, I explicitly refer to John (that is part of what I say), but I also 'activate' (via the phonetic token) the corresponding form type, the lexeme <*John*, /dʒɒn/, proper noun, JOHN>, the concept JOHN (see Saka 1998: 126).

I think Saka's hypothesis is sound. Therefore, I assume that the result of uttering any expression is that various linguistic aspects of it become objects of the context of utterance (of the universe of discourse) endowed with at least a minimal degree of salience. These objects then are potential targets for subsequent referential expressions because the hearer/reader is at least minimally alert to them as the discourse unfolds.

*4.2 What brings linguistic entities into focus?*

As far as I can see, in world-to-language anaphora, the linguistic referent (an expression-type, a lexeme, a meaning) possesses no other salience than that which it acquires indirectly from the uttering of an associated expression-token. That is all the salience it has *prior* to the moment when a subsequent world-to-language anaphor is processed.

This, at least, holds for cases where no specific intonation pattern or typographical markers have been used. I return to these highlighting devices shortly, but for the moment I rest satisfied with the claim that the antecedents in (12) and (22) can all be uttered neutrally and still be recoverable as antecedents of metalinguistic anaphors. I assume that most of my readers had no trouble arriving at the correct interpretation for these utterances, even though the antecedents were not highlighted in any way.

Now, if Saka's story accounts for all the salience of the relevant linguistic referents, then it must also singlehandedly account for the *level* of activation of these referents. Thus, with respect to cases like (12) and (22), where in-focus status is required, the theory in Gundel et al. (2003) predicts that multiple ostension (or some similar story) adequately explains how that demanding cognitive status is achieved here. Yet, this prediction cannot be right, for several reasons. First, we saw earlier (examples 19, 20) that the explicit mention of a referent may not suffice to bring it into focus. One has reason to doubt that mere deferred ostension will succeed where even explicit mention may fail. Second, it is not sensible to assume that the exact same event (some words have been uttered) can lead to such different outcomes as the following: making a referent 'uniquely identifiable' (when the anaphor is a definite NP), placing it in short-term memory (demonstrative pronouns), or bringing it into focus (unstressed pronouns). Third, and probably worst of all, Gundel et al's prediction would seem to have an absurd consequence: since all (the entities associated with) the tokens uttered prior to the anaphor in (12) or (22) are ostended to the same extent (as a direct result of being

uttered), this must mean that they are *all* in focus (for at least some period of time). Clearly, this is an undesirable conclusion.[7]

So Saka's story cannot be all there is to the salience of *bureaucrat* in (12) and *dead man switch* in (22). But since the multiple ostension hypothesis is all we have to explain what happens *to the left* of the anaphor, we must look for an explanation *to its right*. This explanation is disappointingly straightforward: it lies in the combination of the anaphor with a metalinguistic predicate. Try replacing that predicate with a neutral one that applies equally well to linguistic and non-linguistic objects, and the anaphor can now hardly be interpreted as a case of heteronymous mention:

(23)  I was attacked by Zonkins.

      a.  **It**'s a strange word.

      b.  How do you spell **it**?

      c.  **It**'s strange, isn't it?

      d.  How do you like **it**?

(24)  This grinder uses a dead man switch to activate the grinder. **It**'s funny, isn't it?

(25)  After several hours of bouncing from one bureaucrat (notice **it**'s French) to another I was allowed into the hallowed chambers.

Whereas, in (23a-b), *it* is naturally construed as referring to the name *Zonkins*, it can hardly be understood to do so in (23c-d). Yet, the predicates *strange* and *like* do not bar the metalinguistic interpretation, since they are neutral with respect to a world-oriented or language-oriented reading. In (24), *it* is now very unlikely to be interpreted as referring to the phrase *dead man switch*. (25), it appears, is different. The preferred interpretation for *it* seems

---

[7] As a reviewer notes, the three reasons given may be no more than different ways of identifying a single cause: the fact that the degree of salience of a referent can still change *after* the occurrence of the expression that makes that referent available. See below.

