

If I were you, I wouldn't trust myself

Anne Reboul

► **To cite this version:**

Anne Reboul. If I were you, I wouldn't trust myself: indexicals, ambiguity and counterfactuals. Acts of the 2nd International Colloquium on Deixis "Time, space and identity", Mar 1996, pp.151-175. halshs-00003800

HAL Id: halshs-00003800

<https://halshs.archives-ouvertes.fr/halshs-00003800>

Submitted on 2 Feb 2005

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

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

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

If I were you, I wouldn't trust myself Indexicals, ambiguity and counterfactuals

Anne Reboul
CRIN-CNRS & INRIA-Lorraine

Je n'est pas toujours un autre...

Abstract: This paper aims to defend what could be called the current theory of indexicals, and more specifically the first and second person pronouns, against a peculiar type of counter-examples which might be invoked against it. Very roughly, the current theory of indexicals (based on Kaplan 1977) says that *I* or *you* are interpreted not through a conceptual "description" which would be lexically linked to them (i.e. would be their semantic content), such as *the speaker* or *the addressee* of the present utterance, but rather through a specific procedure, lexically linked to them, such as *look for the speaker* or *look for the addressee*. These procedures would be their (non-conceptual and non-descriptive) semantic content. This analysis means that any utterance where an indexical occurs and where it is ambiguous as to whom the indexical refers to may be used as a counter-example if certain conditions are realised: the analysis of indexicals being what it is, if the situation is clearly perceived (i.e. if it is clear who is the speaker and who is the addressee), it should not be possible for an indexical to be ambiguous. Thus, authentic examples of ambiguity in indexicals are *ipso facto* counter-examples to the current theory of indexicals.

I will first distinguish between mere misunderstandings and what appear to be authentic examples of ambiguity, such as the title of this paper (*If I were you, I wouldn't trust myself*). I will show that these examples mainly occur in the consequents of counterfactuals whose antecedents are identity statements (where indexicals occur) and show the proximity between these counterfactuals and counteridenticals. I then will show that the antecedents are used non-literally and that the "ambiguity" of the indexicals in the consequents comes from the difficulty of choosing which specific implication the speaker intended the hearer to recover. Thus, the so-called ambiguity of indexicals in the consequent of such counterfactuals reduces to a misunderstanding as to the description of the situation to which the procedure attached to the indexicals should be applied.

1. Introduction

In this paper, I will only treat first and second person pronouns¹, and I will be especially interested in the cases in which they are, or seem to be, ambiguous. I should make it clear right away that I am only interested in ambiguity, that is, I am not interested in misunderstandings due to a bad perception of the situation in which the communication occurs. Hence I will distinguish between the situations in which there might be ambiguity in the use of indexicals and those where there is only a misunderstanding. Then I will concentrate on ambiguity in the use of indexicals which, I think, mainly occurs in the consequent of some rather peculiar counterfactuals, like the one quoted in the title², and discuss this ambiguity relative to the most current analysis of personal indexicals, according to which personal indexicals are unambiguous and, in fact, are designators to which a referent is attributed through a procedure and not through a description of any kind. I will try to account for this ambiguity by pointing out that it is not a truly indexical ambiguity, but

¹ Below, I will occasionally say *indexicals* when I mean first and second person pronouns: this should not be understood as including other kinds of indexicals.

² I borrow this example from Gerard Sabbah (personal communication).

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

rather the result of a general ambiguity occurring in counterfactuals which have an identity statement as antecedent.

I will first outline the current analysis of indexicals. Then I will examine a few examples and show why some of them are not truly ambiguous but come from a “mistaken” use of indexicals³ and why some of them are ambiguous. I will show that this raises a difficulty for the current analysis of indexicals and try to explain why this difficulty is only apparent, using the possible worlds analysis of counterfactuals, the notion of necessity proposed by Kripke (1980) and the notion of non-literal use proposed by Sperber and Wilson (1986).

2. The current analysis of indexicals

Traditionally, it is considered that *I* may be paraphrased as *the producer of this utterance*, while *you* could be paraphrased as *the addressee of this utterance*. Quite a few years ago, Kaplan (1977, published 1989) pointed out that, despite its popularity, this analysis could not be right. He used in his argument the following example:

(1) I do not exist.

He pointed out that if the indexical had “the producer of this utterance” as its descriptive content, then (1) would be semantically equivalent to (1’):

(1’) The producer of this utterance does not exist.

This would imply that (1) is true if and only if the speaker of (1) does not exist and that would, in its turn, mean that (1) could never be true, hence it would be impossible that I do not exist. It should be remembered that, in possible worlds theory, impossibility is equivalent with falsity in all possible worlds.

Yet, as Kaplan pointed out, there is strictly nothing impossible about the non existence of this or that individual and, supposing that I uttered (1), a sentence such as (2), though it happens to be false in this world, could indeed quite well be true in another possible world or set of possible worlds:

(2) Anne Reboul does not exist.

Thus, the falsity of (1) (and (2)) is a purely contingent matter and certainly not a necessary falsehood: hence the traditional analysis cannot be right.

I will not get into the details of Kaplan’s analysis of indexicals. I will rather very quickly outline what I take to be the basis of most currently accepted analyses of indexicals. Indexicals on these views are considered as having a semantic content (in the sense that this is what is lexically attached to them), but this semantic content is not conceptual, it is rather something like a procedure, which outlines the instruction(s) enabling the hearer to determine

³ I will show this mainly to be the case in a man-machine dialogue where the difficulties in comprehension were the result of a bad understanding of how pronouns work and not a matter of true ambiguity in the sentences.

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

the referent of the indexical⁴. In the case of the first and the second person pronouns, the procedures would be, respectively: *look for the producer of this utterance* and *look for the addressee of this utterance*. What would get into the proposition expressed by the utterance would not be either the procedure or a paraphrase such as “the producer/addressee of this utterance” but the referent itself. It should be noticed that this makes provision for the fact that the same word (*I* or *you*) changes its referent with its different uses. It should also be clear that this analysis in no way entails⁵ that we have a representation of ourselves which would be implied in the process of production of any singular utterance in which *I* occurs.

This has an implication as far as ambiguity in indexicals is concerned: such ambiguity should not arise though misunderstandings might occur when the situation or its representation is not clear enough to allow the procedure to determine a referent with any degree of certainty (i.e. in the cases where it is not clear who is speaking or to whom the speaker is addressing his utterance). Thus, any true case of ambiguity should raise a difficulty for the current theory of indexicals.

3. A false case of ambiguity in indexicals

I will now describe or quote a few examples and analyse them to show when it can be thought that there is ambiguity in the indexicals and when it is a mere matter of misunderstanding (in the above sense). I will first begin by a case which occurred in man-machine dialogue⁶:

- (3) In this specific case, the man and the machine took it in turns to work. The man had to choose whether it was his turn or the machine's turn and, to do this, he had to push either a button labelled “your turn” or a button labelled “my turn”⁷. Supposing that the user wanted to work himself, he should then decide whether he should push the “my turn” or the “your turn” button. Here, despite every kind of arrangement (“my turn” corresponding to the man working and the machine waiting and “your turn” corresponding to the the machine working and the man waiting or the reverse), the users failed to grasp the way to use it and, in the end, it was decided to add to the label the names of the protagonists, that is, to desambiguate the labels.

