
Quotational Ambiguity

According to W. Quine,
 Whose views on quotation are fine,
 Boston names Boston,
 And Boston names Boston,
 But 9 doesn't designate 9.

Richard Cartwright used to assign to MIT graduate students in philosophy the exercise of supplying quotation marks to that underpunctuated limerick of his so that it says something correct and sensible. One solution is to put pairs of single quotes around the first and fourth occurrences of 'Boston' and a pair of quotes within quotes around the third. Another is to put the single quotes around the second and third occurrences and the quotes within quotes around the first. One of the lessons of this paper is that neither of these solutions is entirely unexceptionable.

It was Quine's *Mathematical Logic*¹ that was responsible for my becoming a philosopher. I came upon a copy of it in the university bookstore during my freshman year; a year later the instructor in "Advanced Logic" counted my having read it as satisfying the course's prerequisite. If I was a bit murky on alphabetic variance and such laws as *Math Logic's* *159:

If α is not free in φ , $\vdash \ulcorner (\alpha)(\varphi \vee \psi) \equiv .\varphi \vee (\alpha)\psi \urcorner$,

I thought I had a pretty good understanding of such arcana as the ancestral and quasi-quotation. I was, I was convinced, an ace on the ordinary kind

First published in Paolo Leonardi and Marco Santambrogio, eds., *On Quine*, Cambridge: Cambridge University Press, 1995, pp. 283–296. Reprinted with kind permission of Cambridge University Press.

I am grateful to David Auerbach, Martin Davis, W. D. Hart, James Higginbotham, David Kaplan, Michael Kremer, Harold Levin, Ruth Marcus, Charles Parsons, W. V. Quine, Nathan Salmon and Göran Sundholm for helpful comments. Research for this paper was carried out under grant no. SES-8808755 from the National Science Foundation.

¹(Quine, 1955).

of quotation.

Less than a decade later I was explaining

‘yields a falsehood when appended to its own quotation’
yields a falsehood when appended to its own quotation.

to unfortunates in introductory philosophy who were expecting the meaning of life. Eventually, though, I wound up teaching courses with titles like “Paradox and Infinity” or “Logic II,” for which the Good Stuff was more appropriate.

In MIT’s “Paradox and Infinity” a few years ago, as I was going over “yields a falsehood . . .,” an undergraduate suggested that what I had written on the board, something like:

‘blue’ appended to the quotation of ‘red’ = ‘ ‘red’ blue’

was ambiguous, and that I needed two kinds of quotation marks, semantic and syntactic, to say what I wanted to say.

“Groan,” I thought. “Another cockamamie undergraduate suggestion. No undergraduate has anything to teach *me* about quotation.” I couldn’t have been more wrong.

The student’s name was Michael Ernst (‘Michael Ernst’?); he is now a graduate student in computer science at MIT. And this paper is about his observation.

When I explained Ernst’s observation to Cartwright, he doubled over in surprise and uttered an oath. I wrote to Quine about it, who replied, “Dear George, Thanks for Ernst’s paradox. I am delighted with it. But I find I am unable to cope with it, even when I have stopped laughing. Yours, Van.”

Here’s the problem. *Expressions*, or *strings*, are (or may be identified with) finite sequences of symbols. We’ll use letters of the Greek alphabet as variables over expressions.

Now where α and β are any expressions at all, β appended to α is the expression obtained by first writing the symbols of α , in the order in which they occur in α , and then writing immediately after these the symbols of β , in the order in which they occur in β . Thus, for example, ‘apple’ appended to ‘pine’ is the expression ‘pineapple’. The operation of appending is associative: γ appended to (β appended to α) is identical with (γ appended to β) appended to α . Thus it does not matter how we add parentheses to such a “term” as: ‘VAKIA’ appended to ‘OSLO’ appended to ‘CZECH’. (‘OSLO’, it is well known, is in ‘CZECHOSLOVAKIA’.)

So, it would seem, ‘b’ appended to ‘a’ is the two-symbol expression ‘ab’, consisting of the letters ‘a’ and ‘b’ in that order.

