Quotation is an apparently very simple procedure, yet even here semantic theorizing has been able to find opportunities for involved argument. The Fregean theory holds that the expression inside the quotation marks is the referring expression, while the role of the quotation marks is merely to indicate a context inside which expressions do not have their standard semantic value. The referring expression, when placed inside quotation marks, refers to itself. The Tarski-Quine theory holds that in quotation the referring expression is the whole formed by quotation marks and the expression between them. This entire expression, the quotation, should be viewed as lacking any semantically significant articulation, as one of the basic elements of the lexicon, "a single word." The quotation marks are from the semantic point of view like the serifs adorning my letters in the font here chosen. The Davidsonian theory holds that quotation marks are the true referring expression in quotations; the word inside them is not really a word, but a thing, a token, and is there to help fix the content of the referring expression. Quotation marks are like an abbreviation of the demonstrative description 'the shape of this thing', pointing toward the token inside them. Finally, there is still the Geach-Tarski-Quine mixed descriptive theory, about which I shall say something below.1

Its initial intuitive implausibility notwithstanding, there are forceful considerations favoring the Davidsonian theory against its rivals. But in a paper recently published in this JOURNAL, Corey Washington2 argues for the Fregean theory (what he calls the identity theory).

* I would like to thank Ramón Cirera, Ignacio Jané, and Begoña Navarrete for comments that led to several improvements. Several discussions with Corey Washington have also been greatly helpful. Research for this paper has been funded by the Spanish Government's DGICYT, as part of the project PB90-0701-C03-03.


Washington's argument has failed to convince me, however, although it has helped me to appreciate the strength of Donald Davidson's account, or, to be more precise, of a minor modification of that account. I shall present, first, what I take to be Davidson's arguments against the other theories. Then I shall offer my conservative modification of his original account, and show how it helps us to see other deficiencies in them, especially in the Fregean theory. Finally, I shall explain why Washington's argument against the Davidsonian view does not withstand close scrutiny.

I

Davidson's arguments against the Tarski-Quine theory are powerful and convincing. I shall mention them here only because they point to facts that any correct theory should honor. First, quotation has properties that it could not have if the Tarski-Quine theory were correct. Quotation is a systematic and productive device. This systematicity is not purely orthographic or syntactic, but semantic, too. Alfred Tarski contends that the interpretation of quotations "which seems to be the most natural one" has it that

[t]he single constituents of these [quotation mark] names—the quotation marks and the expressions standing between them—fulfill the same function as the letters and complexes of successive letters in single words. Hence they can possess no independent meaning. Every quotation-mark name is then a constant individual name of a definite expression [. . .] and in fact a name of the same nature as the proper name of a man (op. cit., pp. 159–60).

W. V. Quine argues that "from the standpoint of logical analysis each whole quotation must be regarded as a single word or sign, whose parts count for no more than serifs or syllables" (op. cit., p. 26). According to Tarski and Quine, therefore, quotations are just orthographically articulated expressions, whose orthographic proper parts do not have any bearing on their semantic interpretation. But this view cannot be right. Compare: the systematicity of numerals is semantic, not merely orthographic or syntactic, because a competent user who knows what is necessary to interpret the finite stock of numerals he has encountered knows thereby something that determines the meaning of numerals not in that set. The systematicity of quotations is semantic precisely in that sense: a competent user who knows what is necessary to interpret the quotations he has in fact confronted knows thereby something that determines the meaning of new quotations. Thus, when we introduce a new expression in the

3 For a clear distinction between the two properties, see the appendix to Jerry Fodor, Psychosemanitcs (Cambridge: MIT, 1987).
language, its quotation name already has a meaning. This would lack any explanation if they were names “of the same nature as the proper name of a man.” The Tarski-Quine theory would require in that case the further introduction into the lexicon of the quote name of the new expression.

This is not the main defect of this theory, however, for there are equally incorrect theories of quotation that would nonetheless define a productive and systematic device. To understand what is truly specific in quotation, we must consider Davidson’s criticism of a different theory, suggested also by Tarski and Quine and adopted (in a slightly different variant) by Peter Geach, which cannot be attacked on the former grounds: the property defined by this theory is systematic and productive in the required semantic sense. According to this view, we first have a limited stock of quotation names of an equally limited stock of entities; quote names of letters, say, or of ordinary lexical units. These function according to the Tarski-Quine view; that is, they lack any semantically relevant structure. Then we form complex descriptions, having resort to the operation of concatenation. A description built along the lines suggested by the theory could be: ‘r’ concatenated with ‘u’ concatenated with ‘n’.

