SUBSTITUTIVITY AND SIDE EFFECTS

Graeme Forbes

1. Introduction

A major thread that runs through Ruth Barcan Marcus's collection of papers *Modalities: Philosophical Essays*,[1] is the ongoing dialectic that she has pursued with Quine over the question of the intelligibility of modal language, particularly *de re* modal language (see Marcus 1990a for a summation). In earlier work by Quine (e.g., Quine 1955), the main symptom of the unintelligibility of *de re* modal language is said to be the failure of coreferential "singular terms" to interchange *salva veritate* within the scope of modal operators. From this it is supposed to follow that the notion of objectual satisfaction is inapplicable to *de re* formulae, hence "quantifying-in" makes no sense. A response that was once favored by Marcus is to reconstrue the semantics of quantification substitutionally, as in (Marcus 1961). But as she had already pointed out (Marcus 1948), Smullyan (1948) had anyway demonstrated that Quine's premise, that the rule of Identity Elimination (=E) fails in modal languages, is incorrect. Quine's later "animadversions" about *de re* modality (e.g., in Quine 1966), accuse it of commitment to an invidious "Aristotelian essentialism", but work by Marcus and Parsons (Marcus 1967; Parsons 1969) shows that this charge has little force.

1. All page references to Marcus's writings are to this volume, (Marcus 1993).

And there matters rest so far as *de re* modality is concerned. But of course, this is not the whole story. For the argument that where =E fails, quantifying-in makes no sense, is still of relevance to other cases in which =E **does** seem to fail. First, there a Quinean paradigm of failure of substitutivity (Quine 1961:22),

(1) Giorgione is so-called because of his size; Giorgione is Barbarelli; therefore, Barbarelli is so-called because of his size.

Here, a use of =E leads us astray. In addition, quantifying-in produces something **uninterpretable** (that is, no complete proposition can be assigned to the quantified sentence, though the premise from which it is derived by Existential Introduction (∃I) is a complete proposition):

(2) Someone is such that he is so-called because of his size.

He cannot be a pronoun bound by the initial *someone* and **also** an indexical or a name. More carefully, a standard understanding of English excludes this (we could devise some conventions that would make (2) artificially interpretable, as we could for any nonsense string). But *he* would have to be a name to be a suitable antecedent for *so-called*. (2) is true if interpreted substitutionally, but it is a bug, not a feature, of the substitutional interpretation of *someone is such that* ϕ that it assigns a truth-value to (2).

In another example (Fine 1989:222–3; see also Linsky 1967:104),

(3) The man behind Fred saw him leave; the man behind Fred = the man in front of Bill; therefore, the man in front of Bill saw him leave

we again find both failure of =E and uninterpretable results from quantifying-in:

(4) Someone is such that he saw him leave.

Syntax excludes interpreting *him* as coindexed with *he* and *someone* (Chomsky's Condition B, that a non-reflexive pronoun must be free in its governing category), which means that, barring a surreptitious conversion of its semantics from anaphoric to deictic, *him* can only be a variable that is free in (4).

However, there is at least one case where we have the combination of (apparent) failure of =E with the acceptability of quantifying-in, the case of attitude ascriptions. For although

(5) Lex fears Superman; Superman = Clark; therefore Lex fears Clark

seems wrong, we would not object to

(6) Lex fears Superman; therefore, someone is such that Lex fears him

on the grounds that the conclusion is **uninterpretable**.[2]

One response to these data is to accept that the illustrated uses of =E are all as incorrect as they appear, while of the three uses of ∃I displayed in (2), (4) and (6), the one in (6) involves a special maneuver that explains its validity. I will develop this response in the rest of the paper, initially against the backdrop of an attempt by Marcus to formulate =E with the right restrictions to block incorrect uses.

2. Cartwright and Marcus on =E

As discussed in (Marcus 1975), Cartwright (1971) accepts that (1) is a counterexample to =E, but distinguishes this rule of inference from Leibniz's Law, which may be formulated as *if a = b then every property of a is a property of b* or as *if a = b then whatever is true of a is true of b*. Cartwright holds that the Law is not threat-

2. We might object to the inference on the grounds that it invalidly generates existential commitment where there is none in the premise. I object to this objection in (Forbes 1996:357–362).

ened by (1), because the expression *is so-called because of his size* does not express a property, or a condition that can be true or false of things. Marcus is sceptical that the fate of the inference-rule can come apart from that of the Law so easily (1975:103), but it seems to me that we can make a stronger objection, namely, that the Law does no better with (1) than does the inference rule. For it is perfectly fine to attribute to Giorgione the property of being so-called because of his size; equally, it is unproblematically true of Giorgione that he is so-called because of his size. Moreover, Giorgione is Barbarelli. Yet Barbarelli does **not** have the property of being so-called because of his size; equally, it is **not** true of Barbarelli that he is so-called because of his size.

Marcus's own proposal is that the rule of inference is only applicable to logical forms, in which various kinds of ambiguity to which ordinary discourse is subject have been eliminated (1975:105). She points out that

(7) Giorgione is so-called because of his size

could mean that Giorgione is called 'Shorty' because of his size if it immediately follows an utterance of *Shorty is so-called because of his size*. She then proposes the following formulation of =E, improving on Cartwright's version (*op. cit.* p.108):

(8) For all proper names α and β (indexed to preserve univocality), $\ulcorner \alpha = \beta \urcorner$ expresses a true proposition just in case, for all sentences P, S and S', if S is a restatement of P in logical form and if S' is like S save for containing an occurrence of β where S contains an occurrence of α, then S expresses a true proposition only if S' does.

However, while the requirement that the rule be applied only to logical forms can deal with problem cases in which substitution affects syntactic structure, (8) does not seem to me to make much headway with (1) and (3).[3] The minor premise of (3) may be written

(9) The man behind Fred$_i$ saw him$_i$ leave

using indices to mark anaphoric relations (assume there is only one Fred and only one Bill). If we wish to replace *the man behind Fred* with the coreferential *the man in front of Bill*, we either have to use the same index *i* on *Bill* or a different one. If we use *i*, we get a falsehood if the man in the middle did not see Bill leave; if we use another index, we get something uninterpretable ('him$_i$' would be a free variable); and if we use no index, *him* becomes deictic.[4] (8) fails to help us here.

Does (8) do better with (7)? Though Marcus is not explicit, her point about the context-sensitivity of *so* and the context-insensitivity of logical forms indicates that at logical form, she would eliminate *so-called* in favor of *called NN*, where *NN* is the expression the context provides by which the relevant entity is called. This allows free use of =E; for example, we can substitute the sole occurrence of *Giorgione* used as a name of big Giorgio in *Giorgione is called* Giorgione *because of his size*.

But this proposal about the logical form of (7) seems unacceptable to me, as does its analogue for (9), which requires *him$_i$* to be replaced at logical form by *Fred*. It belongs to the semantics of an anaphoric pronoun that its semantic value is recovered in the course of interpretation from an autonomously referring ex-

3. For structure-affecting substitutions, see (Fine 1989:III). I assume substitution replaces all and only the contents of some node in the minor premise's parse tree.