to be heteronymous mention. Some members of the audience at the 2005 Dortmund conference on Constraints in Discourse suggested that *French* might be more 'language-biased' than the other alleged neutral predicates. That may be so. But note that, in the absence of additional contextual information, an utterance of *Champagne is French* is ambiguous between a world-oriented and a language-oriented reading (with a preference for the former). This casts doubt on that explanation. My feeling is rather that the heteronymous reading is facilitated by the occurrence of *it* in the parenthetical clause *notice it's French*, which intrudes at an arbitrary spot in the sentence structure: in this case, it splits a place adjunct into two fragments, thereby disrupting the humdrum flow of information and introducing a new topic. This intrusion may draw attention to the words to its immediate left, one of which is indeed French, increasing the likelihood that *it* is going to be rightly interpreted as shifting reference from the world to language.[8]

The facts about metalinguistic predicates in (23)-(25) suggest that *givenness*, in the strict sense, is inadequate to explain world-to-language anaphors. The same applies to accessibility understood strictly as *prior accessibility*. Most likely, in (23)-(25), the referent is not in focus *before* the occurrence of anaphoric *it*. Here, therefore, is my preferred account of what goes on:

the anaphor, being an anaphor, triggers a presupposition that its referent has already been introduced into the discourse. The presupposition initiates a search for this referent. Processing of the metalinguistic predicate then helps direct the addressee's attention to linguistic entities. That way, a linguistic referent is *brought into focus*.

---

[8] At this stage, these are only sketchy and speculative remarks that need substantiating. But a proper study of the role of parentheticals would take me too far afield.

## 5. Role of the predicate elsewhere

The previous discussion has shown that givenness (or accessibility) cannot in and of itself explain how pronominal metalinguistic anaphors select their referent. I have suggested that, in order to account for the pronoun's ability to pick out a linguistic referent, it is necessary to look to the right and consider the metalinguistic predicate. My main aim has been to show the inadequacy of givenness/accessibility as a unique explanatory factor. In section 7, I outline a general framework for anaphora resolution that takes account of the contributions of givenness/accessibility and of the predicate with which the anaphor combines, but also allows for other factors. In the meantime, however, I wish to look further into the role of the metalinguistic predicate.

It is striking that neither of the theories I have appealed to has anything to say about the lexical contribution of the predicate governing the anaphor, or, more generally, about the contribution of the co-text to the right of the anaphor. However, the they might be justified in neglecting the right-hand co-text provided, simply, that the predicate barely played a role in the sorts of anaphora that they were considering (i.e. outside world-to-language anaphora).

But that is just not true. In a study of referents introduced by clauses rather than NPs, Gundel et al. do treat the predicate as a determining factor. Still, they do so only implicitly:

(26) a. John insulted the ambassador. **It** happened at noon.

b. John insulted the ambassador. ??**It** was intolerable. (cf. Gundel et al. 2003: 285)

The claim here is that some clausally introduced entities license subsequent reference by means of *it* while others do not. Thus, if the entity in question is an event, *it* should be usable, whereas if the entity is a situation (or 'fact', in some terminologies), *it* should be dispreferred and, for instance, replaced by *that*. Variation in the acceptability of anaphoric *it* means that

described events are more highly activated than described situations. In the case at hand, if *John insulted the ambassador* introduces a situation rather than an event, then it is more difficult to refer to this situation by means of the personal pronoun *it*.

The problem is that, by the time *John insulted the ambassador* has been processed, it is still an open question whether that sentence denotes an event or a situation. This, by the same token, also means that the degree of salience of the described entity is not fixed yet. Not until the predicate of the following sentence has been processed can the first sentence be said to denote an event or a situation. Thus, (26a) is construed as an event because of the subsequent occurrence of *happened*. Similarly, interpretation of *intolerable* in (26b) turns John's insulting the ambassador into a situation. In 'salience-talk', we are induced to say that the degree of salience of the entity introduced by *John insulted the ambassador* cannot be determined until *happened* or *intolerable* have been processed. In particular, it appears that the in-focus status of the event of John's insulting the ambassador depends in part on the occurrence of the verb *happened* in the right-hand co-text.

These examples show that the influence of predicates on the salience of their anaphoric arguments extends beyond world-to-language anaphora. But, as I said, it is striking that Gundel et al., though their analyses implicitly acknowledge this, do not point out the role of the predicates to the right of anaphors. Neither, for that matter, does Ariel in the papers mentioned earlier.