Davidson’s second argument against the Tarski-Quine theory also applies against what is left of it in the account given of the basic items in the modified Geach-Tarski-Quine theory. The gist of the argument is that the theory does not explain the essential element of “picturing” in quotations. Nothing prevents me from introducing a new lexical unit to name a name; for instance, I may well introduce the expression ‘villar’ to refer to the ordinary name of the city where the Olympic Summer Games took place in 1992. According to this convention, then, ‘villar’ refers to ‘Barcelona’. If someone asks me, however, while in Barcelona, “What is the name of this city?” and I answer “The name of this city is villar,” I have not succeeded in giving him (that is, unless he already knows my convention) a way to speak of the city in the ordinary, conventional way. Moreover, he still lacks the conventional resources to express in a direct way the thought that Barcelona has a port; the most that he can say is something like “The city called villar has a port.”

4 Unless, of course, he misinterprets my Tarski-Quine quotation for a real quotation, understanding me as if I said ‘The name of this city is “villar”’, which is what would probably happen in any real-life situation. In this case, he would utter ‘Villar has a port’—which fails to express his thought, and besides, words lacking harbors, is obviously false.
in the modified Geach-Tarski-Quine theory; for with ordinary quo-
tation, we do give people the resources to make utterances using the
expression referred to by the quotation we have used, without fur-
ther ado (that is, without further knowledge on their part of a nam-
ing convention). If we did not have real quotations, if we were stuck
with the Tarski-Quine variety, we would need to invent them.\(^5\)

To use quotation marks in a logic course when speaking about an
object language is annoying. The cagey logician fond of the Tarski-
Quine theory who uses ‘∧’ instead of “∧” to refer to the object-
language conjunction sign tells us at the outset “we do not care how
the conjunction sign of the object language looks, for we shall not
use it; we shall merely refer to it.” A solution less mysterious for the
beginner is to declare that signs are used “as names of themselves.”
According to the Fregean theory, this is what we ordinarily do in
quoting expressions. Quotation marks are there merely to warn us
that the expression is not intended to make its ordinary contribution
to the truth conditions of the sentence in which it occurs. What
really refers, according to the Fregean view, is the expression inside
quotation marks; and it refers to itself.

Davidson’s objections to this theory are not very convincing (op.
cit.). It all seems to come down to the fact that the theory does not
explain in what consists the intuitive “picturing” occurring between
the quoted material and the expression referred to by the quotation.
But I do not see any problem here for the Fregean view. The tokens
we use in linguistic interchanges possess reproducible properties,
necessary for them to be linguistic entities; properties that deter-
mine a “reproductively established family.”\(^6\) We conveniently reify
them and call the result the type to which a given linguistic token
belongs. To be able to understand a linguistic token requires, at the
very least, to recognize those aspects, to recognize the type. When a
token of ‘Quine’ occurs in an utterance, it is that set of aspects
common to this and other utterances of similar tokens that in fact
does the referring to the Harvard philosopher; it is something that
recurs, types and not tokens, that can refer or have meaning. When
we see the token of an expression surrounded by quotation marks
we recognize the referring aspects as usual, and then the quotation
marks tell us that here the set of aspects constituting the referring
expression refer to themselves.

\(^5\) Quine acknowledges the pictographic character of quotations (op. cit., p. 26),
but apparently fails to see that his own preferred account does not accomodate it.

\(^6\) See Ruth Millikan, *Language, Thought and Other Biological Categories* (Cam-
bridge: MIT, 1984), ch. 1.
This theory, then, seems to account well for the facts we have exhumed in discussing the Tarski-Quine theory. It is productive and systematic in the required way, and it explains why, in giving you a quotation name of an expression, I thereby confer on you, without the further disclosure of any referring convention, the capacity to use that expression in speech and thought.