I think that it is interesting to note that (3) has a similarity with the bottle and cake in *Alice in Wonderland*:

- (4) There seemed to be no use in waiting by the little door, so she went back to the table, half hoping she might find another key on it, or at any rates a book of rules for shutting people up like telescopes: this time she found

⁴ By the way, this is probably the current analysis of all indexicals, not only of first and second person pronouns.

⁵ Kaplan is (rightly) clear about this in his 1977/1989 paper.

⁶ This story was told by Pierre Dillenbourg (personal communication).

⁷ I should make it clear that the choice between the two was not the object of the system: it was merely a device to allow the user to make it clear to the machine which of the user or the machine should work.

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

a little bottle on it, (...) and round the neck of the bottle was a paper label, with the words 'DRINK ME' beautifully printed on it in large letters.

(L. Carroll, *Alice in Wonderland*, 29).

- (5) Soon her eyes fell on a little glass bow that was lying under the table: she opened it, and found in it a very small cake, on which the words 'EAT ME' were beautifully marked in currants.

(L. Carroll, *Alice in Wonderland*, 31).

It might seem, at first sight, that example (3) is very different from examples (4) and (5): this comes from the fact that the utterances *drink me* and *eat me* in (4) and (5) are orders, while the labels *my turn* and *your turn* in (3) are (apparently) merely assertions. Yet, there is absolutely no reason why this should make a difference to the indexicals used in the different labels and indeed the proximity between the three examples comes from the fact that, in all cases, objects are involved to which the indexicals refer: in (3), the indexicals might refer to the computer (and one of them must indeed refer to the computer), while in (4) the indexical refers to the content of the bottle and in (5) it refers to the cake. This rather strong similarity makes it highly difficult to understand *a priori* why examples (4) and (5) are not presented as (and indeed are not) ambiguous in *Alice in Wonderland*, while example (3) seems to raise an insolvable ambiguity, which led the designer of the system to modify the labels in such a way that the ambiguity disappeared. Thus, the question becomes: why should the labels be ambiguous in (3) when they obviously are not in (4) and (5)?

I think that the answer to this question is very simple: (3) is a case of what I called a misunderstanding, that is, it is a case where the situation is unclear and (and this is one of its interest) where it remains unclear. The situation is unclear at the beginning because the man has to decide who is speaking before he pushes any button in order to indicate that he wants to be working: if he decides that he is speaking, he should push the button labelled "my turn"; if he decides that the machine is speaking, he should push the button "your turn". In each case, his decision may be right or wrong, depending on whether the system designer has decided that the machine is the speaker or that the user is the speaker. Thus, it is not surprising that at his first use, the user may have difficulty in deciding which button he should push: yet it seems that, after the first trial, he should be able, depending on whether he has been successful or not, to determine who is *I* and who is *you*. However, from what the designer said, this was something which was obviously very difficult if not impossible.

Why should this be so? As everyone has pointed out, *I* and *you* do not denote the same person all the time, even (or rather especially) inside the same conversation: the speaker who said *I* becomes the hearer and is addressed as *you* by his hearer who now says *I*. Thus, it could be said that the difficulty above comes from the fact that the users of the system were required to treat indexicals as if they were proper names, that is as rigid designators on whose referent no change of situation could bear. This goes so evidently against the common use and understanding of indexicals that it should not be surprising that no way out of the confusion could be found. Changing the labels could not help as it did nothing to solve the problem of the essential flexibility of

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

indexicals and as there is no “natural attribution” of a permanent referent in such a situation.

Thus, (3) is not an example of ambiguity in indexicals: it is a mere misunderstanding, arising out of a bad design, itself due to a mistaken analysis of indexicals and their interpretation. This does not, however, mean that no examples of true ambiguity in indexicals can be found.

4. Examples of ambiguity in indexicals

I will begin with two examples:

- (6) If I were you, I wouldn't trust myself/me.
- (7) 'I shouldn't, if I were you', said George benignly.
'On the contrary,' said Willy the Twig, 'being you, of course you wouldn't, but if you were me you certainly would'.
(E. Peters, *Rainbow's end*, 19).

Example (6) may well be the epitome of what can be meant by ambiguity in indexicals. Let us suppose that the speaker of this utterance is Fred, while his hearer is Larry: in the antecedent of the counterfactual, *I* designate Fred and *you* designate Larry. There does not seem to be any kind of ambiguity here and, in fact, the ambiguity appears in the consequent, where both *I* and *myself/me* seem to allow an interpretation in which they designate Fred and an interpretation in which they designate Larry, yielding the interpretations below⁸:

- (6') a. If I_F were you_L , I_F would not trust myself/me_F.
b. If I_F were you_L , I_F would not trust myself/me_L.
c. If I_F were you_L , I_L would not trust myself/me_L.
d. If I_F were you_L , I_L would not trust myself/me_F.

This might seem due to the fact that (6) is a rather peculiar example, but example (7), which seems fairly common, presents exactly the same type of ambiguity⁹, the only difference being that only the first person pronoun is involved in the consequent, yielding two different interpretations, rather than four:

- (7') a. I_G shouldn't if I_G were you_W .
b. I_W shouldn't if I_G were you_W .

That this is the case is shown by Willy the Twig's answer which gives an explicit interpretation of the two solutions. Thus, this seems to be a clear example of ambiguity in indexicals, just as (6). It should be noted that in both cases the ambiguity is to be found in indexicals which occur in the consequent of a counterfactual. Whether it can or cannot occur in other contexts remains to be

⁸ Where indice *F* corresponds to Fred, while indice *L* corresponds to Larry.

⁹ Indices *G* and *W* designate respectively Georges and Willy the Twig.

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

seen. I will first point out the problem which this raises for the current analysis of indexicals and then show what I take to be the source of this ambiguity.

5. The current theory of indexicals and the ambiguity of indexicals

As said above, the current theory of indexicals does not seem to allow any possibility of ambiguity: thus any example of ambiguity, as long as the ambiguity cannot be reduced to a misunderstanding would falsify it. This paragraph will explain why this is so and will also distinguish the difficulty raised by ambiguities such as those which can be found in (6) and (7) for the current theory of indexicals and the difficulty which these ambiguities raise for the hypothesis of personal infallibility.

Why should ambiguity be dangerous for the current analysis of indexicals? According to this analysis, an indexical has a procedure lexically attached to it, and this procedure directs the hearer's quest for the referent. It is sufficiently constrained for the referent to be found in any cases but those where the situation itself is unclear. Thus, the procedure should only fail when the situation is not clear but not in any other case. Yet, examples such as (6) and (7) apparently show that a failure of the procedure is possible even though the situation appears to be perfectly clear. Let's come back to examples (6) and (7): there does not seem to be any way in which the situation could be described as unclear as far as the procedures attached to *I* and *you* are concerned. In (6) Fred is talking and this is certainly not less true in the consequent than it is in the antecedent. In (7), George is speaking, and, again, this is true in both the consequent and the antecedent. Thus it would seem that there should not be any possibility of ambiguity or misunderstanding or uncertainty as to the referents of the indexicals in these examples: the two first person pronouns in the consequent of (6) should refer to Fred and the first person pronoun in the consequent of (7) should refer to George. Yet, clearly, such an ambiguity does exist.