4. If we treat *the* as a **quantificational** determiner, the inference in (3) becomes: (a) (*the x: man*(*x*) ∧ *behind*(*x*, *Fred$_i$*))[*saw*(*x*, *y$_i$ leave*)]; (b) (*the x: man*(*x*) ∧ *behind*(*x*, *Fred*))[(*the z: man*(*z*) ∧ *in front of*(*z*, *Bill*))[*x* = *z*]; therefore (c) (*the z: man*(*z*) ∧ *in front of*(*z*, *Bill$_i$*))[*saw*(*z*, *y$_i$ leave*)], assuming minimally that anaphoric links to **positions** must be maintained in a correct inference. This inference instantiates a schema that is valid for extensional languages if no special allowance is made for anaphoric links, though exactly how it involves use of =E will depend on the precise details of the inference system for restrictive quantifiers and *the*. The quantificational treatment of *the* will either return the same puzzle, or a new one that can be resolved only with similar resources.

pression that anchors it, a process which is simply missed by the replacement proposal. The semantics of *so* is equally missed by a replacement proposal. *So* is a demonstrative: (7) means *Giorgione is called by* ***that*** *name because of his size*. We should therefore expect a formal semantics for a language with the substitution-blocking *so* device to use apparatus of a kind standardly deployed in the semantics of demonstratives, e.g., the LD framework of (Kaplan 1989b).

Despite these objections to Marcus's proposed version of =E, I agree with her that the rule is correct if formulated in a way that is appropriate for the expressive resources of the language in question. Our two examples of misapplication of =E indicate that what we need is a restriction to the effect that a substitution may be made in $S(a)$ using the major premise $a = b$ so long as the substitution has no **side-effects**. The side-effects in our fallacious substitutions are, in (1), that the referent of *so* is altered, and in (3), that the antecedent of *him* is changed. But what side-effects should a restriction on =E mention? Non-trivial substitutions[5] in natural language usually have typographic side-effects, such as changing the total number of letters in $S(a)$. At the other extreme, we could count loss of truth as a side-effect, but **any** unsound rule would be saved by a no-side-effects restriction that counted this as a side-effect. We will try to sharpen the notion of side-effect after a closer look at the semantics of our examples.

3. Logocentric demonstration

Since we wish to focus on the semantics of *so-called*, it is a problem that the examples that use it also include *because of*, about whose analysis there is little agreement. I will work round this by considering a formal language $\mathcal{L}_\gamma$ which con-

5. A trivial substitution is one which uses something of the form $a = a$ as the major premise.

tains a three-place predicate $\gamma(x, t, z)$ to be read *x is called t for z*, as in *Giorgione is called* Giorgione *for his size* and *Giorgione is called so for his size* (I assume *so-called* is derived from *called so*). I propose to ignore any extensionality problems this may give rise to. t is either *so*, or for some syntactically simple constant c, t is a quotation name $\underline{c}$ of c. $C_{\mathcal{L}_\gamma}$ is the set of individual constants of $\mathcal{L}_\gamma$ and includes $\underline{c}$ for each syntactically simple c.

$\mathcal{L}_\gamma$ is otherwise a standard first-order language. A **structure** $\mathcal{A}$ for $\mathcal{L}_\gamma$ is a pair $(\mathcal{D},\mathcal{V})$ where (i) $\mathcal{D}$ is a non-empty domain perhaps containing members of $C_{\mathcal{L}_\gamma}$ that are not quotation names; and (ii) $\mathcal{V}$ is a valuation function subject to the special requirement that $\langle x, c, z\rangle \in \mathcal{V}(\gamma)$ only if $\mathcal{V}(c) = x$ (Giorgione can be called *Giorgione* for some reason only if *Giorgione* refers to Giorgione).

An $\mathcal{L}_\gamma$-**discourse** Δ is a sentence or set of sentences of $\mathcal{L}_\gamma$ in which occurrences of *so* are distinctively numbered, and for each individual constant c, occurrences of c are distinctively numbered (numbering is just a convenience).

A **context** μ is a function from a set of positive integers (perhaps empty) to a set of pairs each of the form $\langle j,\alpha\rangle$, where j is a positive integer and α is an individual constant of $\mathcal{L}_\gamma$. $\mu(i) = \langle j,\alpha\rangle$ means that the occurrence of *so* numbered i in the sentence or discourse being evaluated refers to the occurrence of the name α numbered j. If $\mu(i) = \langle j,\alpha\rangle$ we define $\mu_1(i) = j$ and $\mu_2(i) = \alpha$.

A context will be unsuitable for a structure if the former assigns to some *so* a denotation which is not in the domain of the latter. Therefore we say a structure $\mathcal{A}$ is **suitable** for a context μ iff, $\mu_2(i) = \alpha$ for some i only if $\alpha \in \mathcal{D}_{\mathcal{A}}$.

A context μ will not match a discourse Δ if there are more *so*'s in Δ than μ is defined for or if some *so* is assigned a reference that does not occur in Δ. Therefore we say that a context μ is **defined for** an $\mathcal{L}_\gamma$-discourse Δ iff (i) μ is defined for i if there is a *so* in Δ numbered i, and (ii) if $\mu(i) = \langle j,\alpha\rangle$ and there is an occurrence of *so* in Δ numbered i, then there is an occurrence of α in Δ numbered j.

A context μ for an $\mathcal{L}_\gamma$-discourse may be displayed using arrows that link occurrences of *so* to the expressions they demonstrate in the discourse. For example, for the conditional

(10.1) If Caravaggio was called so for his birthplace and Guercino was called so for his squint, then Guercino was called so for his squint and Caravaggio was called so for his birthplace

or in $\mathcal{L}_\gamma$,

(10.2) $\gamma(C^{(1)}, so^{(1)}, his_C\ birthplace) \wedge \gamma(G^{(1)}, so^{(2)}, his_G\ squint) \rightarrow$
$\gamma(G^{(2)}, so^{(3)}, his_G\ squint) \wedge \gamma(C^{(2)}, so^{(4)}, his_C\ birthplace)$

there are various contexts available. The most likely is the context μ^* defined by

(11) $\mu^*(1) = \langle 1, C\rangle; \mu^*(2) = \langle 1, G\rangle; \mu^*(3) = \langle 2, G\rangle; \mu^*(4) = \langle 2, C\rangle.$

We can display μ^* by decorating (10.1) and (10.2) as follows:

(12.1) If Caravaggio was called so because of his birthplace and Guercino was called so because of his squint, then Guercino was called so because of his squint and Caravaggio was called so because of his birthplace.

(12.2) $\gamma(C, so, his_C\ birthplace) \wedge \gamma(G, so, his_G\ squint) \rightarrow$
$\gamma(G, so, his_G\ squint) \wedge \gamma(C, so, his_C\ birthplace).$

With contexts displayed explicitly, there is no need to write in superscripts on the occurrences of *so* or the individual constants.

The semantics of $\mathcal{L}_\gamma$ is given by a recursive definition of the concept

$$(\mathcal{A}, \mu) \Vdash_h \phi$$

read as "ϕ is true in the structure $\mathcal{A}$ and context μ relative to the assignment h". Here ϕ is an $\mathcal{L}_\gamma$-wff, and h is a partial function into $\mathcal{D}$ defined for all the variables free in ϕ; $\mathcal{A}$ is suitable for μ; and μ is defined for $\Delta = \{\phi\}$. If ϕ is a formula in which a specific context is displayed, as in (12.2), we may speak of its truth-value (relative to h) in $\mathcal{A}$ *simpliciter*.

The clauses of the semantics are as would be expected, μ being appealed to only to give the reference of an occurrence of *so*. Following Kaplan (1989b:544–6), we define a general notion of *content* in a structure $\mathcal{A}$ and context μ relative to an assignment h, which appeals to $\mathcal{V}_\mathcal{A}$ for individual constants, to h for free variables, and to μ for *so* (I postpone the details of the treatment of descriptions):

(13.1) $[\![c^{(k)}]\!]_{\mathcal{A},\mu,h} = \mathcal{V}_\mathcal{A}(c)$, c a syntactically simple constant with optional superscript; $[\![\underline{c}]\!]_{\mathcal{A},\mu,h} = \mathcal{V}_\mathcal{A}(\underline{c}) = c$;

(13.2) $[\![so^{(i)}]\!]_{\mathcal{A},\mu,h} = \mu_2(i)$;

(13.3) $[\![v]\!]_{\mathcal{A},\mu,h} = h(v)$, v an individual variable for which h is defined.

