1. **Two Observations**

Common prejudice has it that nothing and nobody bears more than one name. When being introduced to someone, you do not ask for one of their names; neither do you expect what you find on a place name sign to be just one way of calling the town. To be sure, the town may have different names in different tongues, and a person may be called different names by different people or on different occasions, but within one language and context, one tends to think of naming as a one-one mapping. Expectations seem deep-rooted in this respect. This is why semantics texts usually have to remind students that distinct proper names may denote the same object nevertheless.\(^1\)

The observation I am interested in is not that the naive expectation is not borne out. Rather it concerns the more basic fact that there is such a naive expectation in the first place. For clearly, from the pre-theoretic point of view, the case of Mark Twain and Samuel Clemens is exceptional.\(^2\)

- **Observation 1**
  
  Names are generally expected to be unique.

It seems as if the observed expectation is not always fulfilled. Exceptional though Mark Twain may be, the fact that he has two names is not. Quite generally, pen names and other pseudonyms, historical place names, under-cover identities, erroneous re-namings etc. all go to show that the contention that names are unique is a naive overgeneralisation. Once this has been recognized, double names are made out all over the place. In particular, they are found in the semanticists’ and philosophers’ of language favourite environment of attitude reports, where they appear to fulfill the specific function of presenting their bearers from different angles, and where they defy substitution *salva veritate*. Starting with Frege’s Venus,\(^3\) the never-ending debate has focussed on the question of how presentational function and non-substitutivity of double names pertain to their meanings and how they relate to the interpretation of attitude reports at large. However, more recently dug up examples indicate that neither double naming-cum-presenting nor substitution failure are a matter of attitudes or even intensional contexts in general:

\[
\text{There are simple sentences [footnote suppressed] which evoke anti-substitution intuitions quite similar to those evoked by attitude-reporting sentences.} \quad \text{Saul (1997a: 102)}
\]

---

\(^0\) This is a considerably revised version of the *Sixth Weatherhead Lecture in Philosophy of Language*, which I had the honour to give at Tulane University in October 2002. Part of the material has also been presented at a workshop on Robert Stalnaker’s philosophy, at the *Sinn und Bedeutung 7* (both in Konstanz), to the Berlin *Semantikzirkel*, and in talks at the Tübingen linguistics department and at Reid Hall, Paris. I owe a great debt to comments and criticism that I received from various members of my audiences. In particular, I would like to thank Manfred Bierwisch, Graeme Forbes, Hans-Martin Gärtner, Ulrike Haas-Spohn, Aaron Konopasky, Cécile Meier, Friederike Moltmann, Robert van Rooij, Robert Stalnaker, Peter Staudacher, as well as two anonymous reviewers for the *Journal of Semantics*.

\(^1\) Cf. Dowty *et al.* (1981: 17): ‘It is […] allowed for one and the same individual to have two or more names (just as “Samuel Clemens” and “Mark Twain” are names of the same person).’

\(^2\) Observation 1 is usually regarded as irrelevant to semantic theory, though it has been connected to the *pragmatics* of proper names; cf. Lerner & Zimmermann (1991: 364), where the following principle is defended along largely Gricean lines:

\[\text{(M) Do not use different names for the same thing in the same context.}\]

\(^3\) It is questionable, of course, whether Frege’s (1892) original examples, the German noun phrases *der Morgenstern* and *der Abendstern*, are proper names at all (as opposed to definite descriptions). Frege himself did not seem to take the matter important.
The point is:

- **Observation 2**
  Non-substitutivity of coreferential names may occur in extensional contexts.

The example Saul uses to drive the point home is:

(1) **Clark Kent went into the phone booth, and Superman came out.**

Here a substitution of **Superman** by **Clark Kent** appears unwarranted, even though the two names are supposed to be known to belong to the same person. Moreover, Saul already mentioned, Observation 2 can also be substantiated by examples involving changing place names:\footnote{4 Saul (1997a: 103). I have allowed myself to adapt her example (**St Petersburg** vs. **Leningrad**) to my German habitat.}

(2) **I never made it to Karl-Marx-Stadt, but I visited Chemnitz last year.**

In the debate following Saul’s discovery, various proposals for accounting for Observation 2 have been made, some of which will be discussed in Sections 2 and 3. In the course of the discussion it will be revealed that one aspect of the data appears to have gone unnoticed. The aspect concerns certain substitutional **asymmetries** that go against a hidden assumption made throughout the debate, viz. that the various names of a given object are, as it were, semantically on a par. In Section 2, substitution puzzles like (2) will be accounted for in semantic terms, reducing the asymmetry of changing place names to (systematic) ambiguity. In Section 3 it will be shown that such a semantic approach runs into problems when it comes to cases like (1) involving hidden identities of persons. In Section 4 I will therefore explore a pragmatic explanation that builds on the prejudice that no object is supposed to have more than one name. It will turn out that the pragmatic approach fares better with cases like (1) than its semantic does. On the other hand, unlike the latter it cannot naturally deal with (2). If this is right, the two examples appear to call for different accounts and are not as parallel as it seems.

2. **Changing Names: A Semantic Account**

In this section, I will introduce a rather obvious semantic way of dealing with examples like (2). At the end of the section I will discuss the kind of asymmetry announced above and adapt the semantics of changing names accordingly (which will not present a great problem). Only then, in Section 3, will I turn to the question of whether the semantic approach of the present section can be extended to double names in general.

2.1 **Restricted Names**

As far as Chemnitz is concerned, we may distinguish three phases: the time **before**, **while**, and **after** it was called **Karl-Marx-Stadt**. The middle phase will be referred to as the **S-Phase**; the other eras are **C-Phases**. The distinction concerns the history of the **official** names of the city of Chemnitz. An official name of a place is one that is used in legal documents, government publications and the like, and at a given time it is unique (as far as I know). However, as the case of Chemnitz shows, official names may change over time. Such changes take place instantaneously, usually at midnight or immediately after a declaration. Concomitant to the instantaneous change of an official name there is frequently a gradual change of speakers’ habits. For instance, once Chemnitz had been officially renamed in 1953, more and more people in that city started using the new name when referring to their hometown, thereby eventually replacing the old official name. And the habit spread through East Germany and other countries. Still, for whatever reasons, a few tra-
ditionalists may have continued to use the old name; as far as their referential habits were concerned, the official change was non-existent. After the German reunification history repeated itself, with the place names swapping their roles. And again the instantaneous change of official names was followed by a gradual change of people’s referring habits, leaving out a few hardliners who may continue to use the name Karl-Marx-Stadt (exploring linguistic imperialism) and those traditionalists who did not have to change (having seen it coming all along). However, the majority appears to have behaved conformistically, gradually switching from the ideology-loaded name back to the original one. In the present context, we may restrict attention to the conformists’ usage of the names in our time, i.e. long after the fall of the iron curtain in 1990. Neither linguistic change, nor socio-logical variation will be relevant to what I am saying, but it is important to realize that these phenomena exist, if only to not confuse them with the problems at stake.

Even if linguistic conformists like me have long stopped calling Chemnitz Karl-Marx-Stadt, we still understand and use sentences like (2). Hence although we do not normally call the city by its communist name, we have not stopped using that name altogether. We do use it, but only in particular connections. For instance, (2) means that Fritz never visited Chemnitz between 1953 and 1990, but only before 1953 or after 1990. Thus, intuitively speaking, the name Karl-Marx-Stadt is used to refer to Chemnitz during a certain time. What I will be concerned with is precisely this use of the name.

There are various ways of implementing the intuition that Karl-Marx-Stadt refers to Chemnitz at a certain time. I will only look into one of these accounts, not because they are all equivalent – they are not – nor because the account is more adequate than its contenders – it isn’t; but it is simpler, and the differences between the various possible analyses do not bear upon the present discussion. The idea is to interpret the name as a second-order property (or quantifier) that applies to properties that Chemnitz had during the S-Phase:

\[(3) \quad \text{Karl-Marx-Stadt as a restricted name}\]

A predication of the form Karl-Marx-Stadt \( P \) uttered in a context \((w_0, t_0)\) is true of a situation \((w, t)\) if and only if

(a) in \((w, t)\), Chemnitz has the property the predicate \( P \) expresses in \((w_0, t_0)\) and

(b) \( t \) is part of some S-Phase in \( w \).

A predication is not itself a sentence, but rather corresponds to an infinitival like Karl-Marx-Stadt be polluted. The analysis rests on the assumption that predications, like sentences, describe – or are true of – situations. Situations may be thought of as points in Logical Space, i.e. maximally specific possibilities. Following standard terminology, the situations described by an expression will be referred to as indices. (More precisely, the term ‘index’ relates to the rôle of being described by an expression.) For the present purposes one may take indices to be possible worlds \( w \) at times \( t \) and thus identify them with pairs \((w, t)\), as in (3).\(^6\)

(3) makes reference to the context of utterance. This is because in general the content of an expression may depend on the circumstances under which it is uttered, which is reflected

---

5 A more adequate analysis along the same lines would treat the restriction (3b) as a presupposition (rather than part of the assertion) and the name as a designator (rather than a quantifier). However, the differences only come out in certain environments that I have chosen to ignore here.

6 Note that, according to condition (3b), what the S-Phases are depends on the index world \( w \). This would have to be justified independently by looking at intensional constructions in which the index parameter is bound. I will ignore this complication here.
by the general format of semantic analyses. For ease of exposition I will assume that the context of utterance is itself a situation \((w_0, t_0)\).\(^7\) To see how (3) works, one may apply it to a simple sentence:

(4) In a context \((w_0, t_0)\), Fritz lived in Karl-Marx-Stadt is true of a situation \((w, t)\)
iff at some time \(t'\) that is both before the time of utterance \(t_0\) and part of an S-Phase, the relation of dwelling holds between Fritz and Chemnitz in \(w\).

One assumption behind (4) is that the main verb must be split into the lexical form live-, which expresses a relation between individuals and places, and the past morpheme -ed, which is interpreted as relating index time (the time \(t\) talked about) and context time (the time \(t_0\) at which the utterance is made) in the obvious way: \(t\) is before \(t_0\).\(^8\) Another assumption is that the past tense morpheme must be moved to a sentence-initial position in order to be interpreted; this may be motivated by the fact that, in a sentence like The king was away, the subject may (but need not) refer to whoever happened to be king at the time talked about, and not (or not necessarily) to whoever happens to be king at the time of utterance. Hence (4) can be derived from a fuller analysis along the lines of (4*), the details of which are left to the reader:

\[
(4^*) \quad [\text{Fritz lived in Karl-Marx-Stadt}]^{w_0,t_0}(w, t) \\
\iff (\exists t) [- \text{ed}]^{w_0,t_0}(w, t) (\lambda t [\text{Fritz live-in Karl-Marx-Stadt}]^{w_0,t_0}(w, t)) \\
\iff (\exists t) [ t < t_0 \land [\text{Fritz live-in Karl-Marx-Stadt}]^{w_0,t_0}(w, t) ] \\
\iff \ldots
\]

According to (4), the sentence Fritz lived in Karl-Marx-Stadt is true of a situation if that situation is set before the time of utterance, when Fritz used to live in Chemnitz while it was called Karl-Marx-Stadt. (4) says under which circumstances a certain sentence is true of a given situation. Under normal circumstances the situation described by an unembedded declarative sentence will be the context of utterance. Hence (4) also says under which (normal) circumstances an utterance of that sentence will be true tout court, viz. if it is true of the context of utterance \((w_0, t_0)\), which in turn is the case if, at some time \(t'\) that is both before the time of utterance \(t_0\) and part of an S-Phase, the relation of dwelling holds between Fritz and Chemnitz in \(w_0\). This, I take it, is an adequate truth-conditional account of the sentence, confirming the analysis (3) of the name Karl-Marx-Stadt. Further evidence comes from considering the present tense version of the same sentence:

(5) In a context \((w_0, t_0)\), Fritz lives in Karl-Marx-Stadt is true of a situation \((w, t)\)
iff at the time of utterance \(t_0\), which is part of an S-Phase, the relation of dwelling holds between Fritz and Chemnitz in \(w\).

(5) is quite analogous to (4). The only further assumption is that the present tense identifies context time and index time: \(t = t_0\). Again, the result of the analysis is summed up in (5), according to which the sentence can only be used truthfully during the S-Phase, i.e. at a time while Chemnitz is called Karl-Marx-Stadt. In particular, since the place no longer bears that name, it would not be true if uttered now. Again, this prediction is no doubt correct.