## 6. Robustness of the proposed account

One way of testing the role played by the metalinguistic predicate is to check what happens when the salience of the linguistic referents is enhanced by means other than a metalinguistic predicate.

I hinted above that antecedents could be highlighted in such a way as to attract hearers' attention to them *qua* expressions. However, the following set of examples suggests that such highlighting is not enough to endow a linguistic referent with the required degree of salience:

(27) I was attacked by 'Zonkins'.

　　a. **It**'s strange, isn't it?

　　b. How do you like **it**?

(28) This grinder uses a *dead man switch* to activate the grinder. **It**'s funny, isn't it?

Even if the scare quotes or italics are realised prosodically as very emphatic stress, it is unclear that construal of *it* as a heteronym is facilitated.

Some further evidence for the role of the metalinguistic predicate comes from 'the other half' of metalinguistic anaphora, i.e. the 'unshifted' cases. Consider what happens to (29) if the metalinguistic predicate is replaced with a neutral one:

(29) "We're in a forest, with spiders and who knows what other yuckies…" **this** she said as she wrinkled her nose. (www.ofelvesandmen.com/Stories/D/Di-and-LK/RoadTrip2-3.htm)

(30) "We're in a forest, with spiders and who knows what other yuckies…"

　　a) **this** she hated, so she wrinkled her nose.

　　b) **that** was pretty bad.[9]

In (29) and (30), the linguistic referent – the female character's utterance – is mentioned by means of the direct quotation. Yet, once again, it appears that without a metalinguistic

---

[9] The predicates *hate* and *be pretty bad* are neutral in the sense that they can apply to linguistic entities just as well as they do to extralinguistic ones. (cf. the predicates used in (23)-(25))

predicate the heteronymous interpretation of the anaphor is hardly available. Notice moreover that we are dealing with demonstrative pronouns, i.e. *less* cognitively demanding forms than personal pronouns, requiring only that the referent be activated (30a) or familiar (30b), according to Gundel et al. (1993). Yet, even though the referent is mentioned explicitly, and the anaphoric forms demand less salience, the intended interpretation cannot be readily accessed.

Things, however, are not as clear-cut as with world-to-language anaphors. Take example (21) again and consider what happens if a pronoun is substituted for the definite NP and a neutral predicate for the metalinguistic one:

(21)  The term "berber," while still used by some, is problematic. **The term** is of Greek derivation, meaning "foreigner" or "non-Greek speaker."

(31)  The term "berber," while still used by some, is problematic. **It** is strange/We don't like **it**.

Here, it seems that something like default assumptions about topic continuity ensure that *it* is going to be interpreted correctly as a heteronym. Alternatively, the presence of the introductory predicate *the term* in the antecedent sentence may affect the availability of the heteronymous interpretation of *it* in (31). Note also that if *this* or *that* were used instead of *it*, the metalinguistic interpretation would not be secured. That can probably be explained by (something like) Grice's maxim of quantity: if a more cognitively demanding form can be used, then a less demanding one cannot be used without a special reason: a plausible candidate here is topic change. Clearly, further work needs to be done on the factors that affect the availability of a metalinguistic interpretation for an unshifted anaphor. But this goes beyond the scope of the present study.

18

## 7.  Plausible explanations

In this section, I will extend the argument that prior accessibility or salience cannot be the sole explanatory factor for metalinguistic anaphora resolution. In 7.1 and 7.2 I discuss two more factors that may play a part, the lexical meaning of the antecedent, and the possibility that some linguistic referents are inherently more salient than others. The conclusion I am led to is that all these factors need to be brought together into a general framework, something along the lines of Recanati (2004). This is not incompatible with Gundel et al. and Ariel. Instead, it is to be seen as an extension of their theories, one that further develops the notion that anaphora resolution is primarily a pragmatic affair.

The story so far offers an explanation for why *a* referent can be made salient enough. But it does not say how a *particular*, non-random referent is picked out. Access to the referent is provided via identification of the right antecedent, but how is this antecedent selected?

### 7.1 The lexical meaning of the antecedent

First, there are instances in which the lexical meanings of both the antecedent and the metalinguistic predicate play a major part. Take (8) again:

(8)  A: I think of him as a family man.