The clauses for atomic and complex formulae may now be stated as:

(14.1) if $\phi(t_1,\ldots,t_n)$ is an atomic wff of $\mathcal{L}_\gamma$, where each t_i is either a variable or a constant or an occurrence of *so*, then $(\mathcal{A}, \mu) \Vdash_h \phi(t_1,\ldots,t_n)$ iff: $\langle [\![t_1]\!]_{\mathcal{A},\mu,h},\ldots,[\![t_n]\!]_{\mathcal{A},\mu,h}\rangle \in \mathcal{V}(\phi)$;

(14.2) $(\mathcal{A}, \mu) \Vdash_h \neg\phi$ iff $(\mathcal{A}, \mu) \nVdash_h \phi$;

(14.3) $(\mathcal{A}, \mu) \Vdash_h \phi \wedge \psi$ iff $(\mathcal{A}, \mu) \Vdash_h \phi$ and $(\mathcal{A}, \mu) \Vdash_h \psi$;

(14.4) $(\mathcal{A}, \mu) \Vdash_h$ (*some/every* v*:* ψv)$[\phi v]$ iff for some/every x in $\mathcal{D}$ such that $(\mathcal{A}, \mu) \Vdash_{h\hat{}(x\mapsto v)} \psi v$, we also have $(\mathcal{A}, \mu) \Vdash_{h\hat{}(x\mapsto v)} \phi v$.

In (14.4), $h\hat{}(x\mapsto v)$ is the assignment that extends h by assigning $x \in \mathcal{D}$ to v (the syntax does not allow two quantifiers with the same variable in a single sentence if one is within the scope of the other). The truth of a closed $\mathcal{L}_\gamma$-sentence σ in a structure $\mathcal{A}$ and context μ is defined by

(15) $(\mathcal{A}, \mu) \Vdash \sigma$ iff $(\mathcal{A}, \mu) \Vdash_{\langle\rangle} \sigma$.

That is, truth in a structure and context is truth relative to the empty assignment $\langle\rangle$. As quantified subformulae are evaluated in the course of deriving a truth-condition, the empty assignment is extended by (14.4) to those variables for which a value is needed.

This leaves us to define validity and consequence. Sentences (10.1) and (10.2) have the appearance of logical truths, since, ignoring superscripts, they seem to be instances of $(\phi \wedge \psi) \rightarrow (\psi \wedge \phi)$. But with the *demonstrata* of their occurrences of *so* unspecified, they are no more logical truths than ϕ(*that*) → ϕ(*that*) is in Kaplan's system (nothing indicates that the two occurrences of *that* are coreferential – see the discussion in Kaplan 1989a:587). However, with the context μ^* defined in (11) provided to fix the reference of the *so*'s, it is clear that changes in the interpretation of the remaining non-logical vocabulary will not affect truth-value; that is, sentence-type (10.2) expresses a logically true proposition in the context μ^* (there are other contexts in which it does not even express a truth). So one notion of validity is that of a sentence which is **valid in a context**:

(16) $\vDash_\mu \sigma$ iff μ is defined for $\{\sigma\}$ and for every structure $\mathcal{A}$ suitable for μ, $(\mathcal{A}, \mu) \Vdash \sigma$.

But there are also sentences whose validity is not sensitive to choice of context, particularly *so*-free sentences. For these we have the stronger concept of **universal** validity:

(17) $\vDash \sigma$ iff for every context μ defined for $\{\sigma\}$, $\vDash_\mu \sigma$.

A *so*-free sentence is universally valid iff it is a logical truth of conventional first-order logic.

These are adaptations of the two notions of validity for a language with a device analogous to *so-called* proposed in (Forbes 2000a), and we can assess (1) using them. First we convert it to a single sentence, *if Giorgione is so-called for his size and Giorgione is Barbarelli, then Barbarelli is so-called for his size*, then translate into $\mathcal{L}_\gamma$:

(18) $\gamma(G^{(1)}, so^{(1)}, his_G\ size) \wedge G^{(2)} = B^{(1)}) \rightarrow \gamma(B^{(2)}, so^{(2)}, his_B\ size)$.

(18) is valid in the context μ^* defined as $\mu^*(1) = \langle 1, G\rangle$, $\mu^*(2) = \langle 1, G\rangle$ (i.e., both occurrences of *so* refer to the initial *Giorgione*), since in this context, substitution has no unwanted side-effects. But the "intended" context is rather $\mu^\dagger$, defined by $\mu^\dagger(1) = \langle 1, G\rangle$, $\mu^\dagger(2) = \langle 2, B\rangle$, and (18) is not valid in $\mu^\dagger$, as the actual situation shows. Hence (18) is not universally valid.

We could get the same result by defining a companion notion of semantic consequence:

(19) $\Sigma \vDash_\mu \sigma$ iff μ is defined for $\Delta = \Sigma \cup \{\sigma\}$ and for every $\mathcal{A}$ suitable for μ, if $(\mathcal{A}, \mu) \Vdash \tau$ for every $\tau \in \Sigma$, then $(\mathcal{A}, \mu) \Vdash \sigma$.

This requires us to take the premises and conclusion of (1) as a single discourse for which μ is defined, with *so* having two occurrences. Displaying the context $\mu^\dagger$ of the previous paragraph, we will then have

(20) $\gamma(G, so, his\ size), G = B \nvDash \gamma(B, so, his\ size)$

(without the display, replace $\nvDash$ with $\nvDash_{\mu^\dagger}$).

It is now clear that the failure of quantifier introduction on G in the minor premise of (20) has nothing to do with G or its position being "not purely referential". The semantics of individual constants in the framework we have just set out is entirely standard. Nor is the problem just that $\mathcal{V}(\textit{so})$ is stipulated to be a subset of $\mathcal{D} \times C_{\mathcal{L}_\gamma} \times \mathcal{D}$, since this is consistent with

(21) (*some x: person*(*x*))[γ(*x*, *so*, *his size*)]

being logically false. (21) should rather come out uninterpretable. One way of achieving this would be to write the syntax of $\mathcal{L}_\gamma$ so that γ(*x*, *so*, *his size*) is not well-formed; but it is perfectly well-formed – trouble only arises when a *so* tries to link to the x. The best option seems to be to modify (13.2), which defines the content of $\textit{so}^{(i)}$ in a context μ and structure $\mathcal{A}$. Instead of setting $[\![\textit{so}^{(i)}]\!]_{\mathcal{A},\mu,h}$ equal to $\mu_2(i)$, we make the assignment of a content conditional:

(22) $[\![\textit{so}^{(i)}]\!]_{\mathcal{A},\mu,h} = \mu_2(i)$ if $\mu_2(i) \in C_{\mathcal{L}_\gamma}$ and $\mu_2(i) \neq \underline{c}$ for some $c \in C_{\mathcal{L}_\gamma}$; $[\![\textit{so}^{(i)}]\!]_{\mathcal{A},\mu,h}$ is undefined otherwise.