\(^7\) In his (1979) and (1980) David Lewis gave independent reasons why this identification should not be made. In the present context, it is harmless. I will, however, switch to more fine-grained contexts in the formal account in Section 4.5.

\(^8\) The treatment of tense as a relation between index and context times is somewhat inspired by Katz (2001) but largely simplified.
Asymmetry

Given the above analysis (3) of Karl-Marx-Stadt, one may expect Chemnitz to denote the same place during the C-Phases:

(6) **Chemnitz as a restricted name**

A predication of the form Chemnitz P uttered in a context \((w_0, t_0)\) is true of a situation \((w, t)\) iff
(a) in \((w, t)\), Chemnitz has the property the predicate \(P\) expresses in \((w_0, t_0)\) and
(b) \(t\) is part of some C-Phase.

Example (7) confirms this analysis:

(7) **Fritz lived in Karl-Marx-Stadt, not in Chemnitz.**

At least one plausible construal of (7) comes out on the analyses (3) and (6), once the ellipsis is expanded in the obvious way. According to it, (7) merely says that Fritz lived in Chemnitz during (part of) the S-Phase. This reading is given in (8), which can be derived as indicated in (8*):

(7) may also have a stronger reading, according to which Fritz did not live in Chemnitz at any time before 1953 or after 1990, and I suspect that it, too, can be derived using the above interpretations of the rules but a slightly different Logical Form. However, I do not want to go too deeply into subtleties of tense and ellipsis and thus leave the matter open. In any case, (6) cannot be the whole story about the name Chemnitz, as the paraphrase (9) of (7) shows:

(9) **Fritz lived in Chemnitz while it was called Karl-Marx-Stadt.**

In order to see what the above account of the name Chemnitz does to (9), let us assume that the underlined temporal modifier restricts the time talked about, the index time, to parts of S-Phases – after all, this is how the term S-Phase was introduced above – and that, moreover, the temporal conjunction while expresses that the main clause is true of a situation that is a temporal part of a situation of which the subordinate clause is true. Then, according to (6), (9) is analyzed as in (10) – or, more explicitly, (10*):

---

9 The difference between the two readings lies in the variable scope of the existential closure operation associated with the tense morpheme. Cf. Katz (2001) for more on this.
(10) \[ [\text{Fritz lived in Chemnitz, while it was called KMS}]^{w_0,t_0}(w,t) \]

\[ \iff \quad (\exists t) \left[ \text{while}^{w_0,t_0}(w,t) [\lambda t \, [\text{KMS}]^{w_0,t_0}(w,t)] \right] \]

\[ \lambda t \, [\text{it be-called KMS}]^{w_0,t_0}(w,t) \right) \]

\[ \iff \quad (\exists t) \left[ (\exists t') \left[ t' \subseteq t \land t' < t_0 \land [\text{it be-called KMS}]^{w_0,t_0}(w,t') \right] \right] \]

\[ [\lambda t \, [t < t_0 \land [\text{Fritz live-in Chemnitz,}]^{w_0,t_0}(w,t)] \right) (t) \]

\[ \iff \quad (\exists t_k) (\exists t_c) (\exists t) \left[ \text{S-Phase}(w,t_k) \land \text{C-Phase}(w,t_c) \land t < t_0 \land t' \subseteq t_k \land t \subseteq t_c \land L_{w,t}(f,c) \right] \]

Since the S-Phase does not overlap the C-Phases (in our world), no time satisfying the underlined condition exists, so that (9) would be bound to be false, no matter where Fritz used to live. Surely, this is an unwelcome result; but it is obvious what’s going on here. In (9), Chemnitz is used as a name of the town as such, independently of what it is called. In fact, in the meta-language, I have been using the name in that sense all along (as the reader may have noticed). Hence apart from the reading given in (6), the name should also receive a more standard interpretation:

(11) \textbf{Chemnitz as a neutral name}

A predication of the form Chemnitz \( P \) uttered in a context \( (w,t_0) \) is true of a situation \( (w,t) \) iff

- in \( (w,t) \) Chemnitz has the property the predicate \( P \) expresses in \( (w_0,t_0) \).

According to (11), the name \textbf{Chemnitz} may be combined with any predicate whatsoever, resulting in a predication attributing the property expressed by that predicate to the city of Chemnitz. Under this analysis, the name \textbf{Chemnitz} does little more than designating the object Chemnitz.\(^{10}\)

One may suspect that (11) is all there is to the name \textbf{Chemnitz} in that it covers the more specific uses when the name is combined with predicates that are true of Chemnitz during the C-Phases. However, the interpretation given in (6) is still needed as a separate reading. Otherwise (7) would come out contradictory, as the reader may verify. Since (7) may well be true even, the name \textbf{Chemnitz} must be ambiguous between a \textit{restricted} reading (6) and a \textit{neutral} reading (11). But the name \textbf{Karl-Marx-Stadt} does not show this kind of ambiguity, witness the oddness of (12):\(^{11}\)

(12) \textbf{Fritz lived in Karl-Marx-Stadt while it was [still] called Chemnitz.}

In fact, the name \textbf{Karl-Marx-Stadt} invariably appears odd when relating to a time that is not entirely part of the S-Phase. In other words, the following pattern emerges, where the upper and lower judgement marks ‘#’ [‘odd’] and ‘✔’ [‘ok’] relate to the Chemnitz and Karl-Marx-Stadt variants, respectively:

\[ (11) \text{is essentially Montague’s (1970) typed-shifted account of proper names as rigid designators. I am making use of it because it minimally deviates from the analyses (3) and (6).} \]

\[ (12) \text{Of course, (12) is only odd in the (present-day) conformists’ dialect, which is the one I am looking at.} \]

It thus turns out that Chemnitz, having a neutral sense, is always fine (in positive contexts), whereas Karl-Marx-Stadt is inherently restricted. This is the earlier-mentioned asymmetry between the two names. Given the above analyses, the easiest way to account for it is to treat the name Chemnitz as ambiguous between a neutral and a restricted reading. Of course, this ambiguity cannot be the kind of accidental homonymy exemplified by words like bank or can. For the case at hand is not an isolated one. Very often when a place changed its name in history, the older name(s) remain in currency – though not as neutral names for the place as such but only as restricted names of the place in olden times; and in these cases the ordinary name may also be used in a contrastive sense, restricted to the bearer in modern times. Byzantium, Constantinople, and Istanbul all refer to the same city, but only the third one is its proper neutral name in present-day English. So the ambiguity must be systematic, to be explained in general semantic or pragmatic terms. When a place changes its official name, only one of its names remains or becomes the one that has a neutral reading. Changing place names, then, confirm rather than refute naive prejudice.12

3. Double lives I: Semantic Accounts

Example (2) illustrates that truth-preserving substitution of co-designative names in extensional contexts may be blocked. Since the distinction between neutral and restricted names offers a semantic account of this phenomenon, it is tempting to try and adapt it to apparently parallel cases like (1) that do not involve changing place names.

3.1 Restricted Names

Let us first generalize the crucial ingredients of the above analysis. A neutral name is just a name. It rigidly designates one particular object; it can be used to refer to that object under all kinds of circumstances. And even though one may represent it by a quantifier, this type shift is only made for convenience and does not change the designating function of the name. The general format of a neutral name is given in:

\[ \text{Fritz lived in } \{ \text{Chemnitz} \} \text{ from } \{ 1840 \text{ till } 1860 \} \{ 1940 \text{ till } 1960 \} \{ 1940 \text{ till } 1999 \} \{ 1960 \text{ till } 1980 \} \{ 1960 \text{ till } 1999 \} \{ 1994 \text{ till } 1999 \}. \]

\[ 12 \text{ The following variation on (14), due to an anonymous referee, appear to cast doubt on the generality of the observation:} \]

(14') Fritz continued to live in Karl-Marx-Stadt even after its name was changed to Chemnitz.

However, a comparison with (14") shows that the alleged neutral use of the name Karl-Marx-Stadt is licensed by the presupposition triggered by the verb continue:

(14") Fritz continued to love his wife even after they had been divorced.

For (14") to be true of a time \( t \), the person loved by Fritz obviously does not have to satisfy the description his wife at \( t \), but only at a time before \( t \) at which the presupposition holds that Fritz used to love her. Similarly, for (14') to be true of a time \( t \), the town lived in by Fritz does not have to satisfy the description (3) at \( t \), but only at a time before \( t \) at which the presupposition holds that Fritz used to live there.
Semantic behaviour of neutral names (Montagovian version)

If \( N \) is a neutral name, there is some (uniquely identifiable) individual \( x \) (the bearer of \( N \)) such that any predication \( N P \) uttered in a context \((w_0,t_0)\), is true of a situation \((w,t)\) iff

- \( \text{in } (w,t), x \) has the property expressed by the predicate \( P \) in \((w_0,t_0)\)

\((NN)\) is a straightforward generalization of the interpretation (11) of Chemnitz as a neutral name; I only replaced the word Chemnitz by an arbitrary neutral name \( N \) and its bearer, the city of Chemnitz, by the bearer of \( N \). According to \((NN)\), combinations of neutral names and predicates express that the bearer of the name has the property expressed by the predicate. As mentioned above, this is the standard semantics of proper names.

In order to arrive at a general notion of a restricted name, we need to generalize (3) and (6). Both analyses are conjunctions of two conditions (a) and (b), where (a) is precisely as the condition • in the interpretation of a neutral name and (b) is some extra condition relating to official names. The most obvious generalization is to replace (b) by something about official names in general. Since the aim is to adapt the analysis of place names to double identities, this won’t do: for one thing, it is not always clear what the official name of a person is – just think of Superman; for another thing, even if one of the names of a person is clearly the official one, it need not be the current one, the one that is used as a neutral name to refer to that person – Mark Twain being a case in point: present-day speakers normally refer to that author by his \textit{nom de plume}, whether or not they know that his official name was Samuel Langhorne Clemens. So if the distinction between restricted vs. neutral names is to carry over to cases of double identity, condition (b) must be allowed to bring in other aspects than official usages of names. Before discussing which aspects these may, let me give a very general version of the interpretation of restricted names, one that simply adds an arbitrary condition to the predication (a):

\((RN)\) \( N \) is a restricted name iff \( N \) is not a neutral name and there is some (uniquely identifiable) individual \( x \) (the bearer of \( N \)) and a condition \( \phi \) such that any predication of the form \( N P \) uttered in a context \((w_0,t_0)\) is true of a situation \((w,t)\) iff

(a) \( \text{in } (w,t), x \) has the property the predicate \( P \) expresses in \((w_0,t_0)\) and

(b) \( \phi \)

According to \((RN)\), a restricted name is like a neutral name in that it refers to its bearer to which a predicate may attribute a property; but unlike a neutral name, it makes this predication subject to an additional condition \( \phi \). The nature of that condition is left maximally open in that satisfaction of \( \phi \) may depend on (i) the context of utterance, (ii) the time and world talked about, (iii) the bearer of the name, and even (iv) the property expressed by the predicate. For the time being one may think of \( \phi \) as expressing a property the bearer must have at the index, thus only showing dependencies (ii) and (iii); for instance, in the case of Karl-Marx-Stadt, the relevant property could be being officially called Karl-Marx-Stadt. However, in Section 3.3 we will see reasons to consider a more liberal notion of condition \( \phi \), one that does not correspond to a property of the bearer.

3.2 The Apparel Oft Proclaims the Man

If (1) and (2) are as parallel as Saul and intuition suggest, the non-substitutivity ought to have a common source. Hence, following the analysis of the preceding section, either Superman or Clark Kent or both would have to be restricted names. The tricky question is not so much which of them is, but rather what the relevant condition might be. A simple answer is suggested by the paraphrase (15) of (1), due to Forbes (1997: 111):
Clark, so-attired, entered the phone booth and Superman, so-attired, emerged.

Forbes not only came up with a paraphrase, he also gave a logical analysis, which I am repeating in (16) with slight notational changes:  

\[
\text{entered}(\text{Clark}, b) \land (\alpha: L(\alpha, \text{so}) \land \text{attired}(\text{he}, \text{Clark}, \alpha)) \\
\text{then}(\text{emerged}(\text{Superman}, j) \land (\alpha: L(\alpha, \text{so}) \land \text{attired}(\text{he}, \text{Superman}, \alpha)))
\]

Following this analysis, the condition so-attired in (15) relates an individual – Superman in both cases – and the object Labelled by the referent of so. According to (16), the first so in (15) is construed as referring to the subject of the first conjunct, i.e. the name Clark; and this name can be understood as Labelling a certain way of dressing, an appearance. Since, on the same analysis, the second so refers to a different name, viz. Superman, the second occurrence of so-attired expresses a different condition, provided that that name Labels a different kind of attire. The idea is, of course, that Clark does Label a different kind of attire, viz. a business suit etc., whereas Superman Labels a certain kind of overall.

