Reflections on Reflexivity

Author(s): Nathan Salmon


Published by: Springer

Stable URL: https://www.jstor.org/stable/25001461

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at https://about.jstor.org/terms

Springer is collaborating with JSTOR to digitize, preserve and extend access to Linguistics and Philosophy
Although two or more are often lumped together as if they were the same, or virtually the same, at least five different theories should be sharply distinguished concerning the contributions to propositional content made by the pronouns occurring in sentences like the following:

(1) John loves himself
(2) John loves his wife.

Linguists will note that in both sentences the pronoun — either 'himself' or 'his' — is c-commanded by 'John'.

In Salmon (1986b) I cited M. J. Cresswell as one theorist (among many) who claims that (1) expresses the same proposition as 'John loves John'. On the Simple Anaphor Theory the pronoun occurrence in (1) or (2) is simply another singular term, one that takes on the same semantic content as its antecedent, referring anaphorically to John. The Simple Anaphor Theory treats (1) as expressing the proposition that John loves John, and (2) as expressing the proposition that John loves John's wife. We may represent these propositions as:

\[
\langle C('John'), C('John'), C('loves') \rangle
\]
\[
\langle C('John'), C('the wife of'), C('John'), C('loves') \rangle,
\]

where \( C \) is the semantic content function for English. Fancier representations are possible, but this will suffice for the present purpose. By adopting this form of representation I follow the Frege-Russell tradition in assuming that the semantic content of a sentence is not, for example, the set of possible worlds with respect to which the sentence is true, but rather a structured, composite entity whose constituents are (roughly) the semantic

* The present essay was written largely in response to a draft of a version of McKay (1991), and to Soames (1989/90). I appreciate the comments provided by Thomas McKay and by a second, anonymous referee.

1 The notion of c-command from theoretical linguistics corresponds, roughly, to the logician's notion of scope. An expression occurrence in a sentence c-commands a nonoverlapping expression occurrence in that sentence iff the first branching node that dominates the first constituent also dominates the second.

2 Alternatively, the pronoun may be regarded as inheriting a modified (e.g., intensionally rigidified) content from its antecedent, which serves to "fix the reference" of the pronoun, in something like Kripke's sense.

contents of the sentential components. As a facilitating expedient we may further assume that ‘John’ is a Millian term that directly refers to John. We may then represent the two propositions as:

\[(\text{John, John, the loving relation})\]
\[(\text{John, the concept wife-of, John, loving}).^3\]

Nothing that I shall argue here depends on the Millian assumption that the name ‘John’ contributes its referent to the propositions contained by sentences in which the name occurs; my central points are compatible with the Fregean thesis that ‘John’ instead contributes a Sinn that is thoroughly descriptional, or purely conceptual, in nature.

Contrary to the interpretation of several readers, Salmon (1986b) does not reject the Simple Anaphor Theory. The misunderstanding may have arisen because I gave reasons there for rejecting this analysis and presented a rival analysis. I cannot overemphasize that I do not know of any decisive refutation of the Simple Anaphor Theory. My own view is that sentences like (1) and (2) may be ambiguous, that the Simple Anaphor Theory may well capture one anaphoric reading (even if not the only, or even the most natural, reading), and that it is even possible, contrary to popular belief, that the Simple Anaphor Theory correctly gives the only legitimate reading of these sentences (aside from the indexical or deictic reading of (2)).

Whereas it remains a genuine possibility that the Simple Anaphor Theory correctly captures one reading for sentences like (1) and (2), I am inclined to believe that it does not give the whole story. My general dissatisfaction with the Simple Anaphor Theory stems from the fact that it leaves out the element I call reflexivity that seems present in (1) and (2), at least on one reading. The other four theories that I shall distinguish attempt to accommodate the reflexivity evidently intrinsic to these sentences.