So it seems now that two questions should be answered:

- (i) *where does this ambiguity come from?*
- (ii) *can it be explained without discarding the current analysis of indexicals?*

There is still a further question which the possibility of ambiguity in indexicals and especially in the first person pronoun raises:

- (iii) *can any ambiguity in the first person pronoun be consistent with the hypothesis of first person or personal infallibility?*

I will begin by question (iii). Before it can be answered, the notion of *personal infallibility* should be explained: it has been noticed¹⁰ that the speaker cannot be in error as to whom it is that he or she refers to when he or she uses the first person pronoun. This should not be interpreted as meaning that one cannot be in error as to what one is: one could be (and presumably is, from time to time) in error as to one's own properties, but one cannot use *I* in order to refer to

¹⁰ Notably by Nozick 1981.

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

someone else and there is thus no possibility of error. This, it seems, is strongly contradicted by examples such as (6) and (7) in which it seems that the first person pronoun can be interpreted as referring to someone who is not the speaker (Larry in (6) and Willy the Twig in (7)): the very possibility of this interpretation implies that Fred can use the first person pronoun to refer to Larry and that George can use it to refer to Willy the Twig. Thus it seems that the ambiguity in the first person pronoun is not only a counter-example to the current analysis of indexicals, but that it is as well a counter-example for the notion of personal infallibility.

6. Counterfactuals and ambiguity

The first thing to investigate is the existence of examples which closely resemble examples (6) and (7), which are reproduced below:

- (6) If I were you, I wouldn't trust myself/me.
- (7) 'I shouldn't, if I were you', said George benignly.
'On the contrary,' said Willy the Twig, 'being you, of course you wouldn't, but if you were me you certainly would'.
(E. Peters, *Rainbow's end*, 19).

These examples are very similar: they both are counterfactuals and their antecedents are identical. These antecedents are hypothetical identity statements of the form "x is y" or "x = y". Examples which seem highly similar though they are not identical are examples in which the antecedent is an hypothetical identity statement where one at least of the component is not an indexical. Such examples can be found fairly easily as the following example, borrowed from Goodman (1984), shows:

- (8) a. If I were Julius Caesar, I would not be living in the 20th century.

The first noticeable thing about (8) is that the indexical in the consequent is not ambiguous: if this utterance has been produced by Bill, then, clearly, Bill is the individual designated by *I*. Julius Caesar is just not in the picture at all. The other thing is that, as Goodman himself points out, counterfactuals with hypothetical identity statements as antecedents are what he calls *counteridentical*, that is, the same antecedent can receive two basically contradictory consequents, as can be seen in (8b)¹¹:

- (8) b. If I were Julius Caesar, he would be living in the 20th century.

In (8a) and (8b), there is apparently no ambiguity in the antecedent and thus it is hard to understand why the same antecedent should allow two different consequents (between which there does not seem to be any reason to choose). I want to suggest here that those two consequents can follow the same antecedent because, indeed, this antecedent is somehow ambiguous. I also

¹¹ Goodman, in the antecedent of (8), reverses the order of the NPs: "If Julius Caesar were me,...". However, as he himself points out, the antecedent being an identity statement, the order of the NPs should be irrelevant as *I am Julius Caesar* and *Julius Caesar is me* are strictly equivalent from a logical point of view. I will come back to this later.

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

want to suggest that Willy the Twig's reply in (7) emphasises the fact that an antecedent such that *if I were you*, is open to two contradictory consequents in the same way and hence that such counterfactuals are, of necessity, counteridenticals. Thus, according to what as just been said, counterfactuals with identity statements as their antecedents are counteridenticals because their antecedents are somehow ambiguous. I will now try to explain why some identity statements, when they occur as the antecedent of a counterfactual, are ambiguous, which is certainly not the case when they occur in non-counterfactual contexts.

7. Identity statements in counterfactuals

The ambiguity of (at least) some identity statements in counterfactuals raises some questions:

(iv) *Are all identity statements ambiguous in counterfactuals?*

(v) *Why are identity statements ambiguous in counterfactuals when they are not ambiguous outside of counterfactuals?*

(vi) *Are the identity statements ambiguous in (6), (7) and (8) because of the presence of indexicals or would they be ambiguous if names were substituted to the indexicals?*

It should be clear that there are relationships between the first and the last questions: if the answer to (vi) was to the effect that the ambiguity comes from the indexicals, then clearly the answer to question (iv) would be negative. So I will begin by question (vi), come back to question (iv) and then proceed, in the next section, toward a (tentative) answer to question (v).

The answer to question (vi) is fairly simple and passes through the substitution proposed below:

(6'') If Fred were Larry, he would not trust himself/him

(7'') If George were Willy the Twig, he wouldn't do it.

(8'') a. If Bill were Julius Caesar, he would be living in the 20th century.
b. If Bill were Julius Caesar, he would not be living in the 20 th century.

The pronouns in the consequents of (6'') and (7'') give raise to exactly the same ambiguity as they did in (6) and (7), thus indicating that the antecedent is not less ambiguous in (6'') than it is in (6) or in (7'') than it is in (7). In the same way, the counterfactuals in (8'') are just as counteridentical with this antecedent as they were with the antecedent in (8). Thus it seems that the presence of the indexicals - which, anyway, as was said above, are not themselves ambiguous - is not the cause of the ambiguity.

Let us now come to question (iv): obviously the answer to (vi) does not answer (iv). So let us take a few other examples and test if they are or are not counteridenticals:

(9) a.If the Earth were the Moon, Earthmen would be Moonmen.

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

b. If the Earth were the Moon, Moonmen would be Earthmen.

(10) a. If Anne Reboul were Pope Jean-Paul II, he/she would not forbid condoms.

b. If Anne Reboul were Pope Jean-Paul II, he/she would forbid condoms.

Though there may be examples of identity statements in counterfactuals which are not ambiguous (i.e. which are not counteridentical), they seem highly difficult to find. I think that the answer that there are no identity statements in counterfactuals antecedents which do not produce counter-identicals should be accepted, at least provisionally. I will now come to question (v).

8. Identity statements, counterfactuals and counteridenticals

First of all, it is important to establish that identity statements outside of counterfactuals are not ambiguous, that is, if a true identity statement is used as the antecedent of a conditional, the consequent cannot freely be any of two contradictory propositions. This can also be shown from the truth table of material implication which says that when the antecedent is true, the whole proposition is true iff the consequent is true. Given the principle of the excluded middle, this means that only one of (12a) or (12b) can be the consequent of a true conditional which would have (11) as its antecedent:

(11) The Morning Star is the Evening Star.