Forbes’s analysis can be cast in the present framework of restricted names. At first glance, this may appear impossible, because (16) makes reference to an apparatus of demonstratives, which does not seem to play any role in the purely semantic account (RN). However, I think the difference is cosmetic. After all, the original sentence to be analyzed, viz. Saul’s (1), does not contain any of the demonstratives on which the paraphrase (15) and its analysis (16) rely. So, given that we are after a systematic account of how the original sentence is understood, we would have to provide a source for it, either from within the sentence, or from the context of utterance, or both. Now, the only source for the crucial addition of so-attired that I can think of are the two names themselves; for if they were replaced by descriptions or ordinary names, no such paraphrase seems to be called for, or even suitable. But if the two names are the sources of the critical material, we may as well enrich their interpretation with it, thus arriving at the semantic reformulation (17) of Forbes’s indexical account:

\[
\text{Clark} [\text{Kent}] \text{ as a restricted name (attire version)}
\]

A predication of the form Clark P uttered in a context \((w_0, t_0)\) is true of a situation \((w, t)\) iff

(a) in \((w, t)\) Superman has the property the predicate \(P\) expresses in \((w_0, t_0)\) and

(b) Superman is wearing Kentian attire in \((w, t)\).

3.3 Personae

As Saul (1997b: 116) pointed out in her reply to Forbes, one problem with this particular proposal is that attires are not enough:

\[
\text{Lois slept with Superman before she slept with Clark Kent.}
\]

\[
\text{Lois slept with Superman, so-attired, before she slept with Clark Kent, so-attired.}
\]

(18) cannot truthfully describe a situation in which Lois Lane, having saved herself for the hero for so long, finally starts an affair with Clark Kent, and …

’[...] due to her long-standing fascination with the Man of Steel, asks Clark to put on a little cape, take off those spectacles, etc.’

Saul (1997b: 116)

It thus seems that the conditions \(\phi\) to be associated with the restricted names Superman and Clark Kent must be somewhat less superficial than (17b), marking a difference in

\[\text{The co-indexing in (16) is meant to relate (occurrences of) expressions and not their referents. Forbes uses arrows leading from occurrences of so to the relevant occurrence of the corresponding name.}\]
character rather than in outward appearance. However, as Graeme Forbes (1999) pointed out, it is not enough to trade clothes for characters and just have the names label personae instead of ways of dressing; for Superman may express his double personality simultaneously, by simultaneously sitting in a bar in his fancy costume and talking shop to Lois Lane over the phone. What is missing, then, is a connection between the individual predications in (1) and the personae (to use Forbes’s term) indicated by the names. In other words, the truth condition of the first conjunct must make sure that the property of entering the phone booth is attributed to Superman as Clark Kent; similarly, the second conjunct would have to express that he has the property of leaving the phone booth as Superman. So, although in a (typical) situation as described by (1), Superman has both properties, he has one of them as Clark Kent and the other one as Superman. Referring to properties that Superman has as Clark Kent as C-Properties, we may replace (17) by:

(20) Clark [Kent] as a restricted name (adverbial version)

A predication of the form Clark P uttered in a context (w₀,t₀) is true of a situation (w,t) iff
(a) in (w,t) Superman has the property Q that the predicate P expresses in (w₀,t₀) and
(b) Q is a C-Property in (w,t).

It is important for (20) to work properly that being a C-Property is itself a situation-dependent (second-order) property. In other words, what is a C-Property in one situation may not be a C-Property in another situation: though entering a call box is a C-Property in a situation as described in (1), it is not necessarily one in a situation described in (20). Following this lead, it is easy to give a similar analysis of Superman. Moreover, like Chemnitz, Superman may also be used as a neutral name. Not only have I been using it like that here, but this is how the name is generally used by those who are in the know about the identity of its bearer (and for the time being I will ignore the others). Examples (21) and (22) suggest this:

(21) Superman frequently disguises as Clark Kent.
(22) Clark Kent frequently disguises as Superman.

Whereas (21) is a perfectly normal, if quite obvious, thing to say, (22) is somewhat odd. And this is so because in (21) Superman may be understood as a name of the person as such, whereas the subject in (22) does not allow for such a reading. Unlike Clark Kent but like Chemnitz, then, Superman appears to call for both a neutral interpretation as a rigid designator and a restricted interpretation along the lines of (20), trading C-Properties for S-Properties.

What makes a property a C-Property, or an S-Property, in a given situation? To attack this question, let us look at some specific examples first. In (1), the answer is pretty straightforward: Superman has the property of entering the phone booth due to his performing a certain kind of action (entering the phone booth), and while doing so, he is wearing his Kentian attire. However, the phone call from the bar shows that matters are not always that simple: though he is wearing his Superman costume, he is talking to Lois Lane as Clark Kent. Hence, in that scene, the property of making a phone call to Lois Lane as Clark Kent. Hence, in that scene, the property of making a phone call to Lois Lane as Clark Kent. Hence, in that scene, the property of making a phone call to Lois Lane as Superman, whereas sitting in the bar is an S-Property. We may even imagine him making two phone calls at the same time, by uttering I can’t talk now, but I will get back to you tomorrow into two portables one of which is connected to Lois, who is waiting for Clark’s call, while on the other phone line there is Lex Luthor expecting a call from his arch-enemy. In that case, the property of talking to Lois Lane is a C-Property, while the property of talking to Lex is an S-Property. Since the two properties are realized by two different
speech acts performed by the same person with one articulatory act, the difference between them may be captured by a Davidsonian account of adverbial modification as conditions on events (in a broad sense), as indeed Graeme Forbes (1999) proposed. However, (20) also allows for distinctions that may not manifest themselves in eventualities, and I suspect that this freedom is sometimes needed. (23) are cases in point:

(23) Clark Kent is Superman.
(24) Clark Kent admires Superman.

(23) is unequivocally true. In terms of the distinction between Superman’s two kinds of properties this means that being Superman must be a C-Property; but the predication made in (23) does not relate the bearer of the subject name to an eventuality. One may take this to be an indication that the subject in (23) should not be construed along the lines of (20). But then, other things being equal, a unified analysis of the name Clark Kent is certainly to be preferred to one based on a case distinction. Moreover, while (23) might be analyzed as a metalinguistic predication, no such interpretation offers itself for (24), which may be used to describe the kind of character displayed by Superman performing his Kent act.

What, then, are the criteria by which a given property should be judged a C-Property in a given situation? A simple-minded answer to this question employs a descriptive construal of names as abbreviating descriptions, or as bundling properties, of their bearers: C-properties are those that are part of the sense, or the descriptive content, of the name. I am not calling this answer ‘simple-minded’ because I think that names should not be associated with properties but because the properties associated with the name Clark Kent simply cannot be the C-properties referred to in (20). As a case in point, take the property of entering the phone booth, which, as we have seen, is among the C-properties in certain situations. Yet, clearly, entering a phone booth is not among the properties implied by the sense of the name Clark Kent: otherwise speakers using the name in that sense would have to believe that its bearer is someone who is constantly entering phone booths. Neither is the property of entering the phone booth in a particular situation part of the sense of that name; for it is not part of any speaker’s understanding of that name that Superman is entering a phone booth on a particular occasion, however he may be dressed for it.

3.4 More Asymmetry
In order to motivate and formulate the alternative pragmatic analysis of non-substitutivity, let me discuss one more example, one that brings out more clearly what I take to be the point. A few years ago, I was looking for a paperback edition of No Night is Too Long, a piece of crime fiction written by the British author Ruth Rendell. One thing you have to know about Ruth Rendell in order to understand the example is that she writes under two names, viz. Ruth Rendell and Barbara Vine. It so happens that neither of these names is the author’s full official name (cf. Section 4.2 below), but the fact that they refer to one and the same person was never hidden from the reading public. However, as far as the latter are concerned, there is an asymmetry between the author’s two names. For a long time, starting in 1964, she only published books under the name of Ruth Rendell, and only from 1986 onward did she also use the name Barbara Vine for some of her novels, the difference in name supposedly coinciding with a difference in the styles of writing. The book I had been looking for back in 1995 was the latest novel written under the name Barbara Vine at the time. Someone had recommended a bookstore specialised in crime fiction to me, and entering the store I enquired about the book (which they did not have, but that is another story) and immediately found myself engaged in a conversation with the owner of the shop.

14 Forbes (1999) still seems to prefer an analysis of (1) in terms of attires to a more unified account in terms of personae.

15 The example originated in Zimmermann (1996), where I defended a semantic solution along the lines of the previous subsection. I had not been aware of Saul’s work at the time.
a devoted connoisseur of British crime fiction. One of the things he told me was (25), adding that he liked all of these novels – except the most recent one, which was the one I had been looking for.

(25) **I’ve read all of Barbara Vine’s books.**

It was obvious that he was not talking about all of Ruth Rendell’s books, even though we were both clearly aware of the fact that Barbara Vine and Ruth Rendell are the same person. Rather, the utterance was unmistakably about those books that the British queen of crime wrote under the name of **Barbara Vine**. Why is that so? I think that again the asymmetry between the two names comes into play here. For if, instead of (25), the bookseller had uttered (26), I may have concluded that the bookseller meant *all* of the books written by Ruth Rendell, no matter under which name they had been published:

(26) **I’ve read all of Ruth Rendell’s books.**

On a straightforward construal of (26), **Ruth Rendell** is a neutral name, whereas **Barbara Vine** was not used neutrally in (25). Moreover, although the bookseller could have used both names in a restricted sense as in (27), no neutral reading of **Barbara Vine** seems to have been available to him:

(27) **I’ve read all of Ruth Rendell’s books, though not all of Barbara Vine’s.**

The difference between the two names in (27) is thus reminiscent of the asymmetries observed above. And while could be cast in a semantic account based on a distinction between Ruth Rendell’s *R*-Properties and her *V*-Properties, I would like to explore an alternative, pragmatic approach.

4. **Double lives II: A Pragmatic Approach**

4.1 **Neutrality**

Unlike other pragmatic explanations of substitution puzzles, the approach to be presented does not depend on contrastive uses as in (27) but also works if what looks like a restricted name comes on its own, as in (25). However, strictly speaking there are no restricted names:

(NP) **Neutrality Principle**

All names are neutral.

Neutrality is a *semantic* notion concerning the literal meaning of proper names: a neutral name is one that merely presents its bearer, not adding any descriptive information to the content of expressions in which it occurs. And since the semantic function of a neutral name does not go beyond specifying its bearer, any two names with the same bearer will also have the same literal meaning. Consequently, they may appear to be always substitutable *salvo sensu*; but they are not, as we will see.

---

16 In that case the examples (25) – (27) must not be construed as predications. Otherwise a treatment of the two names along the above lines would assign to (25) an unattested reading according to which being an author all whose books the speaker has read, is one of Ruth Rendell’s *V*-Properties – and hence one of her properties; but then the speaker would have to read all of her books (written under whatever name) to make (25) true.

17 Cf. Barber (2000), where a principle similar to (M) in fn. 2 is invoked to arrive at roughly the interpretation given in Section 4.3. The problem with (M) is that it only applies to utterances of two coreferential names in the same context.
Like all linguistic expressions, names are language-dependent. What may be a name of your language, need not be one in mine. Consequently, since languages are (or at least determine) form-meaning mappings, there will hardly be any two speakers of the same language. On the other hand, names may be shared between languages and if they are, they also share their referents. These truisms, which will be given a (partial) formal reconstruction in Section 4.5, are crucial to the pragmatic account of non-substitutivity of coreferential names. For the time being it is important to keep in mind that a name is always a name in some language and that languages may differ in the names they contain.

4.2 A Failure of Uniqueness

Let me give the naive prejudice about names a more official form:

(U) Principle of Uniqueness
No two names of the same language have the same bearer.