On the Linked Anaphor Theory, as on the Simple Anaphor Theory, the pronouns occurring in (the alleged reflexive readings of) (1) and (2) are anaphoric singular terms that derive their content and reference from their antecedents, but their anaphoric character is also alleged to be something that itself shows up in the propositions expressed by (the relevant readings of) (1) and (2). The propositions are held to contain some further element indicating the “linkage” – or identification – between John’s occurrences therein (or, if one prefers, between the occurrences of the Fregean sense of ‘John’ therein). This further propositional element

---

3 This method of representing propositions is developed further in Salmon (1986a, Appendix C, pp. 143–151).
might be represented through something like lines-of-connection, as follows:

\[ \text{\langle John, John, the loving relation\rangle} \]

\[ \text{\langle John, the concept wife-of, John, loving\rangle}. \]

The second proposition, for example, might be thought of as having something like the following import, where ‘α’ and ‘β’ are two distinct names having the same semantic content as ‘John’: α loves β’s wife, and furthermore, that wife-lover is the same as that one whose wife is loved. Something like this theory was proffered in the mid 1950s by Hilary Putnam, and more recently by David Kaplan and William Taschek— for sentences like ‘John loves John’ and ‘John loves John’s wife’ in which a singular term recurs (perhaps in addition to sentences like (1) and (2)).

I know of no decisive evidence against the Linked Anaphor Theory, though there is one sort of consideration that inclines me against it. The relational proposition that is represented as \( \text{\langle John, Mary, the loving relation\rangle} \) may be said to attribute the property of loving Mary to John (or to ascribe the property to John, or to predicate the property of John, etc.). It might also be said to attribute the property of being loved by John to Mary. It does not do either of these things, however, in the same direct way that the proposition (John, the property of loving Mary) does the first, since the attributed property and the individual to whom the property is attributed occur as the sole elements of the latter proposition. Let us say that the second proposition directly attributes the property of loving Mary to John. Whereas the proposition \( \text{\langle John, Mary, loving\rangle} \) may be regarded as directly attributing the binary loving relation to the ordered pair \( \text{\langle John, Mary\rangle} \), and as thereby indirectly attributing the property (singularly attribute) of loving Mary to John, it does not directly attribute

---

4 Cf. Salmon (1986a, pp. 156, 164 and passim).
5 Putnam (1954); Kaplan (1990, p. 95n6); Taschek (1991). Kaplan has suggested calling these lines-of-connection Putnam wires. Putnam’s idea that the very logical structure of a sentence makes a contribution to the sentence’s semantic content in the manner proposed is criticized in Soames (1987) and in Salmon (1986a, pp. 164–165n4).
6 The theory proffered in McKay (1991) seems similar in many respects to the Linked Anaphor Theory. McKay evidently regards the syntactic relation of c-command between ‘John’ and the pronoun (rather than any piece of notation like a word or phrase) as providing the linkage component of the proposition.
7 The latter proposition may be expressed in English by ‘John is someone who loves Mary’. The former might be expressed by ‘John and Mary are such that the first loves the second’. This leaves open the question: Which proposition is expressed in English by ‘John loves Mary’? (There is a brief discussion in Salmon (1981, p. 20); see especially footnote 19.)

This content downloaded from 132.174.249.166 on Thu, 25 Jan 2024 19:05:17 +00:00
All use subject to https://about.jstor.org/terms
any property to any individual. Then the Linked Anaphor theory apparently does not capture the intuition that (1) directly attributes to John the same property that ‘James loves himself’ directly attributes to James, to wit, the Narcisistic property of loving oneself. Similarly, (2) seems directly to attribute to John the property of loving one’s own wife. On the Linked Anaphor Theory, these reflexive properties make no appearance in the semantically contained propositions; they evidently must be inferred, as logical consequences, from the information actually present in those propositions.7

A closely related problem, or potential problem, with the Linked Anaphor Theory is that, on the view that the Simple Anaphor Theory is incorrect because (1) and (2) have reflexive readings, the predicates ‘loves himself’ and ‘loves his wife’ (on the alleged reflexive readings) would seem to be closed predicates, complete and fully determinate in themselves as regards both content and extension, without an attached grammatical subject to serve as antecedent. Both the Simple Anaphor Theory and the Linked Anaphor Theory fail to achieve this result. On those theories, the pronouns in (1) and (2) derive their content and reference from their antecedents (the relevant occurrence of ‘John’).8

---

7 This sort of consideration focuses on the subtle, but very real, differences that separate the complex proposition that $\alpha$ loves the wife of $\beta$, and furthermore that wife-lover is the same as that one whose wife is loved from the seemingly simpler proposition that $\alpha$ is an own-wife-lover. The first, for example, is essentially relational while the second is essentially single-subject/monadic-predicate. It is even possible for someone to entertain and believe the first without believing or entertaining the second.