(12) a. The Morning/Evening Star is a planet.

b. The Morning/Evening Star is not a planet.

Thus the kind of ambiguity which seems to arise naturally from identity statements in counterfactuals certainly is not a general feature of identity statements. Hence it seems that this ambiguity of identity statements in counterfactuals is peculiar to this context and should be accounted for through some kind of peculiarity which would come from it.

I think that this can be explained: counterfactuals have the peculiarity of presenting their antecedents as false¹². In most cases, this does not raise any special difficulty, as their antecedents are usually propositions of a kind whose truth or falsity is a contingent matter¹³. This, however, is not the case for identity statements: as Kripke (1980) has shown, identity statements when they are true are necessarily true and when they are false are necessarily false¹⁴.

¹² Which, by the way, is one the features which makes it impossible to treat them as material implications. It should be noted that though counterfactuals present their antecedent as false, this does not make their antecedents false: a counterfactual could quite well have a true antecedent and this would not prevent the counterfactual from presenting its antecedent as false, i.e. from being a counterfactual. For the sake of simplicity, I will only look here at counterfactuals with false antecedents.

¹³ In possible worlds theory, contingency means that a proposition is true at some worlds and false at others.

¹⁴ In possible worlds theory, necessity is defined by truth (or falsity) at all possible worlds.

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

Thus counterfactuals with an identity statement as antecedent are what Lewis (1973) calls *vacuously true*: in other words, from the antecedent, any consequent whatsoever may be deduced. Up to a point, this might seem a good way of accounting for counter-identicals, where the same antecedent seems to allow contradictory consequents. However, though it is certainly true both that identity statements in counterfactuals are necessarily false (i.e. impossible) and that from an impossible antecedent, one may deduce any consequent, there does nevertheless seem to be a difference between counterfactuals like (10a) and (10b) and a counterfactual as (10c):

(10) c. If Anne Reboul were the Pope, pigs would have wings.

Whereas in (10a) and (10b), though they are vacuously true, the consequent does not seem completely arbitrary, given the antecedent, in (10c) there does seem to be no relation of any kind between the antecedent and the consequent.

This discrepancy should be explained and presumably this explanation should also account for the ambiguity of identity statements in counterfactuals. That is, it should also be able to answer the following question:

(vii) *How can identity statements be ambiguous in counterfactuals?*

9. The ambiguity of identity statements in counterfactuals

To answer the above question, one should answer the more simple question: *How can identity statements be ambiguous?* The answer to that question is simple: identity statements cannot be ambiguous. We thus seem to find ourselves in a quandary: there obviously appears to be ambiguous identity statements in counterfactuals and yet identity statements cannot be ambiguous. The question arises then of whether they really are ambiguous in counterfactuals. On the face of it, it would seem that they are not: an identity statement such as (13) is not ambiguous and it is undubitably false and necessarily false:

(13) Anne Reboul is Pope Jean-Paul II.

None of this is modified by the insertion of (13) in counterfactuals such as (10) a, b or c, and this is shown by (10c) where it is clear that any consequent can be derived from the impossible antecedent. Thus, identity statements do not seem to change their meanings in counterfactuals. So how are counteridenticals possible?

I think that the answer is that identity statements are not any more ambiguous in counterfactuals than they are out of counterfactuals and that a given identity statement has the same meaning, truth-conditions and indeed truth-value in and out of a counterfactual. Yet, the existence and interest of counteridenticals cannot be denied. I think that a more detailed analysis of the consequent in examples (6)-(8) could be illuminating. In these examples, the indexicals are ambiguous, which means that the consequent can be interpreted in two ways. These two possibilities of interpretation show quite clearly that such counterfactuals are counter-identicals just as are examples (9)-(10). In other words, and leaving aside for a moment example (6), the pronoun in the consequent of (7) may designate either the speaker or his hearer. This in itself

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

should clearly indicate that the antecedent, though it certainly is an identity statement, is not interpreted as an identity statement: if it were, there could not be any kind of ambiguity as the pronoun could only be interpreted as referring to one individual, the individual which is both the speaker and the hearer, both George and Willy the Twig. *Mutatis mutandis*, the same reasoning applies to (8), (9) and (10): there should be only one individual available for reference in the consequent, and not two.

What exactly does all this mean? It does not mean that when the hearer processes the identity statement in a counterfactual, he supposes it to be true as he would do with a more “usual” type of counterfactual¹⁵: it means, rather, that he interprets the identity statement non literally, the most usual non literal interpretation being something like *if I were in your position,...* or *if you were in your position with my opinions,...* In other words, the speaker of a counterfactual with an identity statement as antecedent is not speaking literally and does not intend his utterance (or rather the antecedent of the counterfactual which he utters) to be understood literally. In Sperber and Wilson’s terms¹⁶, he wants to communicate a thought of which (this part of) his utterance is a less than literal interpretation.

Before I proceed to outline this hypothesis in more details, I should point out that this solution meets with a difficulty: what I am suggesting is that the antecedent has an uncertain interpretation, not in that the identity statement is ambiguous from a semantic point of view but in that it is used in a less than literal way and that induces a (non-semantic) uncertainty, this uncertainty explaining the “ambiguity” of the referring expressions in the consequent. However, this does not mean that the consequent is used in a less than literal way: only the antecedent is, and the consequent may be but is not necessarily, used in a non-literal way.

What exactly does it mean in Sperber and Wilson’s analysis to say that an utterance¹⁷ is used in a non-literal way? It means that the utterance is a representation with a propositional form of a thought which is itself a representation with a propositional form. What makes the utterance a representation of the thought is the degree of resemblance between the utterance and the thought. This resemblance works between the propositional form of the utterance and the propositional form of the thought and is relative to the number of common implications of these propositional forms when they are processed against a common context¹⁸. If the two representations share all their contextual implications, the resemblance is total and the utterance is a literal representation of the thought. If the two representations share some but not all of their contextual implications, the resemblance is less than absolute and the utterance is a less than literal representation of the thought. And last if the two representations do not share any of their contextual implications, there is no resemblance and the utterance is not a representation of the thought. It

¹⁵ For instance:

“If this glass had fallen down, it would have been broken.”

¹⁶ In Sperber and Wilson 1986.

¹⁷ Or, in this case, part of an utterance.

¹⁸ A *context*, in Sperber and Wilson’s terms, is a set of propositions.

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175
 should be clear that resemblance, and hence literality, is not an absolute but a relative notion.

How could this be applied to the problem of the non-literality of identity statements in the antecedent of counterfactuals? In other words what common implications do identity statements as the antecedents of counterfactuals have with the thoughts (or what we can surmise of them) which they represent? They share on each occasion of use **either** an implication of form *the speaker has some properties that the hearer has* **or** an implication of form *the hearer has some properties that the speaker has*. I should presumably emphasize that this is not meant as a exhaustive list. Hence the ambiguity of identity statements in the antecedent of counterfactuals is not a true semantic ambiguity: it just comes from the fact that there is a certain degree of uncertainty as to which implication was actually intended by the speaker.