(U) is not intended as an empirical generalization. As such it would not only be false but also miss the point. It would be false because there are languages containing two (neutral) names of the same bearer – even though, as I will argue, it takes more to reject (U) than two coreferential names within one utterance. As an empirical generalization, (U) misses the point in that the naive prejudice only concerns speakers’ expectations about language whereas the empirical correctness of (U) partly depends on extra-linguistic facts: speakers may be convinced that their language satisfies (U) and they may still be wrong about these facts. In other words, there cannot be two distinct names that a speaker knows to corefer. So (U) should be read as an ideal that speakers avow to for whatever reason, although they may fail to live up to it. The following formulation captures this subjective construal of (U):

(UP) Uniqueness Presumption
Speakers presume that (U) holds and that speakers presume that it holds.

Bringing in the notion of a name, both (U) and (UP) are language-dependent. According to (UP), a language can only contain two distinct, coreferential names as long as its speakers are not aware of this fact. Once they are, they will drop one of the names, thus becoming speakers of a different language, one that contains fewer names. From then on, any utterances of coreferential names must either be instances of code-switching or else deviations from ordinary usage to be explained by pragmatic strategies of reinterpretation. The latter will be addressed in Section 4.3.

To see how (U) and (UP) square with the examples discussed above, let us start with Chemnitz. Some fifty years ago, the semantic situation was simple. If you wished to name the place, you had no choice: Chemnitz was the only name in town. Clearly, in the language of [Olden] spoken at the time, Chemnitz was no counter-example to (U). Then came the time when the Chemnitzers were split into two linguistic communities. The first of these continued to speak [Olden] whereas in the language [Newspeak] spoken by the major-

---

18 The concept of a name must be taken as pre-semantic lest the Neutrality Principle be circular. Semantically, a name is a (lexical) expression the meaning of which is neutral in the sense to be made more precise in the next subsection. Pre-semantically, a name is a linguistic expression that is conventionally associated with an object (its bearer), the relevant conventions being partly non-linguistic. I do not have anything to say about the nature of these conventions, but I trust there is some way of distinguishing names from other linguistic expressions.

19 In order to avoid the impression that (U) can be justified by recourse to rationality principles underlying cooperative communication, I have refrained from formulating it as an imperative à la Grice (1975). I would like to stay uncommitted on this question.
ity, Chemnitz was now called Karl-Marx-Stadt. Presumably, the two communities were not entirely stable. Apart from natural turnover, there may have been some fluctuation due to changes of political camps, convenience, etc. Moreover, most (if not all) Chemnitzers were bilingual and quite capable of following and conducting conversations in either language. All in all, it may be assumed that the two communities led a peaceful coexistence in Chemnitz – and elsewhere, where the same split could be observed. In fact, the split still exists, though after the reunification of Germany the number of $\mathfrak{N}$-speakers has dramatically shrunk. More importantly, however, both languages$^{20}$ conform to $(U)$, by containing either Chemnitz or Karl-Marx-Stadt, but not both, as a (neutral) name for Chemnitz.

Being bilingual, sometimes speakers of one language would make utterances in the other language. In particular, as long as $\mathfrak{N}$ was the official dialect of the political authorities, an $\Theta$-speaker may have preferred to adapt herself to it, given the right kind of circumstances; at the same time an $\mathfrak{N}$-speaker may have mocked his neighbour by making an utterance in $\Theta$. So sometimes $\Theta$-speakers speak $\mathfrak{N}$ without becoming (permanent) $\mathfrak{N}$-speakers, and vice versa. The code-switching may occur instantaneous and it may trigger pragmatic side-effects. For instance, speakers may out themselves, or misidentify themselves, as members of a particular community by making an utterance in that community’s language; or they may indicate their sympathy, their dislike, or their profound knowledge of the other community by using their language in the right kind of situation. None of this is too exciting; nor is it characteristic of the Chemnitz situation. But I think it can help explain away some apparent counter-evidence to $(U)$. Still, it is hard to see how $(2)$ – repeated here for the readers’ convenience – could be construed as an instance of code switching from $\Theta$ to $\mathfrak{N}$ or back.$^{21}$

$(2)$ I never made it to Karl-Marx-Stadt, but I visited Chemnitz last year.

It seems that something other than, or maybe something on top of, code-switching is going on here. Precisely what this is, will be elaborated in Section 4.4.

In Superman’s case, time is less important than knowledge. As has been emphasized before,$^{22}$ we must distinguish between those ignorant speakers of a language $\mathfrak{I}$ who believe that there are two distinct persons by the respective names of Superman and Clark Kent, and us enlightened speakers of $\mathfrak{E}$ who know that the two names refer to the same person. It is cases like $\mathfrak{I}$ for which a subjective construal of $(U)$ becomes important: after all, the language does contain two distinct names for the same person. But not only does this escape its users; in line with $(UP)$, they even assume that Superman and Clark Kent are the names of two different persons. The situation with $\mathfrak{E}$ is different. I take it those who know the ‘facts’ about the comic paper hero’s double life, call him Superman, which is his neutral name in $\mathfrak{E}$; and according to $(U)$, Superman ought to have no other name in our language. Hence uses of Clark Kent by speakers of $\mathfrak{E}$ must be either ungrammatical or else instances of code-switching. The latter uses are extremely rare, simply because normally people get acquainted with Superman by reading the comics, watching the

$^{20}$ Speaking of two languages is an oversimplification: other coreferential names will cause other dialect splits, so that the total number of languages may well be equal to, or even exceed, the number of speakers.

$^{21}$ Grame Forbes (p.c.) pointed out to me that there are non-temporal ways of interpreting (2), which I suspect might be explained along these lines; but then as far as I can see, such construals are clearly distinct from the temporal readings discussed in Section 2 above.

$^{22}$ Moore (1999); the linguistic turn is mine though. Note that, like anybody in the debate initiated by Saul (1997a), I am pretending that the Superman myth is fact rather than fiction. As to Superman and Clark Kent as names of fictional characters, I can only refer to Parsons (1980).
movies, or listening to an enlightened account, thereby becoming enlightened themselves; so there are no Ignorami and, consequently, it seems that there is hardly any need for speaking their language. Still, as will become clear from the discussion in the next two subsections, code-switching from Е to І does occur precisely in the cases discussed by Saul - though, again, this is not the full story.\textsuperscript{23}

The case of Ruth Rendell and Barbara Vine is slightly more involved. Neither of these names is a genuine nom de plume. The author's maiden name is Vine, Rendell is her married name, and she has two Christian names, Barbara and Ruth, which are apparently enjoying equal currency among her family and personal acquaintances - but:\textsuperscript{24}

I tend to divide friends and relatives into the “Ruth people” and the “Barbara people”. Both names are equally familiar to me, equally "my" names. [...] And I don't mind which I am called so long as people don't try to change in, so to speak, midstream. There is for me something grotesque in a Barbara person trying to become a Ruth person, or vice-versa. [signed ‘RR’]

These remarks about unmotivated code-switching are certainly in line with (U) (and may well have served as a source of inspiration). However, for the present purposes I will concentrate on the community of crime fiction readers, among whom I will also ignore those who do not know Ruth Rendell under either, or only one, of her names. The rest of her readers is divided into the Еxperts who know that both names refer to the same person, and the Laymen who do not. As in the Superman case, there is a clear asymmetry: Еxperts invariably use Ruth Rendell to neutrally refer to the author, preserving Barbara Vine for special occasions, the above interchange between me and the bookseller being a case in point; further examples can be gleaned from the secondary literature.

Now for the Laymen. In line with a certain epistemological tradition, it would be natural to characterize them, or at least some of them, as being undecided as to whether the two names corefer: in some worlds compatible with a typical Layman’s perspective they would, in others they would not.\textsuperscript{25} I will not follow this tradition here, but instead claim that even someone who has read, say, A Fatal Inversion (by Barbara Vine) and The Lake of Darkness (by Ruth Rendell) and discovered certain similarities in style, plot, or characters without being aware that they have been written by the same person, will take it that they have not – simply because the covers give different author names. This claim is somewhat hard to corroborate: after all, doubts about identity may always be elicited, given strict enough standards of knowledge and belief.\textsuperscript{26} But even in the presence of doubt, the distinctness assumption may prevail until the Layman becomes an Еxpert. According to (UP), this is what should happen. And in the interest of simplicity and transparency, I will work on the assumption that it does. For although it will turn out that the pragmatic reinterpretation strategy only depends on a Uniqueness Presumption made by Еxperts

\textsuperscript{23} As regards Е and І, the same caveats apply as in the Chemnitz/Karl-Marx-Stadt case (cf. fn. 20). Moreover, it should be noted that І is the language spoken in the fictitious world(s) of the comic strip, whereas Е is common in the actual world.

\textsuperscript{24} The quotation is taken from an announcement of the first Barbara Vine novel, A Dark Adapted Eye, (Viking 1986); the announcement, which has been published in various places, can also be found on the internet under www.gusworld.com.au/books/vine/why.htm.

\textsuperscript{25} Cf. Stalnaker (1976: 89): ‘Let us consider what happens when a person comes to know that Hesperus is identical to Phosphorus after first being in doubt about it. If the possible-world analysis of knowledge is right, then one ought to be able to understand this change in the person’s state of knowledge as the elimination of certain epistemically possible worlds.’ [Emphasis added]

\textsuperscript{26} See Lewis (1996) for the general mechanism.
about Laymen, it is hard to see why they should make this presumption if it were not for its apparent universality. I thus take it that what distinguishes the Layman from the Expert is his being positive that Ruth Rendell and Barbara Vine refer to different persons. In particular, then, both the Layman and the Expert conform to \((UP)\).

According to \((UP)\), coreferential names are accidents that only happen as long as speakers do not notice them. As soon as the Layman finds out about the true identity of the author he used to call Barbara Vine, he becomes an Expert and erases that name from his language. Hence in his new language \(\mathcal{E}\) there is no room for sentences like (25) or (27) above and repeated below:

\[
(25) \quad \text{I've read all of Barbara Vine's books.}
\]
\[
(27) \quad \text{I've read all of Ruth Rendell's books, though not all of Barbara Vine's.}
\]

In other words, (25) and (27) are not interpretable as utterances in Expert talk. So they must be instances of code-switching, or else they must be construed somewhat differently. In the next subsection, I will argue that the latter is the case.

4.3 A Counterfactual Reading

Though what the bookseller told me did not make any literal sense, I was still able to figure out what he meant. I think this was mainly because I recognized (25) as an utterance in a Layman’s language; Experts do not use the name Barbara Vine to neutrally refer to the author of A Fatal Inversion. So my first inclination might have been to interpret the bookseller’s utterance simply as an utterance in a foreign language, \(\mathcal{L}\). However, even in \(\mathcal{L}\) the speaker would have said that he had read all of Ruth Rendell’s books — whichever name appeared on their covers: as a sentence of \(\mathcal{L}\), (25) is true in that situation iff the bookseller had read all of the books written by the bearer of the name Barbara Vine, i.e. Ruth Rendell. A mere shift of language would not have led me to the right interpretation. What did, I will now argue, was a shift in perspective: instead of taking the sentence uttered as a piece of Laytalk, it rather seems that I interpreted the utterance as if it had been made by a Layman, while at the same time realizing that it was not. Let me explain.

It helps to first take a look at a Layman’s perspective on his own language. I am assuming that he conforms to \((UP)\), but does not realize that \(\mathcal{L}\) fails to satisfy \((U)\). How can this be? After all, according to (38), any two neutral names of the same individual are synonymous. Consequently, knowing \(\mathcal{L}\), the Layman ought to be able to see that Ruth Rendell and Barbara Vine have the same meaning. More generally, if the speakers know their language and thus the meaning of each word, they can tell whether there are any pairs of distinct coreferential names. In other words, if a language fails to meet \((U)\), it would seem that its speakers, knowing their language, ought to be aware of this. However, the Layman is not; and according to \((UP)\), speakers in general are never aware of any failure of \((U)\). One may conclude that speakers do not always know their language. 27

The kind of linguistic ignorance that is at stake here is not at all uncommon. In fact, whether or not \((UP)\) holds, ordinary speakers arguably never fully grasp the meanings of (neutral) names as interpreted in \((NN)\)28 and a fortiori practically never fully know the

\[\text{This is not the only escape from this Fregean dilemma, of course. In an earlier version of this paper I opted for giving up \((NN)\) instead, replacing it by a descriptive account of names along the lines of Lerner & Zimmermann (1991) and Haas-Spohn (1995: ch. 4). Though I am not aware of any principled obstacles to this strategy, I now think the more natural option is to follow Kaplan (1989a: 562) and Stalnaker (2001) and treat referential uncertainties about proper names as lack of linguistic knowledge.}\]

\[\text{— on an internalist analysis of grasping, that is. Cf. Frege (1892: 27), or Lewis (1981).}\]
language they actually speak. And whether or not \((NN)\) holds, speakers may be uncertain about the exact meaning of an expression like **identical twins** although they do use it as part of their language. This tension between what an expression actually means and what a speaker takes it to mean points to an important distinction between two senses of what a speaker’s language is. His **objective** language is the one the speaker actually speaks. Its expressions receive their meanings by social convention, and the speaker may or may not know which meaning is associated with which expression. A speaker’s **subjective** language\(^{29}\) is a language the speaker takes himself to speak. Since, as we have just seen, speakers may be uncertain as to what precisely their language is, they will in general have more than one subjective language in the sense that there are several possibilities of matching form and meaning none of which the speaker can exclude to be in full accord with her (objective) language. For instance, if I am uncertain as to the exact meaning of **identical twins**, then in one of my subjective languages it may be paraphrasable by **twins emerging from the same egg**, in another one it would mean the same as **twins emerging from the same sperm** (though I may exclude that it is synonymous with **twins of the same sex**).