So does Davidson's own theory, according to which the quotation marks are demonstrative descriptions, somehow abbreviating "the shape of this thing," the demonstrative understood as pointing toward what is inside the quotation marks. This, the pointed-at entity, is not conceived of in the theory as itself an expression. The situation would not be logically different if the thing inside the quotation marks were outside, something visibly physical like the word 'Quine' Kaplanwise made out of big neon tubes, to impress forcibly in us the idea that it is simply a token that helps the referring act, and the sentence were \( \langle ' ' \rangle \) is disyllabic \( \rangle \). Remember, \( \langle ' ' \rangle \) abbreviates "the shape of this thing." For \( \langle ' ' \rangle \) is disyllabic \( \rangle \), according to the Davidsonian theory, must be the real sentence; as we said before, linguistic meaning (and truth conditions) is conveyed only by types.

It must be readily acknowledged that this is indeed a bit strange; Washington builds his argument against Davidson on this weirdness, and in so doing he makes clear in what exactly the weirdness consists. I shall discuss the argument in the last section. But, leaving that aside, at this moment in the argumentative dialectic Davidson's account is on a par with the Fregean theory.

We did not take a stand on what those aspects of tokens are in virtue of which they count as linguistic expressions; but Davidson does take a stand. He seems to think that they are shapes. This is usually true, but it is not the whole truth, and it leaves out what is essential to the possession of linguistic meaning. This fact, as I shall point out below, unduly restricts Davidson's theory. But we shall not have to abandon its main elements to avoid that restriction. And in modifying Davidson's original account, we shall take advantage of its real strength and find a better reason than Davidson's to prefer his view of the matter to the Fregean theory.

II

The point of Ludwig Wittgenstein's considerations about ostensive definitions in *Philosophical Investigations* §§27–36 is not, against what might be understood, that ostensive definition is somehow faulty, or even impossible. You can, for instance, define 'indigo' by

---

\(^7\) For a passage in which Wittgenstein's discussion seems to be understood in the way I am claiming is faulty, see Quine, "Ontological Relativity," in *Ontological Relativity and Other Essays* (New York: Columbia, 1969), p. 31.
correlating the word with a sample, and there is nothing wrong with
the procedure. Wittgenstein’s point is that, in so doing, you are not
somehow directly correlating words with their meanings. For mean-
ing is normative; the meaning of an item determines a distinction
between events involving it (say, uses of ‘indigo’ as a sentence) which
are correct and events which are incorrect, and a sample (even if it is
mental), by itself, does not provide for that. A sample can be “inter-
preted” in whatever way, and every interpretation determines a dif-
ferent criterion of correctness, which is to say that the sample alone
does not determine a criterion of correctness. Wittgenstein’s point
is instead the familiar one that it is not the sample, but the use we
make of it, that determines the meaning of the ostensively defined
term. (This is why he proposes to consider the samples as part of the
language; their “meaning,” too, must be accounted for, and cannot
be assumed as somehow already given.) The use is very different
when we define the name of a river by pointing at a stretch of it,when
we define the name of a person by pointing at a temporal
segment of him, when we define the name of a kind by pointing at
an instance of it, or when we define the name of a musical tone or
the name of a color by giving samples.

Let us say that an ostensive sign consists of a demonstrative (‘this’),
a category term (‘sound’) — which may be merely implicit in the con-
text — and a demonstration (the act of pointing, or directing one’s
glance) toward a certain token, a certain physical thing, the stretch
of a river, a (spatio-temporal segment of a) person, a patch of color,
a sound. Ostensive signs can be used for whatever purpose, in giving
a definition or in making a statement. The difference between an
ostensive sign and an expression we define by means of it does not
lie in the contribution each one makes to the truth conditions of the
sentences in which it occurs. This contribution is the same for ‘in-
digo’ as for ‘this color’ (said while pointing at a sample of indigo) in
‘Fred was wearing an indigo sweater’ and ‘Fred was wearing a
sweater this color’, respectively. The difference lies in that a being
cognitively similar to us, assuming he understands demonstratives
and (perhaps only “implicitly”) the appropriate category terms, is
able to get without further ado the content of the ostensive sign;
while the content of the nonostensive term must be explicitly ex-
plained to him. The relations between the pointed-at token and the
content of the ostensive sign, different as they are in the examples I

\footnote{What is the difference between an ostensive sign and a plain demonstrative? Perhaps none. Let us just say that an ostensive sign is a plain demonstrative that can be used in giving ostensive definitions to a normal human being. Maybe the two classes are coextensional; we do not need to bother here about that.}
have given, are somehow natural to us human beings; the relation between expression and content in nonostensive terms is purely conventional.