   B: Funny, I've always considered **that phrase** an oxymoron.

Since *an oxymoron* is predicated of *that phrase*, the latter NP must refer to something that can be judged to be both a phrase and an oxymoron. Therefore, the antecedent must (i) be a complex phrase and (ii) denote two properties that can, in a given context, be interpreted as incompatible. Only an antecedent whose lexical meaning is consistent with that of the predicate will be selected. If there is no more than one such antecedent – here, there is only

*family man* – then anaphora resolution is straightforward. In any case, all words or phrases whose meaning did not match that of the predicate would already have been eliminated.

However, the match between lexical meanings does not systematically play a part in anaphora resolution. Thus, in (9), this explanation is entirely unavailable, because the metalinguistic predicate, *title of an Elvis Costello album*, does not favour any particular antecedent: only a fully context-based, pragmatic, explanation can do the job. The fact that the antecedent of *that* is *get happy*, rather than the italicised *get* or *happy* is entirely a matter of world knowledge. Anyone unfamiliar with the Elvis Costello catalogue would be unable to pick out the right referent.

*7.2 Inherent salience of certain words or phrases*

In discussion, there have been suggestions that some antecedents inherently stand out. Take rare words, like *swimmingly* in (10), or long ones like *repatriated* in (7). I do not deny that rare or unusually long words tend to attract the addressee's attention. Some psycholinguists at least have argued that different words have different activation thresholds (cf. 'the frequency effect'; see Garman 1991: 279-81 or Harley 1995: 71-73). All the same, words are not *inherently* salient, and the 'lexical oddity' thesis therefore falls short of being a satisfactory explanation. High salience is not part and parcel of a lexeme, it is context-sensitive. Two quick examples: specialised, technical words tend to be rare words. This means that they are likely to attract attention when used in an everyday conversational context. Yet, when used in a specialised one, these terms will not be particularly salient. Or consider the fact that even the most ordinary words can become highly salient, provided they are uttered in the right context. Thus the word *baby* heard by a man who is about to become a father.

*7.3 A general pragmatic framework*

The factors that have so far been shown to play *some* part in the resolution of metalinguistic anaphora are:

- the accessibility of the linguistic referents

- a constraint on the distance between antecedent and anaphor (end of section 2)

- the match between the lexical meaning of the antecedent and the predicate with which the anaphor combines (or which is included in the anaphor, in the case of definite and demonstrative NPs)

- unequal threshold levels of contextual salience

All these factors are potentially relevant, and they need to be integrated into a general framework. This framework must be essentially pragmatic, as it appears that no single parameter can account for metalinguistic anaphora resolution: although some factors (distance, for instance) probably play a role in all cases of resolution, it is usually a *cluster of factors* that combine to help identify the right antecedent and select the corresponding referent. And which factors are relevant appears to be context-dependent.

Let us once again focus on pronominal world-to-language anaphors. What happens when a hearer/reader comes across the anaphors in utterances like (9)-(12) or (22)? It is very likely that the hearer/reader will not immediately ascribe to these a linguistic referent. On the contrary, he is likely at first to have in mind an ordinary extralinguistic referent. That is because nothing in the co-text *up to that point* has particularly activated any linguistic entities. Thus, in (9), he would probably expect *that* to refer to the event of getting happy (cf. *So what happens if you* get *happy, and yes I know that's not likely to happen soon*). Similarly,

in (10), the default interpretation for *which* is likely to be the fact that everything went swimmingly (cf. *Yes, everything went swimmingly, which was rather a surprise*).[10]

At this stage, two stories are, I think, conceivable. On the first, a costly 'repair' procedure is assumed to take place: the pronominal anaphor's default interpretation enters into semantic composition with the interpretation of the metalinguistic predicate. This yields an absurd interpretation for the whole clause (something to the effect that "the event of getting happy is the title of an Elvis Costello album" or that "the fact that everything went swimmingly is a peculiar adverb"). The absurdity of this result then induces the hearer/reader to backtrack and re-interpret the pronoun as having metalinguistic reference. Only on this second attempt does the hearer/reader come up with an acceptable interpretation for the whole clause.