As a matter of historic fact, in \(\mathcal{L}\) *Ruth Rendell* and *Barbara Vine* both refer to Ruth Rendell. Hence the \(\mathcal{L}\)ayman’s objective language \(\mathcal{L}\) obviously violates \((U)\). On the other hand, his subjective languages do not; for the \(\mathcal{L}\)ayman is devoted to \((UP)\) (as I continue to assume).\(^{30}\) An immediate consequence of this observation is that \(\mathcal{L}\) cannot be among the \(\mathcal{L}\)ayman’s subjective languages. This loss of objectivity among the subjective alternatives indicates that the \(\mathcal{L}\)ayman has an erroneous belief about his (objective) language, viz. that it satisfy \((U)\). Concomitant to this linguistic error are false beliefs about extra-linguistic matters, like the conviction that *A Fatal Inversion* and *The Lake of Darkness* (both of which may be identified by plot rather than title) have been written by two different women. This is why a \(\mathcal{L}\)ayman who wants to express that he has read all of the books published under the author name *Barbara Vine*, would utter (25). Given that he is a speaker of \(\mathcal{L}\), where *Barbara Vine* refers to Ruth Rendell, objectively he would therefore express that he has read (at least) all of the novels written under the author name *Barbara Vine*. However, subjectively he would be under the impression to have only expressed that he has read (at least) all of Ruth Rendell’s books – including *The Lake of Darkness*. However, subjectively he would be under the impression to have only expressed that he has read (at least) all of the novels written under the author name *Barbara Vine*. Moreover, and somewhat paradoxically, the discrepancy between his subjective understanding and the meaning objectively expressed need not thwart his communicative intents: if the hearer is another \(\mathcal{L}\)ayman, he would be subject to the very same (mis-)understanding; and if she is an \(\mathcal{E}\)xpert, she will recognize his utterance as a \(\mathcal{L}\)ayman’s talk and gather that he does not realize that the person called *Barbara Vine* also wrote books under a different name. In both cases, then, the hearer will understand that the speaker did not mean to say anything about the latter.

\(^{29}\) This concept as well as its formal reconstruction (to be given in Section 4.5) are closely related to Haas-Spohn’s (1995) theory of subjective meaning, over which in Kaplan’s (1989b: 574) terms, a metasemantic stance is taken. – As to the concept of an objective language, I am glancing over a lot of interesting, but largely irrelevant subtleties. For example, there is a limit to linguistic ignorance in that speakers need to know enough about the meanings of enough expressions to count as speakers in the first place; and there is also a good deal of vagueness as to what the socially determined meanings are.

\(^{30}\) Hence in each of his subjective languages, *Ruth Rendell* and *Barbara Vine* refer to different persons. But who are these persons? Whoever they are, they cannot be the same; \textit{a fortiori}, they cannot both be identical to Ruth Rendell. Any further specification of them depends on the metaphysics and epistemology of cross-world identification. Cf. Lewis (1981).
By the time the bookseller uttered (25), it had become common knowledge between us that he was an Expert, which even his Layman-like linguistic behaviour could not change. This is why I did not take him to simply make an utterance in L, to be interpreted literally. Rather, I took him to communicate precisely what he would have communicated had he been a Layman, viz. that he has read all of the books Ruth Rendell wrote under the author name Barbara Vine. In other words, I took his utterance to be a counterfactual speech act. More generally and schematically, a sentence $S$ of $L$ would thus be amenable to the following counterfactual reinterpretation when uttered by an Expert:

\[(28) \quad \text{I would have said } S, \text{ had I been a Layman.}\]

That (28) delivers the intended interpretation in the case at hand crucially depends on the precise construal of the counterfactual condition, the nature of which will be investigated in the next subsection. The idea is that, had the bookshop owner been a Layman, his beliefs about the coreference of the names Barbara Vine and Ruth Rendell would have been different, but everything else would have been as it actually is. In particular, he would have had read the very same books that he actually read. Of course, he would have been under the impression that the ones written under the name Barbara Vine were the only ones written by the bearer of that name. Hence, by uttering (25), he would have been under the impression of only saying that he read all of those books. Assuming that, whether Layman or not, he would have spoken truthfully, it follows that the bookseller, had he been a Layman, had read all of the books written under the name Barbara Vine. Since the books that he had read had he been a Layman, are the books he actually read, one may conclude that he has actually read all of those books. On the other hand, had he been a Layman and uttered (25), he would have been under the impression of not committing himself to having read all of Ruth Rendell’s books. Hence his being a Layman is compatible with his truthfully uttering (25) without having read all of Ruth Rendell’s books. Hence, since the books he would have read as a Layman are the books he has actually read, the counterfactual reinterpretation (28) of his utterance does not imply that he has read all of Ruth Rendell’s books. This, I take it, is an adequate interpretation of the bookseller’s utterance.

Since, by hypothesis, $L$ only differs from $E$ by the presence of an additional name for Ruth Rendell and the consequent violation of the Uniqueness Principle, the counterfactual reinterpretation of an Expert’s Layman’s talk can be given the following rough paraphrase that will come in handy in a minute:

\[(29) \quad \text{I would have said } S \text{ had I believed that the bearer of the name } Barbara \text{ Vine (in } L \text{) is distinct from the bearer of the name Ruth \text{ Rendell (in } L \text{).}\]

Saul’s case of Superman and Clark Kent is quite parallel. As pointed out above, the pertinent construal of (1) arises if the sentence is used in communication between Enlightened speakers: by using the name Clark Kent, which is absent from his language $E$, the Enlightened speaker behaves like an Ignoramus, who would have communicated that a person whom he refers to as Clark Kent entered a phone booth, which was (subsequently) left by a different person, viz. the one he refers to as Superman. Given the speaker’s counterfactual speech act, the Enlightened hearer may conclude that, were it not for the fact that the speaker was Enlightened too, the latter had reason to believe that two distinct persons were involved in the scenario reported by him. If the speaker’s report is based on an eye-witnessing, it is obvious that what would have led him (as an Ignoramus) to the conclusion that the enterer was called Clark Kent and the leaver was called Superman
must have been that they looked like these two persons, which is most likely due to the different attire the true common bearer of the two names wore on the two occasions. Since Superman’s different appearances would have formed the ground of the speaker’s erroneous judgment had he been an Ignoramus, Superman must actually have looked differently in the two parts of the scenario described; for the counterfactual situation only differs from the reported reality in the speaker’s information state, which can hardly affect the way Superman dresses. Hence, according to the speaker’s report, Superman must have changed his outward appearance between entering and leaving the phone booth. Again, I take it that this is an adequate interpretation of the utterance.

Epistemic perspectives seem to have little to do with the difference between Chemnitz and Karl-Marx-Stadt. Hence a treatment of (2) along the lines of (1) does not look all too promising. Even if one somehow managed to render (30) equivalent to (31) and both equivalent to the result of interpreting Karl-Marx-Stadt as a restricted name, this construal does not seem to reflect a general and natural strategy of interpretation.

(30) \( I \text{ would have said } S, \text{ had I been a speaker of } \Theta. \)
(31) \( I \text{ would have said } S \text{ had I believed that we are currently in an S-Phase.} \)

In Section 4.5, I will briefly address the question of what it takes to have the counterfactual speech act of a Newspeaker’s utterance come out in the sense of (31). Whether such formalization is adequate in the first place, I will largely leave open here, devoting the remainder of this paper to the non-temporal examples.

4.4 Mechanisms of Reinterpretation

In order to see how the non-literal construal of the bookseller’s utterance (25) can be arrived at in a systematic fashion, combining literal meaning and contextual background, I will first isolate three features that distinguish the counterfactual speech act from an ordinary assertion. As the formalization to be presented in Section 4.5 will make clear, the reinterpretation process roughly described in (28) may then be seen as a combination of these three features.

By containing a name that is not part of the language spoken among Experts, the utterance was taken, and intended to be taken, as an instance of code switching from Expertese, which had been the language of our conversation so far, to Laytalk. One of the prerequisites of any successful code switching is that the language switched to is intelligible to all participants in the conversation. In the case at hand this means that the bookseller and I were in a position to speak and understand utterances made by Laymen. \( L \) only differs from \( E \) by one name (and the differences that result from this lexical difference), and so for us Experts to understand Laytalk means to realize that Barbara Vine (as used in \( L \)) is but another name for Ruth Rendell – and there can be no doubt among us that we do realize this. More importantly though, apart from intelligibility there is another general prerequisite for code switching, viz. that the language in which the utterance was made is recognized.31 To see what is at stake here, let us briefly recapitulate the bookseller’s alternative utterances discussed earlier:

(25) \( I \text{ve read all of Barbara Vine’s books.} \)
(26) \( I \text{ve read all of Ruth Rendell’s books.} \)
(27) \( I \text{ve read all of Ruth Rendell’s books, though not all of Barbara Vine’s.} \)

31 Strictly speaking, in view of the kind of linguistic ignorance discussed in the previous subsection, demanding full identification of the intended language by the hearer (or the speaker, for that matter) would obviously be too much.
(26) is a sentence in \(E\)xpertese, and I would probably have taken it as such had the book-seller uttered it instead of (25). However, (26) is also a sentence of \(L\) and as such has the same literal meaning as it has in \(E\): Ruth Rendell is a (neutral) name for Ruth Rendell in both languages, which do not disagree in the interpretation of any of the other parts of (26) either. By the same token it is readily seen that (25) and (26) must have the same literal meaning in \(L\). Still, I insist, I would have taken an utterance of (26) as a sentence of \(E\), not of \(L\), because – and as long as – there was no indication of code switching. Fair enough, one may object, but what would have been the difference, given the synonymy just observed? The difference would have been (I reply) that, confronted with a code switch from \(E\) to \(L\), I would not have taken the utterance in its literal meaning. A brief look at (27) reveals this: the speaker would have left \(E\) with the second clause; consequently the first clause is also (preferably) understood as an utterance in \(L\), and as such it is not interpreted literally but as part of a counterfactual speech act. Hence it seems to be the very use of the name Barbara Vine that indicates a switch from \(E\) to \(L\) and triggers the non-literal interpretation. But how come that it is \(L\), rather than some other language that the speaker successfully alludes to? After all, there are other languages containing Barbara Vine as a name for Ruth Rendell (or maybe someone else). For instance, why couldn’t the speaker have uttered (25) as a sentence of a language that contains Barbara Vine instead of Ruth Rendell and otherwise coincides with \(E\)? I think this is a matter of contextual background and expectation: we all know, or can easily imagine, people who have come across the two author names without ever realizing that they are coreferential. Unlike these Laymen, people who are using the name Barbara Vine without using Ruth Rendell are somehow harder to come across or even imagine – for whatever reason: the name Barbara Vine simply has a Laymanish ring.