Before coming to an answer to the problem of the literality of the consequent, when the antecedent is not literal, I would like to sum up briefly both the questions which have been raised and the answers which have been given to them.

10. The state of the questions

I have asked a few questions above. I want now to take stock of the answers which have been given to some of them in order to answer those which have not been before:

- (i) *where does this ambiguity come from?*
- (ii) *can it be explained without discarding the current analysis of indexicals?*
- (iii) *can any ambiguity in the first person pronoun be consistant with the hypothesis of first person or personal infallibility?*
- (iv) *Are all identity statements ambiguous in counterfactuals?*
- (v) *Why are identity statements ambiguous in counterfactuals when they are not ambiguous outside of counterfactuals?*
- (vi) *Are the identity statements ambiguous in (6), (7) and (8) because of the presence of indexicals or would they be ambiguous if names were substituted to the indexicals?*
- (vii) *How can identity statements be ambiguous in counterfactuals?*

Which of these questions have been answered? Question (i) has been answered: the ambiguity of referring expressions in the consequent of counterfactuals which have identity statements as antecedent comes from an uncertainty in the interpretation of these antecedents. Question (ii) still has to be answered, though it should be clear by now that it will be answered through the answers given to other questions ((i) and (vii) mainly). The same thing is true of question (iii). Question (iv) has been given a provisional answer and, given the answer to question (v), should be reformulated as *Are all identity statements non-literal in counterfactuals?* I will return to it in the next section.

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

Question (v) has been answered: identity statements (at least some of them) are not ambiguous either in counterfactuals or out of them; nonetheless they do not receive a unique and straightforward interpretation due to some uncertainty as to what it was that the speaker wanted to communicate. Question (vi) has been answered: identity statements as antecedents of counterfactuals may not receive a single interpretation regardless of whether indexicals or other referring expressions are used. Question (vii) has received the same answer as question (v).

Thus we still have to answer questions (ii), (iii) and (iv). We may add to them question (viii):

(viii) *How could the consequent of counterfactuals such as (6), (7), etc. be literal when the antecedent is not?*

11. Are all identity statements non-literal in counterfactuals?

On the face of it, it seems that the answer should be a resounding yes. And this was the provisional answer to question (iv) given in section 9. I think however that it is not entirely right and that some (necessarily false) identity statements in the antecedent of counterfactuals are used and interpreted in a literal way. To see that this is the case, let us come back to a few of the examples given above:

- (10) a. If Anne Reboul were Pope Jean-Paul II, he/she would not forbid condoms.
 b. If Anne Reboul were Pope Jean-Paul II, he/she would forbid condoms.
 c. If Anne Reboul were Pope Jean-Paul II, pigs would have wings.

I think that a case can be made that the same necessarily false identity statement which is the antecedent of all of (10a), (10b) and (10c) is used non-literally in (10a) and (10b) and literally in (10c). Before I outline why I think this is the case, I should make it clear that this answer is directly linked to the question of the lack of motivation of the consequent of (10c), whereas the consequents of (10a) and (10b) do not seem unmotivated or arbitrary in the same way.

What I am trying to point out is that in (10c), though not in (10a) and (10b), we have a good example of a vacuous counterfactual where anything would be the consequent. This is not the case with (10a) and (10b), where, in fact, it can be said that the consequent determines up to a point the way the antecedent should be understood. How is it that (10c) allows any consequent when (10a) and (10b) have determinate consequents? I want to claim that this can be explained by the fact that the antecedent has been used literally (in the case of (10c)) or non literally and with a specific interpretation intended (in the case of (10a) and (10b)). The fact that the antecedent is used literally or non-literally (and which interpretation is intended) is clearly indicated by the consequent.

This, however, is only part of an answer: it still leaves a question which is: *How does the consequent indicate that the antecedent should be interpreted literally or non literally, and if non literally, with this or that interpretation?* Given that the

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

antecedent can be interpreted literally or non-literally and if non-literally in two opposite ways¹⁹, and that different consequents appear after it in these different cases, it may be surmised that the consequent is a clue to the intention of the speaker as far as the antecedent is concerned. This can be explained from the semantics of the antecedent: if the antecedent is intended literally, it can only be seen as necessarily false, which means that the whole of the counterfactual is vacuously true. Hence any consequence whatsoever can be drawn from it. If the antecedent is interpreted as less than literal, it is still necessarily false, but its truth value is less relevant than are the implications which can be drawn from it. I pointed out above that these implications can roughly be said to fall under two main types:

- (i) implications of the form *A has some of the properties which B has*.
- (ii) implications of the form *B has some of the properties which A has*.

In the case of (10), if the antecedent is intended non-literally and if the interpretation chosen is of the second form, (10a) is rather natural; if, on the other hand, the interpretation chosen is of the first form, (10b) is also rather natural.

Briefly, when the antecedent of a counterfactual is an identity statement, the same antecedent can be followed by three highly different types of consequent: consequents which do not seem to have any kind of relation whatsoever with the antecedent; consequents which do have a relation of some kind but not so much to the antecedent itself than to one implication of the antecedent; consequents which do have a relation of some kind to another implication of the consequent. The first type occurs when the antecedent is interpreted literally, because its necessary falsity is the only relevant factor justifying its use. The second and third type occur when the antecedent is not interpreted literally and when, hence, one of its implications is the factor justifying its use.

12. The asymmetry of non-literal identity statements

Before I come to questions (i), (ii) and (vii), I would like to expand a bit on why such propositions as *A has some of the properties which B has* and *B has some of the properties which A has* may be said to be implications of an utterance saying that *A is B*. First of all it should be pointed out that these propositions are implications of *A is B* regardless of whether *A is B* is used or interpreted literally or not. As will be remembered, Leibniz proposed three laws, which are still widely accepted nowadays (see Wiggins 1980), governing identity. We will only be interested here by the first one: "The first is what is (...) called Leibniz's law, which says that if A and B are identical, then everything that is true of A is true of B, or $[A = B \Rightarrow (f) (fA \equiv fB)]$ " (Ishiguro 1990, 17). In other words, two identical things share all their properties: this entails that they share some of their properties, i.e. that they are similar to a degree²⁰. Thus if A is B, there are some properties of A which are also properties of B and some properties of B

¹⁹ Or at least in two ways which allow contradictory consequents.

²⁰ In other words, identity implies similarity, though the reverse is not true. See Reboul to appear.

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

which are also properties of A, because all the properties of A are also properties of B and *vice versa*. Thus *A has some properties which B has* and *B has some properties which A has* are two completely normal implications of a proposition such as *A is B*. What is different between the true proposition *A is B* interpreted literally and the necessarily false proposition *A is B* interpreted non-literally is the fact that, in the first case, the speaker presumably intends that the implication *all the properties of A are properties of B and vice versa* will be communicated whereas in the last case he does not intend this implication to be drawn from his utterance. However, he does intend in this last case that **either** an implication of the form *A has some of the properties which B has* **or** an implication of the form *B has some of the properties which A has* should be drawn.