The essence and main strength of Davidson’s account of quotation, as I see it, lies in its treating quotations as ostensive signs. This provides for useful unification; there is nothing semantically different, according to this view, in (i) the way the quotation in ‘pronounce this word: “expresso” ’ works, (ii) the way the ostensive sign composed of ‘this sound’ and the assumed sound token works in ‘repeat this sound: . . .’, and (iii) the way the ostensive sign composed of ‘this color’ and the assumed color patch works in ‘reproduce this color: . . .’. And this seems intuitively correct; all these cases should be treated similarly by a semantic theory.

Seeing that the essence of Davidson’s proposal lies in its treating quotations as ostensive signs, as I have just proposed, we can easily solve a minor defect in Davidson’s original version of the proposal. As I said before, he says that the content of a quotation is the shape of the quoted material. But this cannot be so, as one of the examples before easily proves: ‘pronounce this word: “expresso” ’. (The instruction must be imagined as given as here, by means of an inscription and not of an utterance.) It is not the shape of the material quoted in the inscription which you are asked to articulate to produce a concrete sound. Another example of the same problem is provided by an inscription of the following explanation: “An onomatopoeic term is a sign that resembles its meaning; for instance, ‘hiss’.” It is not in virtue of the shape of the material quoted in this inscription that ‘hiss’ is an onomatopoeia.

To think of quotations as species of the kind “ostensive sign” solves this problem. The content of an ostensive sign need not be the shape of the demonstrated token. It must only be something naturally related to it. In “Words,” David Kaplan9 defends the idea that tokens belong to a common word when they are linked by a causal chain. I myself prefer to think of linguistic expressions, particularly words, as structures having a common function, in the teleological sense of ‘function’. If we then render ‘function’ in a causal-historical way, the tokens of the same word are then also linked by a causal chain (for they belong to what Ruth Millikan calls “a reproductively established family”; op. cit.). Their common function is essential for our classifying them together; but they must also have some common inherited character, constituted by those features whose reproduction is accounted for by their causal role in the

---

performance of the function. This functional apparatus provides a new commonality over and above observable shape unifying some expressions of Spanish with certain expressions of English (for instance, the spoken version of 'Tony Curtis' typically produced by a Spanish speaker and its English counterpart), and more obviously the spoken and the graphic shape of "the same word." We find here the mechanism through which the tokens in the examples in the previous paragraph produce the content of the ostensive signs invoking them; for those mechanisms are constituted by relations "natural" to us.

To think of quotations as ostensive signs gives us therefore a first reason to prefer Davidson's theory to the Fregean. For in this way we unify what intuitively should be unified, namely, the semantic mechanisms accounting for the way the quotation in 'pronounce this word: "expresso"' works, the way the ostensive sign composed of 'this sound' and the assumed sound token work in 'repeat this sound: . . .', and the way the ostensive sign composed of 'this color' and the assumed color patch work in 'reproduce this color: . . .'.

There is one more reason in favor of Davidson's theory. Ostensive signs are flexible; two uses of some ostensive signs, which as a matter of fact have the same content, might however have had different contents if used in different contexts. It is the same with quotation. There are contexts in which the quotations "Madrid" and "Madrid" would have the same content, but there are easily conceivable contexts in which they would have different contents (for example, contexts in which you manage to refer to the general type "cursive word," or to the more specific cursive version of 'Madrid'). The Davidsonian theory easily accounts for this; the Fregean could do that, too, but only in a convoluted, ugly way (for instance, positing an implicit "the cursive version of the expression . . ." before the first quotation, in those contexts in which the content is different). But this consideration, taken together with the preceding one and with general methodological considerations favoring the virtues of simplicity and conservatism, favors the Davidsonian theory.