If code switching were all that is going on in the bookseller’s utterance of (25), I would have interpreted it literally, thus ending up with the same result as if he had uttered (26). What made the interpretive difference was that I not only recognized the sentence as one of a foreign language, but also took the speaker to act as if he had been a speaker of that language. In other words, code switching indicated a context shift, albeit a hypothetical one. And just like the target language of the linguistic switch, so would the target context of the contextual shift have to be identifiable by his partner in conversation, i.e. me. Hence the hypothetical context hinted at by the speaker was not just any old context in which the language differed from ours. It was the context that was exactly like our actual situation (then, in New Brunswick, New Jersey, by the way) – except that \(L\) and not \(E\) was the language of conversation. For this to be the case, I take it, the speaker would have had to be a Layman, although he would still have been a bookseller, and he would have read and liked the same books as he actually did, although being a Layman, he would have been ill-informed somehow about the authorship of some of them. Maybe, in that hypothetical and counterfactual context, he had never come across, or chosen to ignore, or forgotten about, the author’s announcement quoted earlier in which she reveals her identity. This is hard to imagine, given the fact that the same ‘revelation’ can be found on the cover of practically any book written under the author name Barbara Vine, but then the context we are dealing with is a hypothetical and counterfactual one. And it is this context to which the bookseller draws my attention by code switching.

Of course, the hypothetical context shift is by no means a regular side-effect of code switching, but lies in the very nature of the target code, which is the Layman’s language. Violating \((U)\), the latter is defective, and the (typical) Expert knows this. Hence the bookseller...
must have had a good reason for lapsing into Laymanese, and one that I would have had to be able to reconstruct. Merely using a different name for Ruth Rendell does not appear to be a good reason, especially since this would not have changed the content of what the speaker would have expressed by uttering (26). However, instead of intending the utterance to be interpreted as is, he wanted it to be understood as if it had been made from a Layman’s perspective; in other words, instead of relating it to the actual utterance context, he wanted it to be an expression of his – hypothetical and counterfactual – subjective background. For even though he would have had expressed the same proposition by uttering (25) as he would have done by uttering (26), as a Layman he would not have been aware of this fact. Instead, he would have been under the impression of saying something close to:

(32) I've read all of the Barbara Vine books.

In (32) the italicized noun phrase is a name of the name Barbara Vine, a meta-name (if you wish), which I take to be part of both the Expert’s and the Layman’s language. Of course, this does not mean that the Layman takes (25) and (32) to be synonymous; one is about an author, the other one about her name. But for all intents and purposes, (32) may go proxy for (25) – or so the Layman erroneously believes. According to him, uttering (32) would have had the same communicative effect as uttering (25): in either case he would have informed me that he has read all of the books that were written under the name Barbara Vine. And it is this subjective understanding of (25) that the bookseller is hinting at when he puts on his Layman act. Given the close resemblance of the hypothetical and the actual context, I could then easily figure out that what he would have tried to get across as a Layman was actually the case, viz. that (32) is true of him, the Expert.

The rationale behind the counterfactual speech act is that code switching creates a hypothetical context shift – an actual context shift being out, because the bookseller does not become a Layman by simply speaking like one – and that the context shift invites a subjective interpretation, because the literal meaning of the sentence uttered can somehow be excluded. As to the latter proviso, I am not sure about the reasons. After all, code switching as such may occur for a variety of reasons, and so why didn’t the speaker merely use the Layman’s language in a literal way? I suppose that the fact that he may as well have used (26) then is reason enough for not supposing that merely an otherwise unmotivated code switch took place and thus look for a non-literal interpretation. The fact that the language he uses is defective – as the Experts know, L violates (U) – may then be of help in arriving at a subjective interpretation: the latter is in line with (UP). I admit that this reasoning is still a far cry from a systematic derivation, but I think it is on the right track.

Understanding how the intended non-literal construal of the utterance comes about is one thing. Explaining why the speaker has chosen such a roundabout way of expressing him-
self is another one, which is the more pressing if he had simpler alternatives at hand. And he did. For he could have reached the communicative effect of his counterfactual speech act by simply uttering (32), which is easily recognizable as a piece of \( \mathcal{E} \)xpert talk, ready to be interpreted literally. So why did he not do that? Again I can only speculate. Maybe, by relying on the fact that speaker and hearer were both \( \mathcal{E} \)xperts, the counterfactual speech act established a certain sense of belonging, as a switch into a familiar dialect may do. Certainly more cases will have to be explored to settle this question.

Generalizing from the example, a speaker performs a counterfactual speech act if he uses a sentence that is recognizably part of a language \( L \) not spoken in the utterance context and wants it to be understood as having the communicative effect he would have intended it to have had he spoken \( L \). Once again, the three ingredients that lead from the sentence uttered to its intended interpretation are:

- code switching;
- hypothetically shifting the context; and
- interpreting the utterance subjectively.

About the first one, I do not have anything to add to what I have already said, except that determining the target of the switch, though certainly a matter of contextual background knowledge, cannot be part of the interpretive process. After all, ‘semantics cannot tell us what expression was uttered or what language it was uttered in’ – as David Kaplan observed.\(^{35}\) As to the context shift, I hope to have already made it clear that it may be thought of as an implicit counterfactual conditional along the lines of the paraphrase (28) above. One advantage of this understanding, and of isolating the context shift in the first place, is that it thus becomes amenable to a formal treatment along familiar lines. And as to subjective interpretation, it is well known that two-dimensionalism, i.e. the distinction between context and index, may do some service here. More precisely, if doxastic uncertainty is modelled by sets of utterance contexts, separating out the contexts being in which the speaker does not exclude (on the basis of his beliefs), then the speaker’s subjective construal of a sentence may be identified with all those contexts at which it would have been true. This account of subjectivity, diagonalization, is a well-respected tool in formal pragmatics, and it will be put to work in the following subsection, where some of what I have said so far is made more precise.\(^{36}\)

4.5 **Formal Modelling**

The following formal development is rather sketchy and gappy. To a large extent, this is so because filling in the details would be a distracting and tedious task, but it is obvious that it can be done, and how it can be done. To an equally large extent though, gaps occur when it is not so obvious that they can be filled, or how precisely they should be filled. I still hope that the network of formal definitions, assumptions, and observations presented on the following pages, offers a glimpse of what a fully fledged pragmatic account of the extensional substitution puzzles would amount to.

Somewhat simplifyingly, languages will be identified with sets of (disambiguated declarative) sentences denoting Kaplanian *characters*. More precisely, given the sets \( W, T, \) and \( D \) of possible worlds, times, and speakers respectively, let us pick a set \( C \subseteq W \times T \times D \) and a family \( (L_c)_{c \in C} \) of functions \( L_c : \mathcal{A}_c \rightarrow \wp(W \times T) \), where \( \mathcal{A}_c \neq \emptyset \) is the set of sentences of \( L_c \). The elements \( c = (w_c, t_c, d_c) \) of \( C \) are the (possible utterance) contexts; \( L_c \) is

---

35 Kaplan (1989a: 559), emphasis added. The quotation is taken from a passage that is mainly directed against treating *ambiguity* as context dependence.

36 The usefulness of two-dimensionalism in the analysis of subjectivity has already been speculated on in Kaplan (1989a: 529ff.). The central rôle the diagonal has been first emphasized in Stalnaker (1978).
the speaker’s objective language in the context c. If L is a language (i.e. the speaker’s objective language in some context), \( \varphi \in A_L \) and \( c \in C \), then \( [\varphi]^c_L := L(\varphi)(c) \) is the (horizontal) proposition expressed by \( \varphi \) in c (according to L). A sentence \( \varphi \in A_L \) of some language L is L-true in a context \( c \in C \) if, and only if, \( (w, t) \in [\varphi]^c_L \); the (semantic) diagonal of \( \varphi \) (according to L) is the set \( \Delta_L(\varphi) := \{ c \in C | \varphi \text{ is L-true in } c \} \).

Although a full formal treatment of Neutrality would call for a more fine-grained individuation of languages, the present framework does allow for a general and abstract formulation. For each language \( L \), let \( E_L \subseteq A_L \) be the set of equations between any two distinct names. Then the following formalization is a close as we get to (NP):

(NPF) For any language \( L \) and \( \varphi \in E_L \) it holds that:
- either: \([\varphi]^c_L = W \times T \) for all \( c \in C \),
- or else: \([\varphi]^c_L = \emptyset \) for all \( c \in C \).

We will also make use of a relation of syntactic (and lexical) isomorphism \( \sim \) between languages \( L \) and \( L' \); the following crude characterization will be sufficient: \( L \sim L' \) whenever \( A_L = A_{L'} \) and \( E_L = E_{L'} \).

Subjectivity is captured by a function \( \Pi \) with domain \( C \) and values that are sets of contexts: \( \Pi \subseteq C \), where \( \Pi_{(w, t, d)} \neq \emptyset \) is (speaker) d’s doxastic perspective in \( w \) at \( t \), i.e. the set of those contexts she may be in, for all \( (d \text{ thinks } d \text{ knows in } w \text{ at } t) \); I will refer to the elements of \( \Pi_c \) as the (speaker’s) doxastic alternatives at \( c \). We will say that the speaker (in a context \( c \in C \)) is reliable as long as \( c \in \Pi_c \); although this notion inherits the notorious plague of coarseness from the possible worlds background, it will do for our purposes. The speaker’s subjective languages (in a context \( c \in C \)) are those languages \( L \) for which there are contexts \( c' \) such that \( c' \in \Pi_c \) and \( L_{c'} = L \); the set of subjective languages in \( c \in C \) is \( \Sigma_c \). For simplicity, I will assume full syntactic (and lexical) competence: \( L \sim L_c \) whenever \( L \in \Sigma_c \). Moreover I assume that there is no objective language without subjective language: \( \Sigma_c \neq \emptyset \) whenever \( c \in C \). In case \( L_c \notin \Sigma_c \), the speaker has some misconception about her objective language (and is not reliable); and if \( \Sigma_c \) contains more than one element, then — whether or not she has a misconception — the speaker is uncertain about its exact (semantic) nature. She may also be uncertain as to which proposition she would express by using a given sentence \( \varphi \) of her objective language \( L_c \); for there may be two contexts neither of which she excludes to be in and in which \( \varphi \) expresses distinct propositions — whether or not the (objective) languages at those contexts

\[ \text{Note that, according to this definition, not every mapping from the sentences } A_L \text{ of a given language } L \text{ can be a language. Otherwise one could construct a diagonal language } L^* \text{ by putting } L^*(\varphi)(c) := (W \times T) \setminus L_c(\varphi)(c) \text{ (for any } c \in C \text{ and } \varphi \in A_L \text{)} \text{ — and end up in a contradiction: by definition of a language, } L = L_{c^*}, \text{ for some } c^*, \text{ and so } L_{c^*}(\varphi)(c) = L^*(\varphi)(c^*) = (W \times T) \setminus L_{c^*}(\varphi)(c^*)! \]

\[ \text{Unlike names (cf. fn. 18) equations may not be identifiable without interpretation. If so, both } (NPF) \text{ and the formal version } (UPF) \text{ of the Uniqueness Presumption below would have to be consequences of underlying semantic principles and cannot replace them.} \]

\[ \text{The general approach to subjectivity in terms of alternatives is due to Hintikka (1962), of course. Since any } c \in C \text{ provides a speaker, the perspectives form the basis of self-locating belief, i.e. belief } \textit{de se} \text{ in the sense of Lewis (1979).} \]
are the same: \( c' \in \Pi_c \), \( c'' \in \Pi_c \), but \( [\varphi]_{L_c}^c \neq [\varphi]_{L_{c'}}^c \) (whether or not \( L_c = L_{c'} \)). In such cases of uncertainty – which, as regards to sentences \( \varphi \) of the speaker’s objective languages may well be the rule rather than the exception – she may still be able to use such \( \varphi \) to express whatever \( \varphi \) may express for all she knows, or takes herself to know. By uttering a sentence \( \varphi \in \mathcal{L}_{L_c} \) in a context \( c \) the speaker intends (or would intend to express) a character \( \chi : C \rightarrow \varphi(W \times T) \) just in case \( \chi(c') = [\varphi]_{L_c}^c \), for any \( c' \in \Pi_c \). A sentence \( \varphi \in \mathcal{L}_{L_c} \) of the speaker’s objective language \( L_c \) in a context \( c \in C \) is subjectively true in \( c \) if, and only if, \( \varphi \) is \( L_c \)-true in all contexts \( c' \in \Pi_c \); the metasemantic diagonal of \( \varphi \) is \( \Delta(\varphi) := \{ c \in C \mid \varphi \text{ is subjectively true in } c \} \). And the speaker is, or would be, sincere in \( (L-) \) asserting \( \varphi \) in \( c \) if, and only if, by uttering \( \varphi \) in \( c \) she intends a character \( \chi \) such that \((w_c, t_c) \in \chi(c')\), for all contexts \( c' \in \Pi_c \).