8 Since the anonymous referee was confused concerning my claim that the predicates ‘loves himself’ and ‘loves his wife’ are incomplete on the Anaphor theories, the editor has suggested that I include a word of caution to the reader. The referee evidently misunderstood me to be claiming that on the Anaphor theories, the predicates ‘loves himself’ and ‘loves his wife’ are open expressions (like ‘loves x’ and ‘loves the wife of x’ with the variable ‘x’ occurring free) and their occurrences (or their pronoun occurrences) in (1) and (2) remain free (unbound), so that (1) and (2) are deemed open formulas (“open sentences”) on the Anaphor theories. I am not making any of these claims. The Anaphor theories need not regard anaphors on the model of a bound, or bindable, occurrence of an open expression (like a variable), and indeed it is probably a mistake to do so. Even if anaphors are so regarded, they should probably also be seen as bound in some way by their antecedents. (In that case, singular-term antecedents would emerge as a sort of quantifier, in the manner of Richard Montague.)

Accepting that the Anaphor theories treat the predicates of (1) and (2) as open expressions, the referee defended the Anaphor theories by arguing that, on any theory that sees the phenomenon of anaphora as incorporating a syntactic linkage between an anaphor and its antecedent, it is indeterminate whether the predicates are free or bound in (1) and (2). This is also incorrect. Indeed, if it were correct, it would point to a serious defect, rather than a virtue, in the Anaphor theories, since it is determinate that (1) and (the relevant reading of) (2) are (closed) sentences.

The incompleteness of an anaphoric expression removed from its anaphoric setting is not
On the Polyadic-Predicate Operator Theory, the pronouns occurring in (1) and (2) are not singular terms at all – anaphoric or otherwise. They designate a higher-order entity. In the simplest kind of case, they designate the function that maps any binary relation $R$ between individuals to (the characteristic function of) the class of individuals that reflexively bear $R$ to themselves, $(\lambda R)(\lambda x)[xRx]$. On this theory, the 'himself' in (1) and the 'him' implicit in (2), on the alleged reflexive readings of these sentences, are expressions for this higher-order function, and they designate it non-anaphorically.

A special case of the Polyadic-Predicate Operator Theory, the Dyadic-Predicate Operator Theory for certain reflexive pronoun occurrences, is the rival theory (rival to the Simple Anaphor Theory) presented in Salmon (1986b). Some readers have erroneously thought that I endorse the theory there. Salmon (1986b) takes no sides on the question of whether the Polyadic-Predicate Operator Theory is correct, for reflexive pronouns or for other pronouns. In fact, however, the Polyadic-Predicate Operator Theory has difficulties which make it almost certainly false.

First, it makes pronouns generally – including reflexive pronouns – radically ambiguous, between pronominal singular terms on the one hand (at least for their indexical use and for occurrences not c-commanded by antecedents), and polyadic-predicate-operator expressions on the other. In fact, the Polyadic-Predicate Operator Theory regards the several (explicit or implicit) occurrences of 'he' and 'him' in a complex sentence like the following as somehow forming a single, albeit scattered, polyadic-predicate operator:

$$S: \text{John, with his wife's help, fooled his sister into thinking that he was ill.}$$

In this case, the scattered predicate operator would operate on the extension of a complex four-place predicate – even though the needed predicate does not seem to occur as a separate, unified component of the original surface sentence $S$. All of this is implausible purely as a matter of English syntax.\footnote{The needed predicate might be formulated in a quasi-formal version of English as ‘$(\lambda x y z u)[x, \text{with } y's \text{ wife's help, fooled } z's \text{ sister into thinking that } u \text{ was ill}]$. It is not easy to see how the same phenomenon as the openness of an expression containing a free variable.