The second account does not assume such a cognitively costly procedure. Here, the idea is that the extralinguistic interpretation of the anaphor does not undergo composition: it is 'entertained' only until the hearer/reader encounters the metalinguistic predicate, at which stage it is superseded by the metalinguistic reading. The assumption is that this reading was already activated 'by association': as soon as words are uttered, they (or their related types) are endowed with a minimal degree of activation – this is the gist of Saka's hypothesis. Therefore, what happens is that the linguistic referent enters directly into semantic composition with the metalinguistic predicate: there is no cancellation of an initial, primary interpretation *of the whole clause*. The first time the whole clause is interpreted, it is *qua* referring to a linguistic entity.[11]

The second account is essentially that given for figurative language by François Recanati in his *Literal Meaning* (2004: 27-30). Recanati rejects a dominant 'Gricean' model according

---

[10] That there has been no particular previous activation of a linguistic-entity-as-referent is patently true, of course, when the anaphor *precedes* its antecedent, as in example (11).

[11] This paragraph is a follow-up on a comment made by Rachel Giora (p.c.). Giora suggested that my account would predict garden-path effects in cases of world-to-language anaphora, adding that such effects could be tested empirically. I have not had the opportunity to do this, but I think that my account leaves open the possibility that world-to-language anaphors are not garden-path sentences, that some adjustment of anaphor and predicate takes place 'locally' before the whole sentence is processed.

to which an absurd literal interpretation for the whole utterance is processed first, then discarded because it fails to comply with some conversational maxim, thus triggering a repair procedure by means of conversational implicatures. Recanati outlines a different story, one that makes allowances for 'primary pragmatic processes', i.e. associative processes by which the meaning of local (subclausal) constituents is adjusted *before* undergoing semantic composition.

Interestingly, Recanati then shows how this model can be extended to the selection of a referent for an anaphoric pronoun (2004: 31-32) and explicitly underlines the role that processing of the predicate plays in affecting the accessibility of the candidate referents (2004: 31, 33). My account of world-to-language anaphora is in the same spirit. In particular, I take it that the processes involved in interpreting pronominal anaphors are cognitively the same as those described by Recanati: they are associative rather than properly inferential; they apply locally rather than to the whole utterance. However, the analogy is perhaps only partial. The processes described by Recanati are not 'pre-semantic'; they are not part of the processes that lead to determining 'which sentence was uttered'. By contrast, it is reasonable to hold that choosing between an ordinary and a metalinguistic reading of a pronoun may be a facet of disambiguation. Assume I say *I love Chicago!*. I may mean the proper name to refer to a variety of objects, notably a city, a word (e.g. if I like the pronunciation), but also a musical, a band, etc. It seems to me at least that the sentence about the city is a different sentence from the one that refers to the word. If that is correct, selection between an object-level and a meta-level referent for *Chicago* is part of disambiguation. And if, as seems sensible, one extends this analysis to all cases in which a linguistic referent competes with an extralinguistic one, there is a genuine difference between my account and Recanati's.

## 8. Binding or accommodation?

I believe the story told so far can be reformulated in terms of dynamic models of discourse such as DRT. The cognitive constraints associated with the use of anaphors can be cashed out in terms of presupposition. On this view, use of anaphoric *this*, for example, triggers the presupposition that, say, the referent must be present in short-term memory. And, even if it turned out that metalinguistic anaphors are really demonstrative deictics, this could be captured in presuppositional terms as well.[12] Use of indexical *this* would trigger the presupposition that the speaker is demonstrating (e.g. pointing at) the relevant referent. On theories like those set out in Geurts (1998) and Beaver & Zeevat (2007) the presuppositions triggered by anaphors and demonstratives[13] can either be 'bound' to an antecedent (by which the relevant referent has been previously introduced into the discourse and is therefore part of the 'common ground'), or 'accommodated' (see below). If neither happens, presupposition failure occurs and communication is unsuccessful.