(22) by itself changes nothing, since we already stipulated that $\mu_2(i)$ is in $C_{\mathcal{L}_\gamma}$ and is not a $\underline{c}$, but if we are not going to forbid language-users from **attempting** to link *so* illegitimately, we should accommodate such intentions in the syntax and pragmatics and let the semantics carry the big stick. So let us redefine a context to be a function μ whose range is a set of pairs $\langle j,\alpha\rangle$, where j is a positive integer and α is **any well-formed expression** of $\mathcal{L}_\gamma$. Then in conjunction with (22), leaving everything else as it is, the derivation of a truth-condition for (21) "crashes" when it tries to call the value of *so* and finds it undefined.

Explaining why (21) cannot be inferred from the minor premise of (20) is part of the more general project of providing an adequate system of inference for which we will have such results as:

(23.1) $\gamma(G, \textit{so}, \textit{his size}), G = B \;\nvdash\; \gamma(B, \textit{so}, \textit{his size})$

(23.2) $\gamma(G, \textit{so}, \textit{his size}), G = B \;\vdash\; \gamma(B, \textit{so}, \textit{his size})$

(23.3) $\gamma(G, \textit{so}, \textit{his size}) \;\nvdash\; (\exists x)\gamma(x, \textit{so}, \textit{his size})$.

(23.4) $\gamma(G, \textit{so}, \textit{his size}) \;\vdash\; (\exists x)\gamma(x, \textit{so}, \textit{his size})$.

The most natural notion of proof for this consequence relation is one in which a proof is a list of sequents of the form $\Gamma \vdash \sigma$, in which σ and every member of Γ is a sentence-in-context, as illustrated in (23.1)–(23.4). Intro and Elim rules apply to the sentence-in-context on the right of $\vdash$, with appropriate adjustments to the premise-assumption set on the left. The Elimination rule for *so* is:

(24)
$$\frac{\Gamma \;\vdash\; \ldots\gamma(t_1, \textit{so}, t_2)\ldots \quad [c]}{\Gamma \;\vdash\; \ldots\gamma(t_1, \underline{c}, t_2)\ldots}$$

which says that if *so* in a subformula $\gamma(t_1, \textit{so}, t_2)$ of the conclusion formula **targets** an occurrence of c in the conclusion formula or in any member of Γ, then a quotation-name for c (e.g., ⌜'c'⌝) may be substituted for that *so*;[6] t_1 and t_2 are any singular terms.

The Introduction rule is the inverse:

(25)
$$\frac{\Gamma \;\vdash\; \ldots\gamma(t_1, \underline{c}, t_2)\ldots}{\Gamma \;\vdash\; \ldots\gamma(t_1, \textit{so}, t_3)\ldots \quad [c]}$$

If c occurs somewhere in the premise-conclusion discourse outside the indicated

6. A notational variant of (24) labels each line with a context and deletes the appropriate pair from the context each time *so*-E is applied.

subformula $\gamma(t_1, \underline{c}, t_2)$, then the occurrence of $\underline{c}$ in this subformula may be replaced by *so* provided the latter targets one of those other occurrences of *c*. No 'so' in a premise can target a 'so' in a different premise. Proofs are subject to the global condition that if π is an assumption containing an occurrence of a constant *c* which some *so* targets, then π may not be discharged by $\vee$E or $\exists$E, and, in an obvious sense, the link must "follow" π if π is discharged by $\rightarrow$I; for $\neg$I, we stipulate that neither major nor minor premise may contain 'so'.[7]

=E and quantifier introduction rules are the only primitive rules that replace individual constants. Assuming that this is the **only** way that quantifier introduction works (that is, there are no "arbitrary names" – see Forbes 1993 for justification of this), then we must require that a name can be substituted in =E or replaced by a variable in a use of a QI rule only if it is untargeted. Thus the connection between substitution-resistance and the uninterpretability of quantifying-in is quite simple in this case: the device that prevents substitution is the very same one that blocks quantifier introduction.

To end this part of our discussion, here are proofs of (23.2) and (23.4):

(26.1) $\gamma(G, \textit{so}, \textit{his size}) \vdash \gamma(G, \textit{so}, \textit{his size})$ Premise

(26.2) $G = B \vdash G = B$ Premise

(26.3) $\gamma(G, \textit{so}, \textit{his size}) \vdash \gamma(G, \text{'}G\text{'}, \textit{his size})$ 1 *so*-E

(26.4) $\gamma(G, \textit{so}, \textit{his size}), G = B \vdash \gamma(B, \text{'}G\text{'}, \textit{his size})$ 2,3 =E

(26.5) $\gamma(G, \textit{so}, \textit{his size}), G = B \vdash \gamma(B, \textit{so}, \textit{his size})$ 4 *so*-I

7. The obvious sense can be made precise using the notation mentioned in note 6. We also require that $\rightarrow$E and $\wedge$E cannot be applied if the antecedent or eliminated conjunct is linked to by the consequent or retained conjunct. But we can use *so*-E and *so*-I to work round this.

As this proof illustrates, the rules for *so* are context-manipulation rules, so that context can change from line to line in a proof. Another example:

(27.1)	$\gamma(G, \textit{so}, \textit{his size})$	$\vdash$	$\gamma(G, \textit{so}, \textit{his size})$	Premise
(27.2)	$\gamma(G, \textit{so}, \textit{his size})$	$\vdash$	$\gamma(G, \text{‘}G\text{’}, \textit{his size})$	1 *so*-E
(27.3)	$\gamma(G, \textit{so}, \textit{his size})$	$\vdash$	$(\exists x)\gamma(x, \text{‘}G\text{’}, \textit{his size})$	2 ∃I
(27.4)	$\gamma(G, \textit{so}, \textit{his size})$	$\vdash$	$(\exists x)\gamma(x, \textit{so}, \textit{his size})$	3 *so*-I

Since context can change from line to line, the appropriate notion of soundness for these rules is not that of preserving validity relative to a fixed context, but rather, relative to a variable one. More precisely, where μ is defined for $\Sigma \cup \{\sigma\}$ and μ' is defined for $\Sigma' \cup \{\sigma'\}$, we require of each rule R that if $\Sigma' \vdash_{\mu'} \sigma'$ may be inferred from $\Sigma \vdash_{\mu} \sigma$ using R, then if $\Sigma \vDash_{\mu} \sigma$, then $\Sigma' \vDash_{\mu'} \sigma'$. The rules for connectives are all sound, since in their use, $\mu = \mu'$. And the soundness of the *so*-rules is easy to demonstrate. So the system is sound, since every proof begins with a sequent of the form $\sigma \vdash \sigma$, where σ is a sentence in context.

4. Argument anaphora

Our other example of a misuse of =E,

> (3) The man behind Fred saw him leave; the man behind Fred = the man in front of Bill; therefore, the man in front of Bill saw him leave

will detain us less, since the apparatus we deploy to deal with it, deriving from (Evans 1977), is similar to what we developed in the previous section, and does not employ indexing. On this account, the anaphoric pronoun *him* in the minor premise or conclusion of (3) is assigned content by inheritance from an argument expression functioning as its antecedent. We can use arrows in the same way as in §3 except that whereas the content of *so* is the very expression which it targets, the content of an anaphoric pronoun is the **content** of the expression which it targets. If we number occurrences of antecedent expressions and anaphoric pronouns, we can describe which pronoun targets which antecedent in a function μ as before. But in this case we would not regard such a function as defining a **context of evaluation**; instead, it is part of the semantic structure of sentences with anaphoric pronouns: the arrows disambiguate ambiguities of back-reference in ordinary English much as the ordering of quantifiers at LF disambiguates scope-ambiguities in ordinary English.