Two remarks on the last two definitions are in order. First, given that the condition on the intended character \( \chi \) is normally satisfied by a lot of functions that only differ in the propositions they assign to contexts outside the speaker’s doxastic perspective, this is a rather coarse approximation to subjective meaning; but I think it will do for my present purposes. Secondly, it may be noted that the speaker’s sincerity in uttering \( y \varphi \) in \( c \) could have been defined equivalently in terms of the metasemantic diagonal of \( \varphi \):

\[ (33) \quad \Pi_c \subseteq \Delta(\varphi). \]

A full formal treatment of the Uniqueness Presumption would again call for a more fine-grained individuation of languages; but, as in the case of \( (NPF) \), the present framework allows for a general and abstract formulation. According to \( (UP) \), all members of \( \mathcal{E}_L \) ought to be false in all subjective languages \( L \), which is why we adopt:

\[ (UPF) \quad [\varphi]_{L}^c = \emptyset, \text{ for any } c \in C, L \in \Sigma_c \text{ and } \varphi \in \mathcal{E}_L. \]

Whatever counterfactuality is needed will be supplied by a function \( + \) associating with each context \( c \in C \) and language \( L \) a context \( c+L \) that is like \( c \) except for the fact that \( L \) is the speaker’s objective language in \( c+L \). Hence two minimal requirements on \( + \) are \( S[\text{selectivity}] \) and \( S[\text{strong centering}] \), which we take to hold for all languages \( L \) and contexts \( c \); we also assume that code-switching does not change denotations, i.e. that the proposition a sentence of a given language expresses and the truth value it has are preserved in a context in which a different language would be spoken.\(^{41}\)

\(^{40}\) Cf. Stalnaker (2001) on the relation between the semantic and the metasemantic diagonal.

\(^{41}\) One may think of \( + \) as determined by a more general selection function \( f: \varphi(C) \times C \rightarrow C \) along the lines of Stalnaker (1968): \( c+L = f((c' \in C \mid L = L_{c'}), c); (Sel) \) and \( (SC) \) then follow from corresponding assumptions about \( f \). (Int) generalizes Kripke’s (1972: 290) remarks about ‘describing a possible world or counterfactual situation in which people, including ourselves, did speak in a certain way different from the way we speak’: ‘we use English with our meanings and our references’ (ibid.). (Ext) obviously imposes some harmless restrictions on expressive power by excluding reference to foreign (objective) languages, as in the informal paraphrase (29) of counterfactual assertion.
Assumptions about $+$ (for any languages $L$ and $L'$, $\varphi \in \mathcal{A}_L$ and $c \in C$):

(Sel) $L_{c+L} = L$.

(SC) $c + L_c = c$.

(Int) $[\varphi]_{L_{c+L}} = [\varphi]_{L_c}$.

(Ext) $(w_c, t_c) \in [\varphi]_{L_{c+L}}$ iff $(w_{c+L}, t_c) \in [\varphi]_{L_c}$.

(TI) $t_c = t_{c+L}$.

(TM) reflects the fact that counterfactual context shifts only affect the world and and will be needed later. A similar inertia condition on the speaker would be equally natural but does not play a role in the formal reconstruction. Further, more specific properties of $+$ will emerge in the formal treatment of the above examples.

We are now in a position to capture the counterfactual speech act performed by uttering a sentence that is not part of the speaker’s (objective or subjective) language. If $c \in C$ and $L$ is a language such that $\varphi \in \mathcal{A}_L \setminus \mathcal{A}_{L_c}$ then the speaker in $c$ is (or would be) open in counterfactually (L-) asserting $\varphi$ if, and only if, the (same) speaker would have been sincere in uttering $\varphi$ in $c + L$. As we will see, openness is to counterfactual assertion what sincerity is to ordinary assertion; the two concepts are primarily distinguished for reasons of terminological transparency.

The previous definition still depends on the language $L$ to which a hypothetically uttered sentence is meant to belong. Given the above strict identity criteria for languages, a sentence that is not part of the language of the utterance context will be part of many other languages. Hence, in order to interpret a counterfactual speech act, one of these languages would have to be picked out first. I suspect that, where no natural choice can be made, the counterfactual interpretation is simply out. But even in the cases envisaged above, no single language in the sense defined here will suggest itself. It rather seems that a counterfactual speech act can be made in cases where a sufficiently well-defined class of alternative languages containing the sentences uttered is equally obvious to the speaker and the hearer, in which case the resulting interpretation will be somewhat underdetermined. Though such uncertainty about the language hinted at by uttering a non-sentence can clearly be captured within the present framework, I am afraid this additional complication would create more confusion than insight. I therefore abstain from implementing it here, presuming that a single obvious alternative to the language of the utterance context is always available.\footnote{An earlier version of this paper contained an alternative treatment that does justice to the indeterminacy of the language hinted at; it has been omitted for reasons of space.}

We must finally say something about the \textit{communicative effect} of utterances. To this end, let us take the common ground underlying communication in a context $c$ to form a set $\Gamma_c \subseteq C$ of contexts, for which we will assume syntactic competence – $L_{c'} \sim L_c$ whenever $c' \in \Gamma_c$ – as well as trust: the speaker is reliable in all $c' \in \Gamma_c$. It is this common ground that is primarily affected by the speaker’s utterance of a sentence $\varphi \in \mathcal{A}_L$ of some language $L$ in the context $c$, in one of the following two ways.

If $L = L_c$, then the utterance may be taken literally, and at the time $t'$ immediately following it we have the following \textit{communicative effect of an (ordinary) assertion}:\footnote{(34) is in the spirit of Gazdar’s (1979: 46) reconstruction of Grice’s (1975: 45) \textit{Maxim of Quality}. Given trust, (34) implies, but is not equivalent to:}

42

43

25
Assumptions about (d) contains the name (a) \(\Delta \neq \emptyset\) than sincere: 

\[
\Gamma_{(w,f,d)} \subseteq \Gamma_{(w,f,d)} \cap \{c' \in C \mid \text{in } c', \text{ the speaker is sincere in } L\} - \text{by (33)}
\]

If, on the other hand, \(L \neq L_c\), then the utterance may be taken as a counterfactual assertion, in which case the communicative effect is that the speaker had been open rather than sincere:

\[
\Gamma_{(w,f,d)} \subseteq \Gamma_{(w,f,d)} \cap \{c' \in C \mid \text{in } c', \text{ the speaker is open in counterfactually } L\}-\text{asserting } \varphi\}
\]

\[
\Gamma_{(w,f,d)} \cap \{c' \in C \mid \Pi_{c \rightarrow L} \subseteq \Delta(\varphi)\}
\]

In order to see what this amounts to in the cases under scrutiny, we must recast them within the present formal framework. I will only do so for the case Ruth Rendell vs. \textbf{Barbara Vine}, leaving the other examples to the reader. Let us therefore see what the pre-

formalization can do for the following sentences.

(25) I've read all of Barbara Vine's books.
(26) I've read all of Ruth Rendell's books.
(32) I've read all of the \textit{Barbara Vine} books.
(36) I've read all of the \textit{Ruth Rendell} books.
(37) \textbf{Barbara Vine} is \textbf{Ruth Rendell}.
(38) \textit{[The name] Barbara Vine is [the name] Ruth Rendell}.

The two objective languages involved are \(\mathcal{L}\) and \(\mathcal{E}\). The following assumptions about them are fairly straightforward:

**Assumptions about \(\mathcal{L}\) and \(\mathcal{E}\) (and +):**

(a) \(\mathcal{A}_\mathcal{L} \subseteq \mathcal{A}_\mathcal{E}; \mathcal{E}_\mathcal{E} \subseteq \mathcal{E}_\mathcal{L}; (25) \in \mathcal{A}_\mathcal{L} \setminus \mathcal{A}_\mathcal{E}; (26),(32),(36) \in \mathcal{A}_\mathcal{E}; (37) \notin \mathcal{E}_\mathcal{L} \setminus \mathcal{A}_\mathcal{E}; (38) \in \mathcal{E}_\mathcal{E}.

(b) \(\mathcal{L}(25) = \mathcal{L}(26) = \mathcal{E}(26); \mathcal{L}(32) = \mathcal{E}(32); \mathcal{L}(36) = \mathcal{E}(36).

(c) \(\llbracket (37) \rrbracket_\mathcal{L} = W \times T \neq \llbracket (38) \rrbracket_\mathcal{E} = \emptyset - \text{for any } c \in C.

(d) \(\llbracket (37) \rrbracket_\mathcal{L} = \llbracket (38) \rrbracket_\mathcal{E} = \emptyset - \text{whenever } L_c = \mathcal{L} \text{ and } L \in \Sigma_c.

(e) \(\llbracket (25) \rrbracket_\mathcal{L} \neq \llbracket (26) \rrbracket_\mathcal{E}; \llbracket (32) \rrbracket_\mathcal{L} \neq \llbracket (36) \rrbracket_\mathcal{L} - \text{for any } c \in C \text{ such that } L_c = \mathcal{L} \text{ and any } L \in \Sigma_c.

(f) \(\Delta(25) \cap \Pi_c = \Delta(32) \cap \Pi_c \neq \Delta(26) \cap \Pi_c = \Delta(36) \cap \Pi_c - \text{for any } c \in C \text{ such that } L_c = \mathcal{L}.

The syntactic assumptions in (a) are all straightforward: only the</p>

\(\mathcal{L}\)ayman’s language contains the name \textbf{Barbara Vine}; but both the \(\mathcal{L}\)ayman’s and the \(\mathcal{E}\)xpert’s language contain the (italicized) names of the alleged \textit{noms de plume} \textbf{Barbara Vine} and \textbf{Ruth Rendell}. (b) reflects the fact that \(\mathcal{L}\) and \(\mathcal{E}\) agree on the meanings of all expressions that do not contain the name \textbf{Barbara Vine}, which – although speakers are not aware of it – objectively refers to Ruth Rendell (in \(\mathcal{L}\)) and is thus synonymous with \textbf{Ruth Rendell} (in \(\mathcal{L}\) and \(\mathcal{E}\)). (c) expresses the semantic effect of this hidden synonymy on the proposition expressed by the
equation (37) in \( L \), as well as the contrary effect of the referential distinctness of the two meta-names. On a fuller semantic account, \((b)\) and \((c)\) could be derived compositionally along the lines indicated in the previous sections. Together with \((a)\) and syntactic competence, \((UPF)\) implies \((d)\) – but does not imply \((e)\), which would not even follow had we bothered to include a level of subentential compositionality. In fact, \((d)\) does not even imply the considerably weaker inequalities:

\[
\begin{align*}
(e') \quad L(25) &\neq L(26); \quad L(25) \neq L(26).
\end{align*}
\]

Still, \((e)\) is certainly a highly natural assumption given \((c)\): if a name is replaced by a referentially distinct one, the resulting sentence ought to express a different proposition. 44

In fact, the second inequality in \((e)\) should also hold for the Layman’s and the Expert’s respective objective languages \( L \) and \( E \). The idea behind \((f)\) is that, the Layman takes Barbara Vine to be coreferential with the person called Barbara Vine and is thus under the impression that \((25)\) and \((32)\), though not expressing the same proposition, have the same truth value; similarly for \((26)\) and \((36)\), which have been included for illustration only.