In many respects it is more like the incompleteness of a demonstrative expression (‘loves him’) considered in abstraction from any accompanying demonstration that might fix its content and extension in a particular context of utterance. The criticism of the Anaphor theories does not challenge their ability to accommodate the completeness of the sentences (1) and (2). Instead I challenge their ability to accommodate the completeness of certain predicates, considered in isolation from any larger expression in which those predicates occur.}
Moreover, the Polyadic-Predicate Operator Theory fails to achieve the
desired results as regards content. The motivation for the Polyadic-Predi-
cate Operator Theory is the intuition – more or less shared by Peter
Geach, David Wiggins, Tanya Reinhart, and many others – that (1)
and (2) express the same propositions as those expressed by:

\[(\lambda x)[x \text{ loves } x](\text{John})\]
\[(\lambda x)[x \text{ loves } x's \text{ wife}](\text{John}).\]

These propositions are represented here as:

P1: \(\langle \text{John, the reflexive property of loving oneself} \rangle\)
P2: \(\langle \text{John, the reflexive property of loving one's own wife} \rangle\).

Instead of these desired propositions, the Polyadic-Predicate Operator
Theory delivers the (respectively logically equivalent) propositions ex-
pressed by:

\[(\lambda R)(\lambda x)[x R x](\text{loves})(\text{John})\]
\[(\lambda R)(\lambda x)[x R x][(\lambda y)(y \text{ loves } z's \text{ wife})](\text{John}).\]

These propositions (which might be expressed in English by 'John has the
reflexivization of loving' and 'John has the reflexivization of loving the
wife of') would be represented here as:

\(\langle \text{John, } O, \text{ the binary loving relation} \rangle\)
\(\langle \text{John, } O, \text{ the binary relation of loving the wife of} \rangle,\)

where \(O\) is the content of the predicate-operator expression
\((\lambda R)(\lambda x)[x R x]\) (perhaps something like the operation of assigning to any
binary relation \(R\) between individuals the characteristic function of the
class of individuals that reflexively bear \(R\) to themselves). Here again, the
propositions delivered by the theory do not directly attribute reflexive
properties; the desired properties make no appearance in the relevant
propositions, and must be inferred on the basis of the information actually
present in those propositions.

It would appear that the Polyadic-Predicate Operator Theory, extended

(and perhaps not even possible) to formulate the needed predicate in standard English. Perhaps the following will do: 'are four individuals such that the first, with the help of the second's wife, fooled the third's sister into thinking that the fourth was ill'. (See the discussion of the Abstraction Operator Theory below.)

10 Geach (1962, chap. 5, pp. 108–143, especially p. 132 and passim); Geach (1965, pp. 112–113).
to cover pronouns c-commanded by singular-term antecedents generally, is advocated by Scott Soames. He proposes extending the Dyadic-Predicate Theory presented in Salmon (1986b) into a theory according to which anaphoric pronouns with c-commanding singular term antecedents are not themselves singular terms, but rather are abstraction operators which combine with predicates of the sort illustrated by ["\(~\_\) loves \(\_\)'s mother"] to produce predicates . . . represented by [(\(\lambda x\))[x loves x’s mother]]. In the simplest cases the effect of the anaphoric pronoun is to map a two-place relation \(R\) onto the corresponding one-place property of being an object \(o\) such that \(R\) applies reflexively – i.e. of being an object \(o\) such that \(R\) applies to the pair \(<o, o>\).13

Evidently the term ‘antecedent’ must be given a non-standard sense here, since the pronouns are alleged on this theory not to be anaphoric terms. Soames’s characterization of the the proposed theory as an extension of the Dyadic-Predicate Operator Theory of Salmon (1986b) and his characterization of the pronoun in (2) as having the effect of mapping a binary relation onto the corresponding reflexive property, strongly support an interpretation on which he is defending a version of the Polyadic-Predicate Operator Theory. On the other hand, Soames’s use of the phrase ‘abstraction operator’ instead of ‘predicate operator’, and his subsequent discussion, suggest that he may have in mind a variant of the Polyadic-Predicate Operator Theory. According to the fourth theory considered here, the pronoun ‘him’ in (2) is a genuine predicate-abstraction operator, which forms a monadic predicate for loving one’s own wife when attached to the gappy expression ‘\(~\_\) loves \(\_\)'s wife’). Although Soames calls this gappy expression a ‘predicate’, it would in fact play the role of an open formula, like ‘\(x\) loves \(x\)'s wife’, with gaps serving as separate occurrences of a single free variable.