A pronoun may link to a proper name, a variable, or a definite description, so we must stop being coy about the treatment of definite descriptions. For the purposes of this paper, we can follow Frege in taking them to be singular terms which give rise to presupposition-failures if they are improper. And let us say that a presupposition-failure generates a truth-value gap in atomic formulae that is inherited by any formula that includes an undefined atomic one.[8] We stipulate

(28) $[\![\textit{the } v\text{: } \phi]\!]_{\mathcal{A},\mu,h}$ = the unique $a \in \mathcal{D}_{\mathcal{A}}$ such that $(\mathcal{A}, \mu) \Vdash_{h^\wedge(a \mapsto v)} \phi$ if there is such an a; $[\![\textit{the } v\text{: } \phi]\!]_{\mathcal{A},\mu,h}$ is undefined otherwise.

Then if $\rho^{(i)}$ is a pronoun-occurrence linked to an antecedent $\mu(i)$, the rule for $\rho^{(i)}$ is

(29) $[\![\rho^{(i)}]\!]_{\mathcal{A},\mu,h} = [\![\mu_2(i)]\!]_{\mathcal{A},\mu,h}$ if $[\![\mu_2(i)]\!]_{\mathcal{A},\mu,h}$ is defined, $[\![\rho^{(i)}]\!]_{\mathcal{A},\mu,h}$ is undefined otherwise.

The other cases, $[\![c]\!]_{\mathcal{A},\mu,h}$ and $[\![v]\!]_{\mathcal{A},\mu,h}$, are defined in (13.1) and (13.3).

Returning to (3), the pronouns target antecedents in the way displayed below:

8. This is controversial. See (Soames 1989:560) for discussion.

(30) The man behind Fred saw him leave; the man behind Fred = the man in front of Bill; therefore, the man in front of Bill saw him leave.

These links do not block substitution of the targeted term, since the content of the pronoun is the **content** of the term, not, as in the previous case, the term itself. So if Fred = Jed, then the man behind Jed saw him leave. The problem cases, as the example illustrates, arise when terms are embedded in complex terms. The same goes for quantifying in. We can quantify into the description in the minor premise of (30), which would produce English like *there is someone the man behind whom saw leave* or more colloquially, *the man behind someone saw him leave*. In our current formalism, this is

(31) (*some x: person*(*x*))[*the man behind x saw him leave*].[9]

In evaluating (31), *x* receives as a value some element of $\mathcal{V}$(*person*) for each test of the scope-formula, and by (29) that value becomes the content of *him*. What we cannot do is quantify away the entire description, since this produces

(32) (*some x: person*(*x*))[*x saw him leave*].
?

Him's antecedent *Fred* has simply disappeared. Like (21), then, (32) should be uninterpretable, and evaluation in a structure should crash on account of the missing $[\![him]\!]_{\mathcal{A},\mu,h}$. By "crash" we mean that the semantic rules do not permit any of

9. Allowing anaphoric pronouns to link to bound variables in this way effects a disentangling of Evans' views about the relationship between pronouns and variables ("traces") from his quasi-substitutional semantics for quantifiers; see (Larson and Seagal 1995:380–382). However, it means that we are positing the **same** mechanism to cover cases that are often distinguished as *bound variable anaphora* and *referential anaphora*. We therefore incur the obligation to explain phenomena such as VP ellipsis and *do so* anaphora that are standardly explained by appeal to this distinction.

the three possible verdicts about (32) (*true*, *false*, *undefined*) to be reached. This draws a contrast between (32) and sentences with descriptions that are improper in $\mathcal{A}$, since the evaluation rules permit, indeed require, the verdict *undefined* to be reached for sentences of the latter sort. So it is insufficient for the uninterpretability of (32) in $\mathcal{A}$ relative to μ and h that $[\![him]\!]_{\mathcal{A},\mu,h}$ is undefined – it is also undefined in *the woman talking to the man drinking a martini likes him*, if no-one is drinking martinis; and (29) is formulated to allow for this. But (29) simply assumes that $\mu_2(i)$ is defined (note: assumes that $\mu_2(i)$, not $[\![\mu_2(i)]\!]_{\mathcal{A},\mu,h}$, is defined). For (32), assuming *him* is $him^{(1)}$, $\mu_2(1)$ is undefined; therefore (29) fails when it tries to call the value of $\mu_2(1)$, and no condition for (32) to hold in $\mathcal{A}$ is derivable.[10]

How ought we to formulate =E and quantifier introduction rules for languages with pronouns treated in the above manner? The rule of =E in a schematic formulation intended to be neutral on whether or not any links are displayed is

(33) $$\frac{\begin{array}{c} t_1 = t_2 \\ \phi(t_1) \end{array}}{\phi(t_2)}$$

(33) must be qualified to say that each occurrence of t_1 that is replaced in the minor premise **contains no targeted proper constituent**. Similarly, a quantifier introduction rule can put a variable in the position of a term only if **the term contains no targeted proper constituent**.

It follows from this that the connection between substitution-resistance and the uninterpretability of quantifying-in remains as simple here as it was in the

10. Wouldn't it be simpler to appeal to Binding Theory to exclude (32) as ill-formed? But it is only ill-formed if *him* tries to take *x* as antecedent, and there is no reason to propose that it must be an automatic consequence of the kind of quantifier introduction (32) illustrates that *him* must target whatever potential antecedent the substitution provides.

previous case: the device that prevents substitution is the very same one that blocks quantifier introduction. A link needs a singular term at the sharp end to provide another expression with a content; uninterpretability is the sure result of any inferential step that removes a constituent phrase that properly contains the targeted singular term and replaces the phrase with something unstructured, such as a variable of quantification.[11] So it may look as if a pattern is emerging relating the two failures, of substitutivity to preserve truth and of quantifier introduction to preserve interpretability. An example of substitution failure in which the substituted terms are purely referential, the device explaining substitution-failure is essentially the same as one of the two above, and yet quantifier introduction is not blocked, would disrupt the pattern. We end by considering a case which appears to satisfy this description.

5. Attitude ascriptions

The stipulations on =E that we have arrived at for our two languages – replace no targeted term, replace no term with a targeted proper constituent – block applications of =E with unwanted side-effects. Unwanted side-effects are those that bestow on $\phi(t_2)$ a different truth-condition from that of $\phi(t_1)$. Our two examples suggest that we impose the restriction that a substitution may not affect the semantic value, i.e., the contribution to truth-conditions, of any other expression in $\phi(t_1)$; in our formalism, this is captured syntactically in terms of targeting.

11. Once more, there is nothing in our diagnosis that suggests that the description in (30) is in some sense not purely referential, or in not purely referential position. Of course, there may be good reasons, when we compare descriptions with indexicals and names, to deny that the former are purely referential, or indeed referential at all. The present point is just that the description in (30)'s minor premise functions just like descriptions in contexts where there is no problem with =E: the evaluation rule (28) applies equally well to both types of occurrence. This is an endorsement of Davidsonian semantic innocence, as in the *cri de coeur* of (Davidson 1969:172).

What rule of =E is appropriate for languages in which attitude ascriptions can be expressed? If we are in agreement that

(5) Lex fears Superman; Superman = Clark; therefore Lex fears Clark

has a reading on which it is a fallacy, then we have the option of saying that the position of *Superman* in the minor premise is not purely referential. But Marcus would be among the last to assent to this proposal, and besides, we have seen no motivation for it in our previous cases. Such a proposal would, moreover, make the preservation of interpretability in

(6) Lex fears Superman; therefore, someone is such that Lex fears him

puzzling, since it is not obvious how a device of pure reference, such as a variable, could function in a position that is not purely referential.

I want to make an alternative proposal, one that saves the intuition of incorrectness in (5) and that should have some appeal to Marcus. I will then discuss a way of extending the idea from objectual ascriptions, like *Lex fears Superman*, to propositional ones, like *Lex fears that Superman is nearby*, that depends on an idea from Marcus's recent work that is quite controversial.