In order to arrive at the desired construal of the bookseller’s utterance of \((25)\), we need to implement some more specific observations about the particular context involved. Let us identify the scene described in Section 3.4 with an utterance context \( c^* = (w^*, t^*, d^*) \in C \), where \( w^* \) is (or represents, for that matter) the actual world, \( t^* \) is some particular time in May 1995, and \( d^* \) is the bookseller I had been talking to. The following assumptions about \( c^* \) turn out to be sufficient to make the points announced and sketched in the previous subsections:

**Assumptions about \( c^* \) (and +):**

\[
\begin{align*}
(g) \quad & L_{c^*} = E; \\
(h) \quad & \Pi_{c^*} \subseteq \{ c \mid \text{for some } L \in \Sigma_{c^*} \subseteq L, c + L \in \Pi_{c^*} \}; \\
(i) \quad & L_{c^*} (32) = L_{c^*} (32); \quad L_{c^*} (36) = L_{c^*} (36); \quad L_{c^*} (38) = L_{c^*} (38) \\
& \quad - \text{whenever } c \in \Pi_{c^*} \text{ and } c' \in \Pi_{c^*} \text{ such that } c' = c + L_{c^*}.
\end{align*}
\]

About \((g)\) nothing needs to be said. \((h)\), which has been formulated as a specific property of \( c^* \) although it may well be generalizable to all \( c \in C \), describes the effect of the hypothetical language shift we are interested in on the subject’s perspective. One may have expected a more direct connection, viz:

\[
(h') \quad \Pi_{c^*} = \{ c + L \mid c \in \Pi \}.
\]

However due to \((UPF)\), \( L \) cannot be a subjective language, and so this door is closed. 45 Instead, the idea behind \((h)\) is that the contexts compatible with \( d^* \)'s counterfactual belief are as close to the contexts compatible with his actual belief (at the time) as is consistent with the fact that he would be a speaker of what he would take to be the Layman’s language. Since different doxastic alternatives at \( c^* \) may resolve \( d^* \)'s semantic uncertainty in different ways, there need not be a unique such language he takes to be \( E \), i.e. \( \Sigma_{c^*} \).

---

44 Cf. Rooth (1985: 85f.). Due to general limitations of the possible worlds framework, this cannot be expected of all sentences though.

45 More precisely, if we had \( c^* + L \in \Pi_{c^*} \) (which we have been assuming to be non-empty), then \( L_{c^*} = L \) (by \((Sel)\)) and thus: \( L \in \Sigma_{c^*} \subseteq \{ (L_c \mid c' \in \Pi_{c^*}) \} \). Hence, given that \((37) \in \mathcal{M}_{c^*} \) (by assumption \((a)\) ), \( \models (37) \in \mathcal{M}_{c^*} \) for any \( c' \in C \), according to \((UPF)\) – but contradicting \((e)\).
need not be a singleton. In that case each doxastic alternative \( c \in \Pi_* \) will give rise to a way the world may be like according to him as a Layman, viz. like \( c \) except that the language he speaks (and takes to be \( L \)) would be one that \textit{minimally deviates from} \( L_c \) by resolving all semantic uncertainties in the same way as \( L_c \) does. But then not every combination \( c+L \) of a doxastic alternative \( c \in \Pi_* \) and a subjective language \( L \in \Sigma_{c+L} \) will qualify as a doxastic alternative at \( c+L \), which is why we should not endorse the following strengthening of (h):

\[
(h') \quad \Pi_{c+L} = \{c+L \mid c \in \Pi_* \& L \in \Sigma_{c+L} \}.
\]

Finally, (i) says that, as far as \( d^* \)'s subjective interpretation is concerned, certain sentences of \( E \) are not affected by the counterfactual language shift: a minimal deviation from a given context in the speaker's perspective should not require more changes than those due to the shift from one (subjective) language to the other. Given that the languages overlap considerably, most sentences would come out as (subjectively) meaning the same in the two (subjective) contexts, although their meanings do not have to be stable across the subject’s doxastic alternatives. 46 The reader is now invited to verify that the above assumptions (a)–(i) imply the following predictions:

\( j \) \quad \text{\( d^* \) is open in counterfactually \( L \)-asserting (25) in \( c^* \) iff \( d^* \) would be sincere in \( \mathcal{E} \)-asserting (32).}

\( k \) \quad \text{If \( d^* \) is reliable and would be open in counterfactually \( L \)-asserting (25) in \( c^* \), then (32) is \( \mathcal{E} \)-true in \( c^* \).}

\( k \) is immediate, given (j): \( d^* \)'s openness and reliability imply that \( c^* \in \Pi_{c^*} \subseteq \mathcal{L}_{c^*} \) \( \mathcal{L}_c \), i.e.: \( c^* \in \mathcal{L}_{c^*} \), by (g).

In order to also capture the communicative effect of the bookseller's utterance we have to take an intersubjective turn, applying the reasoning that led to (j) and (k) from within the common ground, provided the latter covers whatever we have been assuming about the context \( c^* \). Let us therefore make the following

\textit{Assumptions about} \( \Gamma_{c^*} \) (and +):

For any \( c^0 \in \Gamma_{c^*} \) the following holds:

\( g_{\Gamma} \) \quad \text{\( L_{c^0} = \mathcal{E} \).}

\( h_{\Gamma} \) \quad \text{\( \Pi_{c^0} \subseteq \{c \mid \text{for some } L \in \Sigma_{c^0+L}, c+L \in \Pi_{c^0+L} \} \); \( \Pi_{c^0+L} \subseteq \{c+L \mid c \in \Pi_{c^0} \& L \in \Sigma_{c^0+L} \} \).}

\( i_{\Gamma} \) \quad \text{\( L_c(32) = L_c(32) \); \( L_c(36) = L_c(36) \); \( L_c(38) = L_c(38) \)

\hspace{1cm} – whenever \( c \in \Pi_{c^0} \) and \( c' \in \Pi_{c^0+L} \) such that \( c' = c+L_{c^0} \).

… and immediately obtain the following \textit{predictions} (for any \( cI \in \Gamma_{c^*} \)):

\( j_i \) \quad \text{The speaker in \( c^0 \) would be open in counterfactually \( L \)-asserting (25) in \( c^0 \) iff the speaker in \( c^0 \) would be sincere in \( \mathcal{E} \)-asserting (32).}

\( k_r \) \quad \text{Suppose the speaker in \( c^0 \) would be open in counterfactually \( L \)-asserting (25) in \( c^0 \) \; then (32) is \( \mathcal{E} \)-true in \( c^0 \).}

46 Do all sentences of \( \mathcal{E} \) (and hence in \( \mathcal{L}_L \cap \mathcal{L}_E \)) emerge unscathed from the (subjective) language shift? This seems questionable in view of cases like (26). I am afraid I will have to leave the matter open here.
... as respective corollaries to (j) and (k). Note that, unlike (k), (k₁) can do without any appeal to reliability because we have been assuming trust as a property of the common ground. Alternatively, we could have replaced the doxastic by epistemic perspectives all along, subsuming reliability under the properties of the accessibility relation $\Pi$: $c \in \Pi_c$, for all $c \in C$. I am not sure which option is to be preferred in the long run.⁴⁷

Of course, our assumptions about the common ground $\Gamma_{c^*}$ are crucial here. For instance, in a context in which the speaker is not known to be an Expert and for which $(g_e)$ thus fails, $(j_e)$ and $(k_e)$ should not hold. Neither should the following conclusion, which I take to be the main point of the present subsection:

$$(39) \quad \Gamma_{c^*} \cap \{c \in C \mid \Pi_{c^*} \subseteq \Delta(25)\} = \Gamma_{c^*} \cap \Delta_{\mathfrak{E}}(32).$$

In other words, asserting (32) would have had the same communicative effect in $c^*$ as counterfactually asserting (25) actually had.

I take it that Saul’s Superman example can be treated along the same lines, trading Experts for Enlightened speakers and Laymen for Ignorami. But, as already discussed in the Section 4.3, historical place names call for a more radical revision (if their use is a matter of counterfactual speech acts at all). For although it may seem plausible to reconstruct the purely temporal condition (3b') given in Section 4.3 in terms of a purely temporal context shift, the assumption that $w_{c^*} = w_c$ whenever $L_c = \mathfrak{P}$ would play havoc with subjectivity: the $\mathfrak{P}$-speaker’s actual doxastic perspective will have changed since the days of Olden (if she is old enough in the first place), but even when uttering a sentence of $\mathfrak{P}$ she will perform her counterfactual speech act on the basis of her present perspective. To account for this constancy of the perspective, one would have to assume that $\Sigma_{c^*} = \Sigma_c$ (at least for the kind of contexts $c$ considered above) and see to it that the temporal parameter coincides with the S-Phase (of $w_{c^*}$ and $w_c$, which I suppose would be the same). I am not at all sure how all this tampering with the revision function $+$ could be motivated independently and whether it would ultimately lead to the same – apparently correct – predictions of the semantic account given in Section 2.

There are some general issues that the formalization leaves open. In particular, I have been rather vague on questions of identifying languages and expressions. As a case in point, it is not obvious, which facts determine what the objective language spoken in a given context is. Is it, e.g., possible to have two contexts $c$ and $c'$ that only differ in their objective language? As far as I can see, it is not easy to give this question a precise formulation, let alone answer it in a principled way; but the whole formal framework is incomplete as long as this and related questions are left open. Similarly, it is not clear under which circumstances two (disambiguated) expressions of different languages should count as being identical. In principle, one should be prepared to identify such diverse forms as French Louis-quatorze and German Ludwig der Vierzehnte.

Let me finally point out an inadequacy in the above account of the common ground $\Gamma_{c^*}$. Condition $(g_{c^*})$, which had been adapted from the corresponding assumption $(g)$ about $c^*$, smuggles the objective language $\mathfrak{E}$ into the (inter-) subjective beliefs. More specifically, $(g_{c^*})$ implies that speaker and hearer agree on what the objective language is, which – given minimal assumptions about the connection between perspectives and common

⁴⁷ My impression is that, starting with Gazdar (1979), there has been a certain preference for epistemic attitudes in the literature on formal pragmatic reasoning, with doxastic perspectives only being considered more recently, witness Gunlogson (2003).
grounds\textsuperscript{48} – implies full identifiability of the objective language: $\Sigma_{c^*} = \{\mathcal{E}\}$. This consequence of $(i_r)$ certainly goes against the spirit of the enterprise, according to which speakers in general only have incomplete knowledge of their language. Hence a more adequate account would have to replace $(i_r)$ by a number of conditions on the language spoken that suffice to derive (39) without necessarily fully identifying it as $\mathcal{E}$. Of course, the pertinent conditions are essentially those laid out in (a) – (f), which would have to be reformulated in turn, having only been given for fully specified languages. Though such a reformulation can be given, it turns out to be rather tedious, which is why I have opted for the more transparent formulation above.

5. Concluding Remarks
The upshot of this comparison of two radically different approaches to extensional substitution puzzles is that neither of them is perfect. However, whereas the semantic account fares well with the temporal examples and leads to conceptual difficulties when it comes to double identities, the pragmatic account offers a (to my mind) plausible strategy of interpreting the latter but would have to be unduly stretched if it were to cover the former in an adequate way. One may speculate whether this shows that the initially obvious parallel between changing places and double personalities is not as close as it would seem. However, I still believe it is because I do not see a principled reason why changed names should invariably be encoded in semantics, whereas the use of hidden names are a matter of pragmatics.

Let me end this somewhat inconclusive comparison with a classical puzzle which surprisingly does not find a simple and natural solution in either of the two accounts:

(23) Clark Kent is Superman.
(37) Barbara Vine is Ruth Rendell.
(40) Hesperus is Phosphorus.

In Section 3.2 I argued that the semantic account cannot explain why whoever may utter (23) on whatever occasion would have said something true. But it seems that neither can the pragmatic account\textsuperscript{49}. To be sure, if the sentence is uttered by an Ignoramus and taken literally, it always expresses a (necessarily) true proposition. However, if it is uttered by an Enlightened speaker, it would have to be taken as a counterfactual speech act and as expressing the subjective perspective of Ignoramus, who takes it to be necessarily false; the communicative effect would thus be disastrous. And, of course, the same goes for the other two of the above examples. I think there is a natural way out of this embarrassment. Any truly informative use of an identity statement like (23) would Enlighten an Ignoramus. In such a conversation between unequal partners, the common ground would have to be undecided as to the coreferentiality of the two names (in $\mathcal{I}$). And although the Enlightened speaker uses a sentence of $\mathcal{I}$, the Ignoramus would not think this odd and would therefore not have any reason to reinterpret it. Rather, he would take it at face-value, i.e. in its literal meaning. Given the above treatment of ordinary assertions, he would thus learn that, according to the reliable speaker, the two names are coreferential, thus calling for a revision of his own beliefs – but not of the pragmatic approach to non-substitutivity.

\textsuperscript{48} In particular, if the common ground is identified with mutual belief in the sense of Schiffer (1972), then $\Pi_c \subseteq \Gamma_{c^*}$, which implies that $\Sigma_{c^*} \subseteq \{L_c \mid c \in \Gamma_{c^*}\}$. And even if, as Stalnaker (2002: 704f.) pointed out, this identification in general is rash, it is still arguable that it holds for a context like $c^*$.\textsuperscript{49} I am grateful to Robert van Rooy for bringing up Frege’s (1892) puzzle in this connection.
References