In my examples, there can be no talk of presupposition failure: world-to-language shifts do not usually cause breakdowns in communication. Therefore, the presuppositions triggered by shifters must be accounted for in terms of binding or accommodation. Binding, however, is to be ruled out: the linguistic referent introduced by a shifter is new: it is not part of the common ground by the time the shifter is processed. Therefore, the prediction is that shifter-triggered presuppositions are accommodated. Accommodation comes in when a

---

[12] I have talked of the expressions shifting reference from the world to language as *anaphors*, thus adopting John Ross's terminology in a 1970 squib about *metalinguistic anaphora*, the first discussion of the topic known to me. Many writers, however, talk of *discourse* or *textual deixis* instead (e.g. Lyons 1977: 667; Webber 1988: 116; Huddleston & Pullum 2002: 1460-61; Levinson 2004: 103, 108). I readily admit that there is a strong kinship between my data and deixis, especially the use of demonstrative NPs. Still, evidence seems to go both ways. (For some discussion, see De Brabanter 2004.) I cannot go into this issue here, but I rest content with the idea that there may be more similarities than differences between the two phenomena, a view for which Recanati (2005) makes a convincing case.

[13] In the rest of this section, I shall use the term *shifter* as a cover term (for expressions that host a shift in the universe of discourse), so as to leave the door open to a deixis-based account too.

24

presupposition that cannot be bound is nevertheless interpretable by the hearer. Let us take a simple example. A stranger comes up to me in the street and says:

    (32)  I've lost my dog! Can you help me?

Use of *my* triggers the presupposition that the man has a dog. But this cannot be bound to any information already in the common ground: the man is a perfect stranger to me. Yet, I have no trouble *adding* this information to the common ground. I 'accommodate' it, and no communication failure ensues. Now consider this other example: I hear a tremendously loud bang. A stranger turns to me and exclaims:

    (33)  That sounded like an explosion.

Her utterance is not accompanied by any demonstration. Although Beaver & Zeevat suggest that "the use of the demonstrative presupposes the demonstration by the speaker. If no such demonstration occurs, infelicity results: the reader cannot simply incorporate a referent" (2007: 535), I none the less assume that they would agree that accommodation still takes place provided a referent is salient enough in the situation of utterance. That is the case here given the loudness of the bang.[14]

It is not clear that a similar story can be told about the presuppositions triggered by shifters. As pointed out in 4.2, there is no notable event foregrounding the intended linguistic referent, nothing like the loud bang in (33). This means that the Beaver/Zeevat account

---

[14] Ginzburg 2001: 21 has a similar example (his (24a)):
        [Context: a shot is heard, followed by a woman's scream:] A: Oh boy, she sounds scared.
Ginzburg's comment is that "deixis makes anaphors felicitous without an overt antecedent" (2001: 21). But this is deixis *without a demonstration*.

predicts presupposition failure in most cases of world-to-language shift, a prediction that cannot be correct, because hearers generally have no trouble identifying the right referent.

At this stage, the only way for a DRT-based account to address this issue is to introduce an extra layer of representation for speech events themselves,[15] with special constraints designed to capture the short span of time during which uttered linguistic forms remain accessible. A step in this direction is taken by Corblin & Laborde (2001), in their study of the French counterparts of *the former* and *the latter*. Another such attempt, though not within DRT, is Ginzburg & Cooper's (2004) HPSG formalisation of 'clarification ellipsis'.[16] Ginzburg & Cooper focus on cases like *A: Did Bo finagle a raise? B: finagle?*, where B's elliptical question is susceptible of either a 'clausal reading' (= *Are you asking if Bo finagled a raise (of all actions)?*) or a 'constituent reading' (= *What does it mean to finagle?*). They show that any model capable of accounting for examples of this kind must include several ingredients, two of which I will single out as they are directly relevant to an account of world-to-language anaphora. One ingredient is 'utterance reference'. The idea is that clarification ellipsis – especially in its 'constituent reading' (which is often metalinguistic) – usually involves reference to utterance *tokens* (or 'speech events'), so that the syntactic-semantic representation of clarification ellipsis "must include references to (previously occurring) utterance events" (2004: 306). Note that this is exactly what is needed in the case of world-to-language anaphors. The second constraint is termed 'sub-utterance accessibility', and it originates in the observation that "any semantically meaningful sub-utterance can be clarified using [clarification ellipsis] under conditions of phonological or partial syntactic parallelism" (2004: 306). Once again, note that a similar constraint weighs upon any model of world-to-language anaphora: the antecedent of a shifted anaphor can be just any word or phrase in the close co-text. As a matter of fact, the constraint is stronger in the case of world-

---

[15] Note that the difficulties highlighted in this section do not crop up with unshifted anaphors: these can be directly bound to an antecedent that refers to linguistic material.