Quotation is an interesting phenomenon by itself, but there is almost always a hotter issue lurking behind when philosophers discuss it, namely, opacity. The Fregean wants his account of quotation as an intuitively plausible first step to introduce, with it as a model, a very problematic account of the semantics of indirect discourse, attitude reports, and modal statements. The Davidsonian sees his own views therefore as a first stab at restoring semantic innocence. I am no exception to the rule; but although I think that the applica-
tion of these points to the general problem of opacity would also favor Davidson’s theory, I cannot dwell on the matter here.

III

Washington levels two arguments against Davidson’s theory. The first is that the theory is unduly inflexible, because it allows for reference to a single type of entity (a shape) for every quotation. I have shown in the previous section that this is not part of what we may well consider the essence of Davidson’s proposal. On the other hand, the Fregean theory, unmodified with some proviso along the lines I myself suggested to the Fregean before, actually is unduly inflexible. According to the pristine Fregean theory, the quotation of an expression has always the same content, the expression itself. But this is not so, as we saw before; by quoting tokens of the same expression we can refer to many different things: the expression (‘“gone” is dissyllabic’); different types instantiated by the tokens (‘“gone” is cursive’); different types somehow related to the token (say, the graphic version of the uttered quoted material, or the spoken version of the inscribed quoted material, as in ‘“gone” sounds nice’); different tokens somehow related to the quoted token (‘What was the part of the title of the movie which, by falling down, caused the killing?—“Gone” was’); the quoted token itself (‘At least one of these words is heavier than “gone,”’ which you should imagine written in big wooden letters); etc. Viewing quotations as special cases of ostensive signs nicely accounts for the semantics of these examples. The flexibility of quotation does not give us an argument for Frege and against the Davidsonian who holds that quotations are ostensive signs; just the opposite is true.

Washington’s main argument tries to turn the weirdness of Davidson’s theory into a clearly false consequence of it. The weirdness I am pointing to is the fact that the quoted material, according to Davidson’s proposal, is not actually part of the uttered expression. The role of the uttered material is the same as that of a demonstrated object in the extralinguistic context, as we have pointed out before. That is, those properties in virtue of which the quoted material in an utterance of ‘Barcelona’ has nine letters is a linguistic expression are not included among the linguistically significant properties of such an utterance. The quoted material figures there simply as a token, it is “a mere thing.”

Now, Washington’s point is that, although a sentence including a quotation could be uttered with appropriate correlates of quotation marks (‘quote-unquote,’” finger gestures, a special intonation), many times we do not do so. But, according to Davidson’s theory,
the quotation marks themselves are the subject of “‘Barcelona’ has nine letters”; remember that the quoted material is not an expression. Therefore, when uttered without any of the oral correlates of quotation marks, “‘Barcelona’ has nine letters” will lack a subject. It will therefore be grammatically ill-formed (589). A vast majority of our utterances involving quotation will then be syntactically ill-formed, against our clear intuitions, it is assumed.\(^\text{10}\)

Although this nicely makes clear what is peculiar in Davidson’s theory, it strikes me as a very weak argument. Let us consider what from a Davidsonian viewpoint are similar cases. In Spanish, the only usual oral difference between the assertion ‘Vino un policía ayer’ (‘A police officer came yesterday’) and the question ‘¿Vino un policía ayer?’ (‘Did a police officer come yesterday?’) is a difference in intonation. Now, there are contexts in which you can intonate the uttered sentence as an assertion, while intending for it to be understood as a question, and you probably will have success. (There are many other examples like this, in every language: expressions which can orally only be distinguished by their stress, and which you can intonate in some contexts one way meaning the other, and get away with it, etc.) How are we to describe these situations theoretically? There are several available possibilities, and we do not need here to choose between them. First, we must decide whether we shall posit a syntactical correlate of the semantic distinction between assertion and question. If so, we must decide what to count as realizations of that syntactical correlate: we could as well admit contextual clues as such realizations, and not only the conventional intonation, stress, etc. Now, a reasonable possibility is for our grammar to require conventional indications of the difference between questions and assertions. In that case, we would give Gricean pragmatic explanations of the former phenomena. A token of ‘Vino un policía ayer’, intonated as an assertion, would assert (the conventional meaning) that a police officer came yesterday; we would then explain how the audience understands it as a question in the context by resorting to the mechanism of conversational implicatures. (“He cannot be making an assertion, if rational and cooperative, for I know that he lacks the evidence for it, and he knows that I know . . . ; therefore, he must be conversationally implicating that he wants to know whether a police officer came yesterday.”)