It should be noted that, in this case however, these implications are not symmetrical whereas propositions such as *A is B* or *all the properties of A are also properties of B and vice versa* obviously are. This goes with the fact that different consequents will occur after the antecedent *A is B* in counterfactuals depending on whether the implication chosen is the former or of the latter type. Note however that this cannot be explained only on the basis that these implications come from the fact that identity implies similarity: similarity is as much a symmetrical relationship as is identity. Rather, these implications have to be interpreted in relation to the linguistic setting of the (fragment of) utterance from which they are drawn: in the antecedent of a counterfactual. In other words, the relevant implications are to be counterfactually interpreted, as *one or more property which belongs to A is attributed counterfactually to B* or that *one or more property which belongs to B is attributed counterfactually to A*. And here the asymmetry is readily explainable: properties which belong to an individual B and are counterfactually attributed to an individual A do not belong to A and properties which belong to A and are counterfactually attributed to B do not belong to B and may be different from the previous ones.

It could be objected that none of the examples presented above (apart from those with indexicals) seem to be asymmetrical: it seems that we could substitute the terms of the identity statement without changing anything. This, however, is not completely true:

- (8) a. If I were Julius Caesar, I would not be living in the 20th century.
b. If I were Julius Caesar, he would be living in the 20th century.

As was said in note 12, Goodman reversed the order of the NPs in (8b). I did not want to discuss the possible asymmetry of the antecedent of such counterfactuals above, but there is an asymmetry, though it is not a very strong asymmetry. In other words, example (8b) would be much more natural as *If Julius Ceaesar were me, he would be living in the 20 th century*. It is not a strong asymmetry in that both forms are possible, but it exists nonetheless and its very lack of strength does go rather well with a pragmatic analysis such as that which is proposed here.

Finally, it should be noted that this asymmetry, however weak, supports the non-literal analysis of the antecedents of counteridenticals. The same asymmetry is also to be found in metaphors of form *A is B*, as can be seen in examples (14a) and (14b):

- (14) a. L'homme est un roseau pensant.

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

(Man is a thinking reed)

b. Le roseau pensant est un homme.

(The thinking reed is a man)

It also occurs in some non-literal similes, such as (15a):

(15) a. Darkness had fallen like the House of Usher.

(M. Lowry 1984, *Under the volcano*, 21).

b. The House of Usher had fallen like darkness.

Both in (14a-b) and in (15a-b), the substitution changes the meaning, which means that neither (14a) nor (15a) are symmetrical. Thus asymmetry, when it is inconsistent with the semantics of the utterance, gives a strong presumption of non-literality²¹.

13. Can the apparent ambiguity of indexicals be explained without discarding the current analysis of indexicals?

Let us begin by example (7). As I pointed out before, the consequent in (7) is on the face of it ambiguous as it should be possible to interpret the pronoun *I* in the consequent as referring either to Willy the Twig or to George:

(7) 'I shouldn't, if I were you', said George benignly.

'On the contrary,' said Willy the Twig, 'being you, of course you wouldn't, but *if you were me you certainly would*'.

(7') a. I_G shouldn't if I_G were you_W.

b. I_W shouldn't if I_G were you_W.

I have tried to show above that this ambiguity in the consequent comes from the uncertainty in the interpretation of the antecedent, this, in its turn, being due to its non-literal use. In other words, the speaker could be communicating either the implication I_G (*counterfactually*) *have some of your_W properties* or you_W (*counterfactually*) *have some of my_G properties*. What implication, if any, does this have for the current analysis of indexicals? In the first case, where the implication I_G (*counterfactually*) *have some of your_W properties* has been chosen, there does not seem to be any problem, as far as the interpretation of the pronoun in the consequent is concerned, because this corresponds to interpretation (7'a) where *I* designates George, i.e. the speaker, as would be expected from the current analysis of indexicals. The problematic interpretation is the interpretation (7'b) where implication you_W (*counterfactually*) *have some of my_G properties* has been chosen for the antecedent and where *I* in the consequent designates Willy the Twig, the hearer. Thus the question above reduces to whether we can account for this interpretation without discarding the current analysis of indexicals.

²¹ All of this, it should be remarked, is not directly related to the truth or falsity, necessary or contingent, of the proposition expressed. The following example (from John Donne) is true, metaphorical and asymmetric:

"No man is an island."

"No island is a man".

On all these problems, see Reboul 1991 and Reboul to appear.

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

What about example (6) and its four interpretations?

- (6) If I were you, I wouldn't trust myself/me.
- (6') a. If I_F were you_L , I_F would not trust myself/ me_F .
 b. If I_F were you_L , I_F would not trust myself/ me_L .
 c. If I_F were you_L , I_L would not trust myself/ me_L .
 d. If I_F were you_L , I_L would not trust myself/ me_F .

The first thing to remark is that the example, such as it stands, could in fact be seen as two different examples²²:

- (6) a. If I were you, I wouldn't trust myself.
 b. If I were you, I wouldn't trust me.

These two new examples do not give rise to exactly the same interpretations:

- (16) a. If I_F were you_L , I_F would not trust myself $_F$.
 b. If I_F were you_L , I_L would not trust myself $_L$.
- (17) a. If I_F were you_L , I_F would not trust me_L .
 b. If I_F were you_L , I_L would not trust me_F .

Let us begin with (6b) and its two interpretations (17a) and (17b): in both interpretations, the pronoun *me* in the consequent refers to another individual than does the pronoun *I* in the consequent. In the first case, (17a), *I* refers to Fred (the speaker) and *me* refers to Larry (the hearer). This interpretation is linked with the implication I_F (counterfactually) *have some of your $_L$ properties* for the antecedent. In the second case, (17b), *I* refers to Larry (the hearer) and *me* refers to Fred. This interpretation is linked to implication You_L (counterfactually) *have some of my $_F$ properties* for the antecedent.

For (6a) and its two interpretations (16a) and (16b), in both interpretations, the pronouns *I* and *myself* in the consequent refer to the same individual. In the first case, (16a), *I* refers to Fred (the speaker) and so does *myself*. This interpretation is linked with the implication I_F (counterfactually) *have some of your $_L$ properties* for the antecedent. In the second case, (16b), *I* refers to Larry (the hearer) and so does *myself*. This interpretation is linked to implication You_L (counterfactually) *have some of my $_F$ properties* for the antecedent.

Thus, the coreference or absence of coreference between the two pronouns in the consequent depends, not surprisingly, on the linguistic form chosen, that is on the fact that either the simple pronoun *me* or the reflexive *myself* has been used. In other words, *me* is interpreted as referring to someone different from the referent of *I* in the consequent, because if the speaker wanted his hearer to interpret the two pronouns as coreferential, he would have used the reflexive *myself*. Yet, examples (6a) and (6b), in some

²² This is not mandatory, but the interpretations appear much more natural.

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

interpretations, raise a difficulty which is highly similar to that raised by example (7) in interpretation (7'b): both interpretations (16b) and (17b) are problematic for the current interpretation of indexicals in as much as the referent of the first person pronouns is not the speaker of the utterance. Again, the question arises: why do *I* and *myself* in interpretation (16b) and *I* in interpretation (17b) refer to the hearer?