[16] I am grateful to an anonymous referee for providing this reference.

to-language anaphora, because the antecedent need not even be a "semantically meaningful sub-utterance" (cf. examples (7) and (9), where the antecedent is not a referential expression). The provision of a formalised account of metalinguistic anaphora that includes utterance reference and sub-utterance accessibility is, at this stage of my research, no more than a promissory note.

## 9. Conclusion

The present study is in line with the earlier proposals of Gundel et al. (1993) and Ariel (1988, 1991) in that it corroborates the idea that a cognitive principle underlies the choice of anaphoric expressions. However, though an important parameter, prior accessibility is not enough to explain what goes on in metalinguistic anaphora. Elsewhere, Wilson & Matsui (1998) have also shown that it is inadequate to account for other types of anaphora resolution.

In the end, the main dividend of this study might be this: the very unobtrusiveness of world-to-language anaphors – the fact that their resolution does not pose special problems – throws light on what participants keep track of as a discourse (e.g. a conversation) proceeds. Psycholinguists have shown that discourse participants keep some record of which words were uttered.[17] But here, we have *direct linguistic* rather than experimental psychological evidence of that assumption. Furthermore, speakers must to some extent be aware of what can be assumed to be in the minds of other participants. Otherwise, we should expect anaphora resolution to be a haphazard process. But it is not. The fact that it is often successful suggests that speakers assume other participants to temporarily store *uttered linguistic forms* over and above extralinguistic referents. It is likely that participants in a discourse

---

[17] See e.g. Levelt & Kelter (1982) or Levelt (1989). The idea is that, when planning a new utterance, speakers rely on the information stored in their 'discourse record' (Levelt 1989: 111) in order to make their new contribution coherent with previous moves. Levelt & Kelter adduce experimental evidence that lexical information is among those aspects of a discourse that are stored in short-term memory.

temporarily open mental files for linguistic entities too, for otherwise these could not serve as anchors for anaphoric expressions.

Although this idea may appear quite commonsensical, it clashes with some views found in the literature. Thus, in a discussion of anaphoric pronouns, Mark Sainsbury writes:

> In the process of interpretation, we expect understanders to carry previous interpretations forward and for them thus to be available for solving new problems of interpretation. *It would be quite another thing to expect them to carry forward memories of precise linguistic forms.* It is a familiar phenomenon that among people who use two languages interchangeably in their conversations, it often happens that one remembers what the other said, but not in which language she said it. *Interpretation is remembered, but not linguistic form.* […] In interpreting the first part of, say, (1) ("A mosquito is buzzing about our room") we come to know that what the speaker has said is true iff there is at least one mosquito which is buzzing around our room, and *we throw away all other information about the utterance, including the words in which it was couched.* (2002: 57; emphasis mine)

As far as I can see, the present study invalidates views of this sort. Sainsbury is certainly right in assuming that propositional contents are usually remembered much longer than the linguistic material used to convey them. But he fails to allow for the difference between effects on long-term and on short-term memory. What Saka calls *multiple ostension* has a short-lived impact. The various linguistic aspects associated to an uttered token are only kept in mind for a while; long enough, however, to enable the hearer to interpret successfully a subsequent shifted anaphor, provided this anaphor occurs while the ostended items are still stored in short-term memory.

References

Ariel, M. (1988). "Referring and accessibility". *Journal of Linguistics*, 24, 65-87.

Ariel, M. (1991). "The function of accessibility in a theory of grammar". *Journal of Pragmatics*, 16, 443-463.

Beaver, D. & Zeevat, H. (2007). "Accommodation". In Ramchand, G. & C. Reiss (eds.), *Oxford Handbook of Linguistic Interfaces*. Oxford University Press, pp. 503-539.

Borthen, K., Fretheim, T. & Gundel, J. (1997). "What brings a higher-order entity into focus of attention? Sentential pronouns in English and Norwegian". In Mitkov, R. & B. Boguraev (eds), *Operational Factors in Practical, Robust Anaphora Resolution for Unrestricted Texts*. Madrid, pp. 88-93.