My claim about (5) is that it fails for essentially the same reason that (1) fails: replacing *Superman* has the side-effect of changing another expression's content. The difference between (5) and (1) is that the expression whose content changes is **non-overt** in the minor premise and conclusion of (5). The minor premise, (34.1) below, may be paraphrased as (34.2)–(34.4) for the reading on which (5) fails:

(34.1) Lex fears Superman

(34.2) Lex **so**-fears Superman

(34.3) Lex fears Superman ***qua*** Superman

(34.4) Lex fears Superman **as such**.

We shall work with (34.4). In logical form the sentential adverb phrase *as such* is prefixed as a complex sentential operator, consisting in the preposition *as* and the demonstrative *such*. The claim would then be that (34.1) optionally contains a phonetically null operator *O* with a semantics like *as such*; when the operator is present, =E fails.

It is straightforward to adapt the semantics for $\mathcal{L}_{\gamma}$ to a fragment of English capable of expressing objectual attitude ascriptions like (34.4). This fragment, $\mathcal{L}_{\text{ITV}}$, contains a stock of intensional transitive verbs (excluding theme-subject verbs like *amaze*), connectives, individual constants, etc., and in addition the *as such* operator. For (34.4) we would write

(35) *as such*(*fears*(*Lex*, *Superman*))

which expresses a proposition relative to a Kaplan-style context that fixes a reference for the *such*.

We display contexts in the usual way, so that the intended meaning for (35) is given by

(36) *as such*(*fears*(*Lex*, *Superman*))

(36) exhibits the context μ such that $\mu_2(1) = \langle 1, \mathit{Superman} \rangle$.[12] In this context, (35) expresses the same proposition as the context-independent *Lex fears Superman* ***qua*** *Superman*.[13]

The main idea behind the semantics for $\mathcal{L}_{\text{ITV}}$ is that intensional transitives like *fear* are assigned a **set** of alternative extensions instead of a single extension. *As* is

12. Since there is only one *such* and one *Superman* in (35), we suppress superscripts.

assigned a function which looks at the content of *such* in a given context μ and produces another function, one which selects an extension from the available alternatives for each intensional transitive in the scope of the *as such* in question. Which extension-selector is produced by *as* can vary with the content of *such*. So whether or not (35) comes out true in a context μ will depend on the action on *fear* of the extension-selector produced by *as*. When *Superman* is the content of *such*, we get an extension-selector that picks an extension for *fear* containing ⟨Lex, Superman/Clark⟩. When *Clark* is the content of *such*, we get an extension-selector that picks an extension for *fear* not containing ⟨Lex, Superman/Clark⟩.

Formally, we define a structure $\mathcal{A}$ for $\mathcal{L}_{\text{ITV}}$ to be a pair $(\mathcal{D},\mathcal{V})$ where (i) $\mathcal{D}$ is a non-empty domain and (ii) $\mathcal{V}$ is a valuation function for $\mathcal{L}_{\text{ITV}}$ which assigns a subset of $\mathcal{D}$, $\mathcal{D}^2$ or $\mathcal{D}^3$ to each extensional one-, two- or three-place verb. $\mathcal{V}$ assigns a subset of $\mathcal{P}(\mathcal{D}^2)$ to each intensional transitive verb.

Define a **projection function** to be a function whose domain is $\mathcal{PP}(\mathcal{D}^2) \cup \mathcal{P}(\mathcal{D}^2)$ and whose action is to return an element of its input if the input belongs to $\mathcal{PP}(\mathcal{D}^2)$ and to return the input otherwise (thus a projection function projects an extension out of a set of extensions and does nothing otherwise). Then we can define $\mathcal{V}(as)$ to be a function from the individual constants of $\mathcal{L}_{\text{ITV}}$ to projection functions.

We also need to make room for substituting-**permitting** ("transparent") uses of intensional transitives. In such a use, we want (34.1) to have a definite truth-value even though $\mathcal{V}(\textit{fear})$ is a **set** of extensions, maybe with ⟨Lex, Superman⟩ in

13. In his comments at an APA session on intensionality (December 1999) Terry Parsons informed us that medieval philosophers distinguished a category of "reduplicative" statements, which use *as* much as *qua* is used in (34.3), for instance, *A human, as a human, is rational*. Parsons suggested that this might be the best model for the interpretation of substitution-resistance in attitude ascriptions. It is indeed a good question whether or not we want the explicit context dependence that comes with hidden indexical theories and that the reduplicative account avoids.

some but not in others. To this end, define **flatten** to be a function whose domain is $\mathcal{PP}(\mathcal{D}^2) \cup \mathcal{P}(\mathcal{D}^2)$ and whose action is to return the union of its input if the input belongs to $\mathcal{PP}(\mathcal{D}^2)$ and to return the input itself if the input belongs to $\mathcal{P}(\mathcal{D}^2)$. **Flatten** converts a set of extensions into a single extension, and does nothing if presented with an extension (this allows *as such* to have extensional verbs within its scope "inertly").

Say that f is an **extension-determiner** if f is either a projection function or **flatten**. We relativize the definition of truth to extension-determiners in a way that ensures that, if everything is extensional, only **flatten** is used, but when an *as such* is encountered, a projection function is called. Read

$$(\mathcal{A}, \mu, f) \Vdash_h \phi$$

as "ϕ is true in the structure $\mathcal{A}$ and context μ relative to the extension-determiner f and assignment h". The inductive definition of this concept is

(37.1) if $\phi(t_1,\ldots,t_n)$ is an atomic wff of $\mathcal{L}_{\text{ITV}}$, then $(\mathcal{A}, \mu, f) \Vdash_h \phi(t_1,\ldots,t_n)$ iff $\langle \mathcal{V}(t_1),\ldots,\mathcal{V}(t_n)\rangle \in f(\mathcal{V}(\phi))$;

(37.2) $(\mathcal{A}, \mu, f) \Vdash_h \neg\phi$ iff $(\mathcal{A}, \mu, f) \nVdash_h \phi$;

(37.3) $(\mathcal{A}, \mu, f) \Vdash_h \phi \wedge \psi$ iff $(\mathcal{A}, \mu, f) \Vdash_h \phi$ and $(\mathcal{A}, \mu) \Vdash_h \psi$;

(37.4) $[\![such^{(i)}]\!]_{\mathcal{A},\mu} = \mu_2(i)$;

(37.5) $[\![as\ such^{(i)}]\!]_{\mathcal{A},\mu} = \mathcal{V}(as)([\![such^{(i)}]\!]_{\mathcal{A},\mu})$;

(37.6) $(\mathcal{A}, \mu, f) \Vdash_h as\ such^{(i)}(\phi)$ iff $(\mathcal{A}, \mu, [\![as\ such^{(i)}]\!]_{\mathcal{A},\mu}) \Vdash_h \phi$;

(37.7) $(\mathcal{A}, \mu, f) \Vdash_h$ (*some/every v: ψv*)[*ϕv*] iff for some/every x in $\mathcal{D}$ such that $(\mathcal{A}, \mu, f) \Vdash_{h^\wedge(x\mapsto v)} \psi v$, we also have $(\mathcal{A}, \mu, f) \Vdash_{h^\wedge(x\mapsto v)} \phi v$.

Finally, we introduce a concept of truth that is not relative to an extension-determiner or assignment by the definitions

(38.1) $(\mathcal{A}, \mu, f) \Vdash \sigma$ iff $(\mathcal{A}, \mu, f) \Vdash_{\langle\rangle} \sigma$;

(38.2) $(\mathcal{A}, \mu) \Vdash \sigma$ iff $(\mathcal{A}, \mu, \textbf{flatten}) \Vdash \sigma$.