14. The quasi-indexical account

It should just be noted that the main problem arises in (6b) where the two first person pronouns cannot be interpreted as coreferential. This means that (6b) raises a problem no matter which interpretation (either (17a) or (17b)) is chosen. In any case, we have two first person pronouns, *I* and *me*, which are not coreferential and which, hence, cannot both be interpreted as referring to the speaker. In this sense, (6b) is especially crucial for the current account of indexicals. There seems indeed to be a very natural answer to the problem: if both first person pronouns cannot refer to the speaker, then one of them (at least) is not interpreted as the current analysis of indexicals predicts. This means that the current analysis flounders.

There is, however, another solution: if both pronouns cannot refer to the speaker of the utterance, then they are not both interpreted relative to the same description of the (same) situation. How can this suggestion be applied to (6b) in interpretation (17b)? Under this interpretation, *I* designates the hearer, while *me* designates the speaker. What descriptions of the situation should be used to account for these attributions of different referents? One of them, obviously, is the description of the situation of communication in which the speaker, Larry, produces the utterance. The other one can only be the reverse. How could this come to be the case? The first thing that comes to mind is that these descriptions are somehow linked to the implication chosen for the antecedent. Yet neither of these implications is by itself a description relevant to the assignation of referents to the indexicals²³.

There is, however, a possibility. As seen above, in counteridenticals, the antecedent is asymmetrical in that the order of the subject and object strongly influence the choice of the relevant implication. Let us come back to examples (8a) and (8'b):

(8) a. If I were Julius Caesar, I would not be living in the 20th century.

(8') b. If Julius Caesar were me, he would be living in the 20th century.

(8a) is associated with implication *I (counterfactually) have some of Julius Caesar's properties*, while (8'b) is associated with implication *Julius Caesar (counterfactually) has some of my properties*. In much the same way, I have said above that interpretation (17a) and (17b) are respectively linked to implications *I_F (counterfactually) have some of your_L properties* and *You_L have some of my_F properties*. If we come back to (8a) and (8'b), we can see that the asymmetry in the antecedent is not only reflected in the consequent chosen, it also manifests itself

²³ The kind of description which would be relevant would be of form *x is the speaker, y is the hearer*.

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

through the choice of the pronoun in the consequent. Thus, in (8a), the first person pronoun is used, while in (8'b) the third person pronoun is used. This, I claim, is due to the fact that the choice of this or that implication implies the choice of such or such a point of view for the consequent (in this case, either mine or Julius Caesar's). Under this hypothesis, **the personal pronouns in the consequent are not indexicals, but quasi-indexicals** and are interpreted relative to the point of view implied by the choice of a given implication for the non-literal antecedent.

The notion of quasi-indexical has been developed by Castañeda (1979a, 1979b and 1989). While indexicals express our own indexical references, quasi-indexicals enable us to attribute indexical reference to other people. This does not, however, mean that there are two different sets of linguistic expressions, one of which would gather the indexicals, while the other would gather the quasi-indexicals. Rather, some linguistic expressions can be used either as indexicals or as quasi-indexicals. Thus, all the singular personal pronouns can be used as either indexicals or quasi-indexicals and *I* is used as an indexical (in the antecedents of (6a), (6b), (7) or (8a)), and as a quasi-indexical (in the consequents of (6a), (6b), (7) and (8a)). *He* is used as a quasi-indexical in (8'b), representing the (first person) reference which Julius Caesar would make to himself (i.e. Julius Caesar's point of view).

Thus, we can account for the interpretation (17b) in (6b) of *I* as referring to Larry (the hearer): this interpretation is linked with the implication *you_L* (counterfactually) *have some of my_F properties* which implies in the consequent that Larry's point of view should be represented in the consequent. Thus *I* in (6b) is interpreted relative to a description in which Larry is the speaker, a situation which, it should be noted, is different from the real situation, in which Fred is speaking. This hypothesis, that the consequent of the counterfactual is interpreted as representing a specific point of view, depending on the implication associated with the antecedent, can, *mutatis mutandis*, account for both interpretations (7'a) and (7'b) for (7) and (16a) and (16b) for (6a). It can, however, only account for the interpretation of the first pronoun, *I*, in interpretation (17a) and (17b) for (6b). So how can we account for the interpretation of the second pronoun, *me*?

The first thing that comes to mind is that in the course of the interpretation of the consequent, the description changes because the other implication for the antecedent takes the advantage. This, however, would be very weird: after all the interpretation of the antecedent controls the interpretation of the consequent and should not change along the way. This suggestion is thus not tenable. Yet, I think that it is true that the description of the situation changes for the interpretation of *me*, though this change is not directly linked to the implication chosen for the interpretation of the antecedent. It is just the result of the fact that *me* favours a non-coreferential interpretation and that just two descriptions of the situation are available: thus the current one is abandoned for the interpretation of *me*, given that the coreferential *myself* which might have been used has not been chosen. It should be noted that under this view *me* in interpretation (17b) is not a quasi-indexical, but an indexical.

15. Preferential interpretation

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

The analysis outlined above, however, is not the whole story. First of all, though it accounts for interpretation (17b), it does not account for interpretation (17a). It should also be noted that (8a) and (8'b), through the asymmetry of the antecedent, i.e. through the implication favoured by the order of the NPs in the antecedent, receive a **preferential interpretation**. This can be seen through the fact that, though both antecedents could, theoretically, be followed by either consequent *I would not be living in the 20th century* or by consequent *He would be living in the 20th century*, the acceptability of the resulting sentences does seem equivalent:

- (8) a. If I were Julius Caesar, I would not be living in the 20th century.
a'. ?If I were Julius Caesar, he would be living in the 20th century.
- (8') b. If Julius Caesar were me, he would be living in the 20th century.
b'. ??If Julius Caesar were me, I would not be living in the 20th century.

Thus the question arises for examples (6a), (6b) and (7): *Is there in the same way a preferential interpretation for them?*

I think that the two questions, that of an account of interpretation (17a) and that of preferential interpretations for (6a), (6b) and (7) should receive a common answer and that there **are** preferential interpretations for these examples just as there are preferential interpretations for (8a) and (8'b). The preferential interpretation for (6a) is (16a), for (6b) is (17b) and for (7), (7a). I will now try to account for them and to show that these preferential interpretations do not contradict the current analysis of counterfactuals.

Let us begin with (6a) and (16a). In this case, there is absolutely no problem for the current account of indexicals: all the first person pronouns designate the speaker and this is as it should be according to the current analysis of indexicals. The same thing goes for interpretation (7'a) of George's counterfactual and also for Willy the Twig's counterfactual in (7). What about the reason for this preference? As said above, the antecedent in (6a) and (7) leads rather naturally to implication *I_(the speaker) (counterfactually) have some of your_(the hearer) properties*, whereas an antecedent of form *If you were me* would rather lead to implication *You_(the hearer) have some of my_(the speaker) properties*. This is shown by the following examples:

- (18) a. If you were me, you wouldn't trust yourself.
b. ?If you were me, I wouldn't trust myself.

Hence the preferential interpretations for (6a), (6b) and (7).