This Abstraction Operator Theory duplicates the syntactic implausibility of the Polyadic-Predicate Operator Theory by treating gappy expression’s like ‘\(~\_\) with \(\_\)'s wife’s help, fooled \(\_\)'s sister into thinking that \(\_\) was ill’ as unified, semantically significant constituents of sentences like \(S\) above.14 The Abstraction Operator Theory compounds the syntactic im-

13 Soames (1989/90, p. 204). Soames seems to suggest (p. 204n11) that Salmon (1986b) explicitly rejected the extension of the theory from the special case of reflexive pronouns to pronouns c-commanded by singular-term antecedents generally. In the passage that Soames cites, however, I made only the weaker claim that the Polyadic-Predicate Operator Theory is incorrect for some nonreflexive pronoun occurrences (besides pronouns of laziness). I had in mind pronouns occurring indexically and so-called E-type or donkey pronoun occurrences, which are not c-commanded by their antecedents. (I do not presuppose that the latter sort of pronoun occurrences are not simply special cases of indexical pronoun occurrences.)

14 Soames says that he is “assuming that gap-containing formulas of arbitrary complexity may be counted as predicates that receive semantic interpretations which are operated on by abstraction operators” (p. 204n11).
plausibility by treating this gappy expression not as a closed predicate but as an open formula with its gaps serving as bindable free-variable occurrences. The Abstraction Operator Theory also apparently shares with the Linked Anaphor Theory the undesirable feature that English predicates like ‘loves himself’ and ‘loves his wife’, on their alleged reflexive readings, are not complete and determinate in themselves as regards content and extension without an attached antecedent. Unlike the situation on the Linked Anaphor Theory, in which the incompleteness of these predicates arises from lack of a content and referent provided by an antecedent, here the incompleteness arises from lack of the antecedent’s syntactic position – an additional gap, which needs to be bound by the alleged pronominal abstraction operator. On the Abstraction Operator Theory, the ‘himself’ in (1) functions like the abstraction phrase ‘(λx)’ in ‘(λx)[x loves x]’, forming a monadic predicate from an open formula. It cannot abstract the monadic predicate from the dyadic predicate ‘loves’, nor from the “open” expression ‘loves _’; it requires an open formula ‘_ loves _’ with two gaps (the analogue of ‘x loves x’).

The final theory discussed here, the Bound Variable Theory, succeeds where the previous theories fail. On the Bound Variable Theory, (1) has precisely the same content as (1’), (2) precisely the same as (2’), and the pronouns in (1) and (2) function as variables bound by a ‘λ’-abstraction operator – like the final occurrences of ‘x’ in (1’) and (2’). The Bound Variable Theory simultaneously achieves the following results: (i) a complex sentence like S above is not regarded as somehow containing a scattered polyadic-predicate operator or predicate-abstraction operator and a complex, polyadic predicate or gappy open formula to serve as the operator’s operand; (ii) predicates like ‘loves himself’ and ‘loves his wife’ are closed expressions, determinate in content and extension without an attached antecedent; (iii) the pronouns in (1) and (2) are singular terms; (iv) the pronouns in (1) and (2) may be regarded as anaphors; and (v) (1) expresses P1 and (2) expresses P2, thereby directly attributing reflexive properties. Although the pronouns in (1) and (2) may be seen as anaphors on the Bound Variable Theory, the theory has an additional feature stressed by Geach: (vi) it is a mistake to ask for the referent or designation of the pronoun occurrences in (1) and (2) – just as it is a mistake to ask

---

15 There are indeed open predicates that correspond to the gappy expression ‘loves _’, e.g. ‘(λy)[y loves x]’ or ‘is someone who loves x’. But abstraction on the free variable does not yield a monadic predicate for loving oneself (a predicate corresponding to ‘loves himself’). Instead it yields an operator that attaches to a singular term to form a monadic predicate. (Attached to ‘John’ it yields a predicate for the property of loving John.)
for the referent of ‘x’ in (1’) or (2’) (even under an assignment of values to variables).16