So in evaluating (35) in $(\mathcal{A}, \mu)$, or (36) in $\mathcal{A}$, begin by evaluating it with respect to **flatten**, as in (38.2). By (37.4), $[\![such^{(1)}]\!]$ is *Superman*; therefore $[\![as\ such^{(1)}]\!]_{\mathcal{A},\mu}$ will be whatever projection function f^* is determined by applying $\mathcal{V}(as)$ to *Superman* (the name), as in (37.5). Next use (37.6) to switch from evaluating with respect to **flatten** to evaluating with respect to f^*. Finally, use (37.1) to obtain a truth-value for *fears*(*Lex*, *Superman*) by projecting the extension f^*(*fear*) out of $\mathcal{V}$(*fear*) and testing ⟨Lex, Superman⟩ for membership in that extension.[14]

Continuing the parallel with $\mathcal{L}_\gamma$, we can define notions of validity-in-a-context and universal validity (recall (16) and (17), p. 10 *ff*). A context μ is defined for a discourse Δ iff $\mu(i)$ is defined for each i indexing a *such* in Δ, and for each such i, if $\mu(i) = \langle k,c \rangle$ then (exactly) one occurrence of c is indexed k in Δ. This leads to the following analog of (19):

(39) $\Sigma \vDash_\mu \sigma$ iff μ is defined for $\Delta = \Sigma \cup \{\sigma\}$ and for every $\mathcal{A}$, if $(\mathcal{A}, \mu) \Vdash \tau$ for every $\tau \in \Sigma$, then $(\mathcal{A}, \mu) \Vdash \sigma$.

It is immediate that we have

(40) *as such*(*Lex fears Superman*), *Superman is Clark* $\nvDash_\mu$ *as such*(*Lex fears Clark*)

if $\mu(1) = \langle 1, \textit{Superman} \rangle$ and $\mu(2) = \langle 2, \textit{Clark} \rangle$. For we can allow $\mathcal{V}_\mathcal{A}$ to assign *fear* a set of two extensions, only one of which contains ⟨Lex, Superman/Clark⟩, and arrange for $\mathcal{V}_\mathcal{A}(as)$ to choose that one when the content of *such* is *Superman* and the

14. The semantics given here simplifies that in §4 of (Forbes 2000a) in two respects: (i) there is no quantification over mental states, and (ii) there is no quantification over individual concepts and no *so-labelled*. Simplification (i) has to go in a fuller treatment, since it causes extensionality problems. But simplification (ii) I regard as an improvement, since it considerably reduces the hidden plumbing that my earlier theory posited.

other when its content is *Clark*. If we think of the minor premise and conclusion in (5) as optionally containing an unvoiced operator *O* with a semantics like the one just sketched for *as such*, then if *O* is present by default (or in virtue of the implicit priming example (5) generates in the typical reader), we have an explanation of why (5) strikes us as incorrect. It is, moreover, an explanation that should appeal to Marcus, since it is entirely consistent with a "direct reference" view of names; once more, nothing in our approach substantiates the idea that *Superman* or its position in the minor premise of (5) is anything other than purely referential.

On the proof-theoretic side, the language $\mathcal{L}_{\text{ITV}}$ requires the $\mathcal{L}_{\mathcal{Y}}$-restriction on its rule of =E, that targeted constants cannot be substituted. There is also an Elimination rule for *as such*, namely, that from $as\ such^{(i)}(\phi)$ we can infer ϕ. The semantic justification for this is that if $as\ such^{(i)}(\phi)$ holds in $(\mathcal{A}, \mu)$, then this means that each relevant pair of objects belongs to at least one of the extensions in the $\mathcal{V}$-assignment for each intensional verb in ϕ. So if ϕ is evaluated independently in $(\mathcal{A}, \mu)$ it will come out true, for lacking an *as such* it will be evaluated with respect to $(\mathcal{A}, \mu,$ **flatten**$)$. By definition of **flatten**, this leads us to consider the union of the extensions of each intensional verb there may be in ϕ, and each union contains the relevant pair of objects *ex hypothesi*.[15]

The soundness of *as such* Elimination is important for solving the puzzle of why quantifier introduction, as illustrated in

(6) Lex fears Superman; therefore, someone is such that Lex fears him

produces an interpretable conclusion. We have written the semantics in a way that assumes that only individual constants can be targeted; specifically, $\mathcal{V}(as)$ is

15. This argument assumes ϕ is *as such*-free; but it is easy to see from (37.6) that any *as such* with a further *as such* within its scope is redundant.

defined only for individual constants. So if a context μ assigns a variable as the content of some *such*, derivation of a condition for the sentence containing that *such* to be true will crash. But it follows that the conclusion of (6) is uninterpretable only if we suppose it to have been inferred by one step of ∃I (or by steps of ∃I coupled with redundant cycles of I and E rules that produce maximum formulae). The reason (6) strikes us as acceptable is that when we interpret *Lex fears Superman* as (35), *Lex fears Superman as such*, we take (6) to be an enthymeme in which a step of *as such* Elimination is implicitly performed before ∃I is applied.

This means that the case of attitude ascriptions is not essentially different from the others we have investigated: the mechanism explaining substitution-resistance also blocks quantifier introduction, **if** it is present. The superficial difference is that in ordinary English, the mechanism's presence or absence is not indicated on the surface, whereas the difference between, say, *Giorgione is so-called* and *Giorgione is called 'Giorgione*', is easy to see.

6. Propositional ascriptions

Our focus has been on objectual attitude ascriptions because they are the easiest case. Ideally, we would like an absolutely uniform treatment of objectual and propositional ascriptions; but it is unclear how this is to be accomplished. One reasonable philosophical account of the phenomenon that the semantics for $\mathcal{L}_{\text{ITV}}$ reflects is that we can stand in attitude relations to objects **under modes of presentation** of those objects. Furthermore, we may stand in such a relation to a certain object under one mode but not under another. The differing extensions for an intensional transitive like *fear* show which experiencer-theme pairs enter or drop out of the verb's extension as modes of presentation vary. We may say that objectual attitude relations are relations **to** Russellian entities that hold **under** Fregean ones. Names are modally but not cognitively equivalent because only cognitive operators like *as such* invoke the Fregean entities.

A uniform account of propositional and objectual ascriptions would require that propositional attitude relations are also relations to Russellian entities that hold under Fregean ones. But here we run into difficulties, since it seems that the only candidate for the second term of a propositional relation is a proposition. Fregeans and Russellians characteristically disagree over the nature of propositions (there is no comparable disagreement either about objects or about modes of presentation). So it is unclear how a theory of propositional ascriptions could combine Frege and Russell.

In more recent work, Marcus (1990b:240–1) has made the interesting proposal that Russellians abandon the notion of proposition, which she argues is ineliminably "language-centered", and exchange it for the "object-centered" notion of a state of affairs (SOA). So-called "propositional" attitude relations are really relations to states of affairs, which are complexes of objects and properties. The idea that **modal** properties are fundamentally features of SOA's is one to which I am highly sympathetic (see Forbes 1989:Ch. 5), but Marcus's suggestion is more radical. And she argues that substituting SOA's for propositions helps Russellians with issues about the rationality of those who apparently take "assentive" attitudes towards contradictory propositions (1990b:248–51), as in the example

(41) Lex fears that Superman is nearby and Clark is not.[16]

The advantages for Russellians of replacing fearing a contradictory proposition with fearing an impossible SOA are not so clear to me.[17] But once we have distinguished SOA's and propositions in Marcus's way, we can think of propositions as

16. Suppose Lex hopes to use Clark as a hostage to thwart Superman.

modes of presentation of SOA's. We are then in a position to distinguish believing an impossible SOA from believing a **manifestly** impossible SOA, a distinction reflected in the following two formulations:

(42.1) *as such*(*fears*(*Lex*, *that Superman is nearby and Clark is not*))

(42.2) *as such*(*fears*(*Lex*, *that Superman is nearby and Superman is not*)).