Clark, H. H. (1977). "Bridging". In Johnson-Laird, P. N. & P. C. Watson (eds), *Thinking: Readings in Cognitive Science*. Cambridge, Cambridge University Press, pp. 411-420.

Corblin, F. & Laborde, M-Ch. (2001). "Anaphore nominale et référence mentionnelle: *le premier, le second, l'un et l'autre*". In W. De Mulder et al. (eds), *Anaphores pronominales et nominales. Etudes pragma-sémantiques*. Rodopi, 99-121.

Cornish, F. (1999). *Anaphora, Discourse, and Understanding. Evidence from English and French*. Oxford, Clarendon Press.

De Brabanter, P (2004). "'World-to-Language' Shifts between an Antecedent and its Pro-Form". In Branco, A., McEnery, T. & R. Mitkov (eds.), *Proceedings of the 5th Discourse Anaphora and Anaphor Resolution Colloquium (DAARC 2004)*. Lisbon, Edições Colibri, pp. 51-55.

Garman, M. (1990). *Psycholinguistics*. Cambridge University Press.

Geurts, B. (1998). "Presuppositions and anaphors in attitude contexts". *Linguistics and Philosophy*, 21, 545-601.

Ginzburg, J. (2001). "Clarification ellipsis and nominal anaphora". In Bunt, H., Muskens, R. & E. Thijsse (eds.), *Computing Meaning: Volume 2*. N° 77 in Studies in Linguistics and Philosophy, Dordrecht, Kluwer.

Ginzburg, J. & Cooper, R. (2004). "Clarification, ellipsis, and the nature of contextual updates in dialogue". *Linguistics and Philosophy*, 27, 297-365.

Gundel, J., Hedberg, N. & Zacharski, R. (1993). "Cognitive status and the form of referring expressions in discourse". *Language*, 69, 274-307.

Gundel, J., Hegarty, M. & Borthen, K. (2003). "Cognitive status, information structure, and pronominal reference to clausally introduced entities". *Journal of Logic, Language and Information*, 12, 281-299.

Harley, T. A. (1995). *The Psychology of Language. From Data to Theory*. Hove, Erlbaum (UK) Taylor & Francis.

Huddleston, R. & Pullum, G. K. (2002). *The Cambridge Grammar of the English Language*. Cambridge University Press.

Kleiber, G. (1999). "Associative anaphora and part–whole relationship: The condition of alienation and the principle of ontological congruence". *Journal of Pragmatics*, 31, 339-362.

Levelt , W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, Mass., MIT Press, Bradford Books.

Levelt, W. J. M. & Kelter, S. (1982). "Surface form and memory in question answering". *Cognitive Psychology*, 14, 78-106.

Levinson, S. (2004). "Deixis". In Horn, L. & G. Ward (eds.), *The Handbook of Pragmatics*. Oxford, Blackwell, pp. 97-121.

Lyons, J. (1977). *Semantics*, 2 vol. Cambridge, etc., Cambridge University Press.

Quine, W.V.O (1940). *Mathematical Logic*, Cambridge, Mass., Harvard University Press.

Recanati, F. (2000). *Oratio Obliqua, Oratio Recta: An Essay on Metarepresentation*. Cambridge, Mass., MIT Press, Bradford Books.

Recanati, F. (2004). *Literal Meaning*. Cambridge, Cambridge University Press.

Recanati, F. (2005). "Deixis and anaphora". In Szabo, Z. (ed), *Semantics vs. Pragmatics*. Oxford, Oxford University Press, pp. 286-316.

Ross, J. R. (1970). "Metalinguistic Anaphora". *Linguistic Inquiry*, 1, 273.

Sainsbury, R. M. (2002). "Reference and anaphora". *Philosophical Perspectives* 16, 43-71.

Saka, P. (1998). "Quotation and the use-mention distinction". *Mind*, 107, 113-135.

Webber, B.L. (1988). "Discourse deixis: reference to discourse segments". *Proceedings of 26th Annual Meeting of the Association for Computational Linguistics*, pp. 113-122.

Wilson, D. & Matsui, T. (1998). "Recent approaches to bridging: Truth, coherence and relevance".

*UCL Working Papers in Linguistics*, 10, 173-200.