Something similar to the Bound Variable Theory has been advocated by Geach, Reinhart, and others. If (1) and (2) indeed have reflexive readings that the Simple Anaphor Theory fails to capture (as I am inclined to believe), then the Bound Variable Theory would appear to be the most likely of the theories discussed here to yield the correct analysis of those readings. The only problem with the theory that I can see (aside from the fact that it posits a potentially controversial reading – the alleged reflexive reading – for (1) and (2)) derives from the fact that it carries the burden of positing an invisible abstraction operator in the predicates of (1) and (2), on their alleged reflexive (closed) readings.17 One might explain the invisible abstraction operator in the predicates of (1) and (2) by positing a reflexive-nonreflexive ambiguity in (1) and (2), incorporating the Simple Anaphor Theory for the nonreflexive reading, and declaring that the reflexive reading is shorthand for something involving an abstractor phrase like ‘is someone who’ or ‘is something that’. If (1) and (2) have reflexive readings (and I am inclined to think they do), it is not immediately objectionable to regard the sentences, on those readings, as shortened versions of

\[
\begin{align*}
\text{John is someone who loves himself} \\
\text{John is someone who loves his wife.}
\end{align*}
\]

Here the pronouns ‘himself’ and ‘his’ are to be construed in conformity

---

16 Pace Geach, result (vi) does not preclude the possibility of assigning a referent to other sorts of occurrences of pronouns (e.g., indexical or “donkey” occurrences).

Soames claims (op. cit., pp. 204–205n11) that the Bound Variable Theory provides a “different but empirically equivalent way of conceptualising” his own proposal. However, the failure of either the Polyadic-Predicate Theory or the Abstraction Operator Theory to achieve results (i), (iii), (iv) and (vi) would seem to demonstrate that they are not alternative conceptualizations of the Bound Variable Theory. (On the Abstraction Operator Theory, the pronouns in (1) and (2) can be seen as designating a certain “abstraction” function, which assigns a class of individuals, or the characteristic function thereof, to a function from variable-value assignments to truth values. Recall also that the Polyadic-Predicate Operator Theory fails to achieve result (v), and the Abstraction Operator Theory apparently fails to achieve result (ii).)

17 McKay (1991) offers a similar criticism of the Polyadic-Predicate Operator Theory. Soames (op. cit., pp. 204–205n11) characterizes the Bound Variable Theory as treating the pronouns in (1) and (2) “as variables bound by lambda operators introduced at the stage in the evaluation of the sentence at which the antecedent of the pronoun is specified.” I do not know what this means exactly, but whatever it means, it is probably incorrect. The syntax of (1) and (2) should not depend on notation introduced (presumably as constituents of other expressions) only at various stages in their “evaluation.”
with the Simple Anaphor Theory. They are anaphoric here not on 'John', but on the bound variable 'who' (or on its trace, in Chomsky's sense).

The Bound Variable Theory does not require Geach's view that all pronoun occurrences other than pronouns of laziness (even so-called E-type or donkey occurrences) are bound variables. There are always the indexical or deictic occurrences. Still, a similar hypothesis might even accommodate a reflexive reading for sentences in which an anaphoric pronoun occurrence is not c-commanded by its antecedent, as in:

If John married Joan, then John loves her

Anyone who marries Joan loves her.

The former sentence, for example, on its alleged reflexive reading, is supposed to express something like that Joan is loved-by-John-if-married-by-John. The alleged reflexive reading of these sentences, however, seems strained. It is questionable whether the Bound Variable Theory should be extended this far.18

REFERENCES


18 Cf. Soames (1989/90, pp. 211–212). McKay objects that the proposed account does not seem to extend to more complex cases, like 'Mary wants the money that John left to his wife'. The anonymous referee made a similar criticism using the sentence 'Seeing himself promoted would be good for John's self-image'. These constructions, however, do not present insurmountable obstacles to the proposed account. They may be regarded as shortened versions of the following:

Mary wants the money which is such that John is someone who left it to his wife

Being someone who sees himself promoted would be good for John's self-image.

If the Bound Variable Theory is applied to the allegedly problematic sentences, its point is precisely that these sentences have readings something like those given by the displayed sentences, so that the pronouns 'him' and 'himself' are not referring singular-term occurrences anaphoric on 'John' but variable occurrences bound by an abstraction operator.
Salmon, Nathan: 1986a, Frege’s Puzzle, Ridgeview, Atascadero, California.

Department of Philosophy
University of California
Santa Barbara, CA 93106
U.S.A.