Here the *that*-clauses refer to SOA's, and the indicated subsentences invoke modes of presentation of them. It is not difficult to extend the semantics for $\mathcal{L}_{ITV}$ to accommodate the introduction of SOA's: $\mathcal{D}$ becomes two-sorted, an extension for *fear* may include experiencer-SOA pairs as well as experiencer-object pairs, *fear* is assigned multiple extensions, and *as such* chooses between them as it does in $\mathcal{L}_{ITV}$.[18] Like before, substitution is forbidden for targeted expression; we simply expand the range of expressions that can be targeted. More carefully, substitution of expressions that are either targeted themselves, or that contain a targeted proper constituent, is forbidden. So there is no interchanging *that Superman is nearby and Clark is not* and *that Superman is nearby and Superman is not* even though they both refer to the same SOA. Similarly, there is no quantifying into any position within the extent of the delimiter of a targeted expression.

The treatments of objectual and propositional ascriptions are now exactly parallel. But have we paid a high price for this uniformity? So far as I can see, the

17. Marcus says (*loc. cit.*) that the language-centred theorist is "baffled" by the question, does Lex fear Clark is nearby or does he not? She then argues that if we take the object-centred view that a belief is a dispositional relation to a SOA, namely, to act as if that SOA obtained, "a puzzle has been solved". But I find myself baffled by the question what it would be to act as if the SOA that Clark both is and is not nearby obtains.

18. Once SOA-terms are in the language it is more convenient to use a type-theoretic semantics than to build the referents of SOA-terms compositionally. See (Forbes 2000b:§4).

most troublesome consequences that the unified treatment highlights are in the area of action-explanation. For instance, we will hold that

(43) Anyone who fears Superman and believes Superman is nearby will, *ceteris paribus*, take steps to avoid encountering Superman.[19]

This is a correct principle if *fears* and *believes* are in the scope of (the counterpart *O* of) *as such*. But it seems uncontroversial that (43) is simply an instance of the more general

(44) Anyone who fears a certain person and believes that that person is nearby will, *ceteris paribus*, take steps to avoid encountering that person.

However, if (44) is not to feature the kind of quantifier introduction that we have banned, there must be no *as such* in it. But in that case, a step of ∀E followed by =E will produce the apparently false

(45) Anyone who fears Clark and believes Superman is nearby will, *ceteris paribus*, take steps to avoid encountering Superman.

How we should deal with this difficulty is a large issue that I am not going to try to tackle now.[20] But it is some small consolation that, as comparison of (43) and (45) reveals, this problem was already present in our semantics for $\mathcal{L}_{\text{ITV}}$. The viability of a uniform treatment of objectual and propositional attitudes is not threatened by (45); it is the treatment of objectual attitudes that is threatened.

19. Strictly, we ought to say (here and in (44)) *take steps that he or she thinks will significantly reduce the chances of encountering that person if he or she believes the chances of an encounter are too high without such steps being taken*.

20. One way to proceed is to make liberal use of the *ceteris paribus* qualifier; see (Braun 1999) for this approach. A second alternative, bringing us back to a theme of Marcus's earlier work, is to find a non-standard construal of quantification that permits quantifying into the scope of *as such* (but without resuscitating (2)). A third possibility is to query the covering-law model of action-explanation implicit in our discussion; perhaps simulationism offers a way out.

Graeme Forbes
Department of Philosophy
Tulane University, New Orleans, LA 70118-5698
forbes@tulane.edu

BIBLIOGRAPHY

Braun, David. 1999. Russellianism and Psychological Generalizations. Noûs 34 (2):203–236.

Cartwright, Richard. 1971. Identity and Substitutivity. In *Identity and Individuation*, edited by M. K. Munitz. New York: NYU Press.

Davidson, Donald. 1969. On Saying That. In *Words and Objections: Essays on the Work of W. V. Quine*, edited by D. Davidson and J. Hintikka. Dordrecht: Reidel.

Evans, Gareth. 1977. Pronouns, Quantifiers and Relative Clauses (I). *The Canadian Journal of Philosophy* 7 (3):467–536.

Fine, Kit. 1989. The Problem of *De Re* Modality. In *Themes from Kaplan*, edited by J. Almog, J. Perry and H. Wettstein. Oxford and New York: Oxford University Press.

Forbes, Graeme. 1989. *Languages of Possibility*. Edited by M. Davies. Vol. 9, *Aristotelian Society Monographs*. Oxford: Basil Blackwell.

———. 1993. But *a* Was Arbitrary. *Philosophical Topics* 21:21–34.

———. 1996. Substitutivity and the Coherence of Quantifying In. *The Philosophical Review* 105:337–72.

———. 2000a. Objectual Attitudes. *Linguistics and Philosophy* 23 (2):141–183.

———. 2000b. Sets, Types, and Intensional Transitives. *Unpublished ms.*

Kaplan, David. 1989a. Afterthoughts. In *Themes from Kaplan*, edited by J. Almog, J. Perry and H. Wettstein. Oxford and New York: Oxford University Press.

———. 1989b. Demonstratives. In *Themes from Kaplan*, edited by J. Almog, J. Perry and H. Wettstein. Oxford and New York: Oxford University Press.

Larson, Richard, and Gabriel Segal. 1995. *Knowledge of Meaning*. Cambridge, Mass.: The MIT Press.

Linsky, Leonard. 1967. *Referring*. New York: Humanities Press.

Marcus, Ruth Barcan. 1948. Smullyan on Modality and Description. *The Journal of Symbolic Logic* 13:149–150.

———. 1961. Modalities and Intensional Languages. *Synthese* 13:303–322.

———. 1967. Essentialism in Modal Logic. *Noûs* 1 (1):90–96.

———. 1975. Does the Principle of Substitutivity Rest on a Mistake? In *The Logical Enterprise*, edited by A. Anderson, R. B. Marcus and R. Martin. New Haven: Yale University Press.

———. 1990a. A Backward Look at Quine's Animadversions on Modalities. In *Perspectives on Quine*, edited by R.Barrett and R. Gibson. Oxford: Basil Blackwell.

———. 1990b. Some Revisionary Proposals about Belief and Believing. *Philosophy and Phenomenological Research* 50:133–153.

———. 1993. *Modalities*. Oxford and New York: Oxford University Press.

Parsons, Terence. 1969. Essentialism and Quantified Modal Logic. *The Philosophical Review* 78:35–52.

Quine, W. V. 1955. Quantifiers and Propositional Attitudes. In *Reference and Modality*, edited by L. Linsky. Oxford and New York: Oxford University Press.

———. 1961. Reference and Modality. In *Reference and Modality*, edited by L. Linsky. Oxford and New York: Oxford University Press.

———. 1966. Three Grades of Modal Involvement. In *The Ways of Paradox*. New York: Random House.

Smullyan, Arthur. 1948. Modality and Description. *The Journal of Symbolic Logic* 13:31–37.

Soames, Scott. 1989. Presupposition. In *Handbook of Philosophical Logic Volume IV: Topics in the Philosophy of Language*, edited by D. Gabbay and F. Guenthner. Dordrecht and Boston: Kluwer Academic Publishers.