Could we in any way explain why interpretation (17b) is preferred for (6b)? It should, first of all, be noted that the account through the asymmetry of the antecedent is not available in this case: (6a) and (6b) share a common antecedent and if the preferential interpretation of (6b) is indeed (17b), the implication chosen for the antecedent in (6a) is not the same as the implication chosen for the antecedent in (6b) and, hence, only one of interpretations (16a) and (17b) can be explained by the asymmetry of the antecedent. We have already

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

used this explanation for (16a), where it applies rather naturally and something else must be used now.

(17b) as a preferential interpretation can, however, be explained by another factor: as noted above, *me* implies non-coreference with *I* in the consequent of (6b). In the analysis given above of interpretation (17b), this is accounted for through the fact that *I* is a quasi-indexical, used by the speaker, Fred, to attribute to Larry, the hearer, a quasi-indexical reference to himself (Larry). *Me*, because it has been chosen rather than *myself*, indicates non-coreference with *I*: hence, it is an indexical used by the speaker, Fred, to refer to himself (Fred). No such account is available for interpretation (17a): remember that, under the hypothesis outlined above, the consequent expresses the point of view of one of the referents of the two NPs in the antecedent of the counterfactual. Under interpretation (17a), the consequent would, in as much as implication I_F (*counterfactually*) *have some of your_L properties*, express Fred's point of view and *I* would refer to Fred. Given that *me* indicates non-coreference with *I*, this means that *me* should refer to Larry. But there is just no way that this could be the case: *me* cannot be a quasi-indexical in this case, because *myself*, which could have been used as shown in (6a), has not. And if *me* is not a quasi-indexical, it can only be interpreted through the description of the real situation, that is, as referring to the speaker, i.e. Fred. But that would not yield interpretation (17a) and neither would it yield any interpretation consistent with the principle of relevance, given the choice of *me*, rather than *myself* in (6a). It also explains why, despite the fact that implication I_F (*counterfactually*) *have some of your_L properties* should have been favoured because of the asymmetry of the antecedent, it is not: it would lead to no interpretation consistent with the principle of relevance, and thus to no possible interpretation.

15. Conclusion: putting Humpty Dupty together again

Let us come back to the eight questions in § 10.:

- (i) *where does this ambiguity come from?*
- (ii) *can it be explained without discarding the current analysis of indexicals?*
- (iii) *can any ambiguity in the first person pronoun be consistent with the hypothesis of first person or personal infallibility?*
- (iv) *Are all identity statements ambiguous in counterfactuals?*
- (v) *Why are identity statements ambiguous in counterfactuals when they are not ambiguous outside of counterfactuals?*
- (vi) *Are the identity statements ambiguous in (6), (7) and (8) because of the presence of indexicals or would they be ambiguous if names were substituted to the indexicals?*
- (vii) *How can identity statements be ambiguous in counterfactuals?*

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175

(viii) *How could the consequent of counterfactuals such as (6), (7), etc. be literal when the antecedent is not?*

Let us take these questions and remember the answers which they received above: the answer to question (i) says that the ambiguity of indexicals in the consequent of counterfactuals come from the fact that there is some uncertainty as to what the proper interpretation of the antecedent is. The answer to question (v) is that identity statements are not ambiguous either in or out of counterfactuals: rather they can be used non-literally in counterfactuals. The answer to question (iv) is that not all identity statements are used non-literally in counterfactuals: they are only used non-literally in counteridenticals. The answer to question (vi) is that the antecedents of counteridenticals may lead to uncertainty as to how they should be interpreted whether or not they contain indexicals. Question (vii) is equivalent to question (v) and receives the same answer. The answer to question (ii) goes through the answers to question (i) and (v): it is because, though identity statements are not ambiguous in counterfactuals, there may be some uncertainty as to how they should be interpreted (i.e. as to whether this or that implication should be chosen), that indexicals seem ambiguous in the consequent of counteridenticals. Questions (iii) and (viii) still have to be answered.

The answer to question (iii) seems, on the face of it, to be simple now: there is no ambiguity in the first person pronoun. Thus the current analysis of indexicals is not put in jeopardy by examples such as (6) and (7), and neither is the hypothesis of first person or personal infallibility. The answer to question (viii) is not more difficult: according to Sperber and Wilson's analysis of non-literality, the only requirement for something to be used in such a way is, if it is a single word, that this word should "infect" the segment of the utterance corresponding to a proposition, where it occurs. If, as in the case of counteridenticals, it is a sentence embedded in a complex sentence uttered, then, as long as there is a corresponding proposition, there is no problem. As identity statements certainly correspond to propositions, there is no discrepancy between the hypothesis that identity statements in counteridenticals are used non-literally and the theory of non-literality proposed by Sperber and Wilson.

In conclusion, I should like to point out that though there appears to be an authentic ambiguity in the indexicals used in the consequents of counteridenticals, in fact this "ambiguity" is nothing more than an uncertainty as to the description of the situation relative to which the indexicals should be interpreted. Thus, the "ambiguity" of indexicals in counteridenticals is a mere misunderstanding and neither the current account of indexicals nor personal infallibility are menaced.

REFERENCES

Castañeda, H.N. (1979a): "The causal and epistemic roles of proper names in our thinking of particulars", in French, P.A., Uehling jr., Th. E. & Wettstein, H.K. (eds): *Contemporary perspectives in the philosophy of language*, Minneapolis, University of Minnesota Press.

Castañeda, H-N. (1979b): "On the philosophical foundations of the theory of communication: reference", in French, P.A., Uehling Jr., Th. E. & Wettstein, H.K.

in *Acts of the 2nd International Colloquium on Deixis « Time, space and Identity »*, Nancy, 28-30 mars 1996. 151-175
 (eds.): *Contemporary perspectives in the philosophy of language*, Minneapolis, Univ. Of Minnesota Press, 125-146.

Castañeda, H-N. (1989): *Thinking, language & experience*, Minneapolis, University of Minnesota Press.

Goodman, N. (1984): *Fact, Fiction and Forecast*, Cambridge, Mass., Harvard University Press.

Ishiguro, H. (1990): *Leibniz's philosophy of logic and language*, Cambridge, Cambridge University Press.

Kaplan (1977/1989): "Demonstratives", in Almog, J., Wettstein, H. & Perry, J. (eds.): *Themes from Kaplan*, New York, Oxford university Press, 565-614.

Kripke, S. (1980): *Naming and necessity*, Oxford, Basil Blackwell.

Lewis, D. (1973): *Counterfactuals*, Oxford, Basil Blackwell.

Nozick, R. (1981): *Philosophical explanations*, Oxford, Clarendon Press.

Reboul, A. (1991): "Comparaisons littérales, comparaisons non-littérales et métaphores", in *TRANEL* 17, 75-96.

Reboul, A. (to appear): *Réalités de la fiction*.

Sperber, D. & Wilson, D. (1986): *Relevance: communication and cognition*, Oxford, Basil Blackwell.

Wiggins, D. (1980): *Sameness and substance*, Oxford, Basil Blackwell.