\(^{10}\) It seems that a similar point could be made about inscriptions, but in that case Washington thinks that grammatical correctness does require the use of quotation marks. I shall not discuss his views on inscriptions, for what I would have to say can be gathered from what I shall say about utterances.
There are similar options available for quotations, assuming that linguistic theory has finally decided that a Davidsonian account is inescapable. In this case, we shall probably decide that there must be a syntactic correlate, for they can act as the subject, etc. Then we must decide whether we want to require a conventional uniform realization; and here, as before, we may well decide to accept contextual clues (say, that the predicate is metalinguistic) as realization enough. Even not doing so, we could easily explain what happens in the controversial cases by having recourse to the mechanism of conversational implicature. In this case, an utterance of what we intend to be ≪‘Barcelona’ has nine letters≫ without any conventional realization of quotation marks would count as an utterance saying that Barcelona—the city has nine letters. Then we would explain how the audience gets the intended meaning through an easily derivable conversational implicature. (“He cannot be saying of Barcelona, the city, that it has nine letters, for this is false, and he must know that it is false; . . . ; therefore, he must be conversationally implicating, about the expression he has used a token of, that it has nine letters.”)

Washington mentions, while discussing this issue, that several people have answered his criticisms by claiming that quotation marks are “implicit” in those utterances lacking any of the conventional oral correlates of them. He disregards this suggestion by interpreting it as the claim that the missing quotation marks would be theoretically on a par with the “empty categories” posited by contemporary linguistic theories, as, for instance, the implicit subjects in some sentences of languages that, like Spanish, admit of sentences lacking explicit subjects (589–90). I agree with him that this is mistaken. I have shown, however, that there are two further possible ways of understanding the idea of “implicit” quotation marks, both of them, it seems to me, in principle unobjectionable: to consider obvious contextual clues as realizations of them; to explain their appearance by resorting to the mechanism of conversational implicature. In neither of these two understandings are we attributing any linguistic mistake to the speaker.

We must acknowledge, I believe, that the Fregean theory has some advantage regarding the explanation of these facts, for, ac-

11 In the presence of a very salient relevant object, the speaker says (while pointing with his eyes or hand and showing distress): ‘is stupid’. We can take the contextual facts as realization of the implicit demonstrative, and count it as grammatical, or adopt the Gricean interpretation. We do not need to posit a heretofore unknown empty category; likewise with the missing quotation marks.
according to it, quotation marks are just warnings of a shift in reference; when the context leaves it clear that the shift has taken place, it seems natural to overlook them. An analogy might be helpful. Our linguistic theory could well require syntactical correlates of differences in the semantic value of referring expressions; there should then be differences between proper names of different persons which happen to have (the names, of course) the same shape. In view of the facts, though, it seems perfectly proper here to accept contextual clues as such correlates, for this is what we actually find in our uses during the same conversation of names with the same shape to refer to different persons. The Fregean theory can naturally claim that quotations constitute a similar phenomenon; when the shift in reference is made clear by the context, we do not need to use any conventional means to indicate that. Davidson’s theory is here on a flimsier basis; to accept contextual clues as the realization of the subject of a sentence, as in my first interpretation of the idea that the missing quotation marks are “implicit,” seems difficult to swallow. And it must also be acknowledged that the second suggestion (to invoke a conversational implicature) leaves the Davidsonian in a worse position than the one occupied by the Fregean.

Nevertheless, and even taking into full regard such an acknowledgment, our final decision must be the result of a careful weighing of the different theories’ pros and cons; and, on balance, I think our bet should be in favor of the version of the Davidsonian view I have proposed here. It can give a sensible answer to the problems discussed in this section, while maintaining the advantages pointed out in the preceding one.

MANUEL GARCÍA-CARPINTERO

University of Barcelona

TESTIMONY, TRUST, KNOWING*

If we know much of what we think we know, then we do so through testimony. Testimony only succeeds if there is trust. But how could a right to be sure rest upon so fragile a basis as trust? Exploiting a number of such seeming paradoxes, John Hard-

* My thanks to L. Jonathan Cohen, Catherine Elgin, and Georges Rey for helpful discussion.