The quotation of α is the expression that results when the expression α is enclosed in a pair of quotation marks, i.e., the result of writing a left

quote, then the symbols constituting α , and then a right quote. Thus the quotation of 'Boston' is ' 'Boston' '.

The quotation of α is supposed to be an expression whose denotation is α ; e.g., the denotation of the eight-symbol expression ' 'Boston' ' is the six-symbol expression 'Boston'.

And I hope you are all familiar with the calculation:

'appended to its own quotation' appended to its own quotation
 = 'appended to its own quotation' appended to the quotation of 'appended to its own quotation'
 = 'appended to its own quotation' appended to ' 'appended to its own quotation' '
 = ' 'appended to its own quotation' appended to its own quotation'.

The calculation shows us that there is an expression, viz., 'appended to its own quotation' appended to its own quotation, which denotes itself.

So far all is familiar. But now consider the nonsense string κ :

b' appended to 'a

κ consists of the second letter of the alphabet, a right quote, a space, the eight letters of a certain word, a space, the two letters of a certain other word, a space, a left quote, and the first letter of the alphabet, in that order of course.

There are a lot of things one might want to say about κ . It's ill-formed, it's a nonsense string, it does not contain the letter 'c', it begins with the letter 'b', the number of letters in the English alphabet that alphabetically precede the first letter of κ is one, etc.

Consider now another string, λ :

ab

λ too is not well formed, nor does it contain the letter 'c'. Unlike κ , though, it begins with the letter 'a', and the number of letters of the English alphabet that alphabetically precede its first letter is zero.

Let's write: $N(\alpha)$ as short for: the number of letters of the English alphabet that alphabetically precede the first letter of the expression α . Then $N(\kappa) = 1$ and $N(\lambda) = 0$.

Now consider μ :

'b' appended to 'a'

μ is the quotation of κ ; thus the denotation of μ is κ .

But wait! What did we say earlier? μ , note, consists of the quotation of 'b' followed by ' appended to ' followed by the quotation of 'a'. Just as 'apple' appended to 'pine' is 'pineapple', 'b' appended to 'a' is surely 'ab'. Thus μ , which begins with a left quote followed by a 'b' and which is the subject of the second clause of the previous sentence, denotes 'ab', and the denotation of μ is λ , i.e., 'ab'.

So $0 = N(\lambda) = N(\text{the denotation of } \mu) = N(\kappa) = 1$.

What has gone wrong? Obviously, the non-identical κ and λ cannot both be the denotation of μ , and unless they are, our demonstration that $0 = 1$ fails. How did we conclude that they are identical?

In the case of κ we said: μ is the quotation of κ ; thus the denotation of μ is κ . We also said: $\mu \dots$ denotes 'ab', and the denotation of μ is λ . It would thus seem that the inferences:

β is the quotation of α ; so α is the denotation of β .

and:

β denotes α ; so α is the denotation of β .

are problematic. In any event, μ is ambiguous, for on different parsings it denotes the different expressions κ and λ . And that is Michael Ernst's observation.

Let's treat the matters with which we have been dealing somewhat more formally. A theorem about quotation, appending, and denotation will ensue.

Expressions, as we have said, are finite sequences of symbols, i.e., functions from a finite initial segment of the set of natural numbers to a set of symbols. (Thus, e.g., 'cat' is a function whose domain is $\{0, 1, 2\}$ and whose value at 1 is 'a'.) For any symbol s , $[s]$ is the expression whose sole symbol is s . It is often important to distinguish between s and $[s]$, but it will turn out that we can here safely identify the two.

We define the *length* $\text{lh}(\alpha)$ of the expression α to be the least natural number not in the domain of α . Intuitively, $\text{lh}(\alpha)$ is the number of occurrences of symbols in α . Thus $\text{lh}(\text{'cat'}) = 3$, $\text{lh}(\text{'cattle'}) = 6$ and $\text{lh}(s)$, i.e., $\text{lh}([s])$, = 1.

If α and β are expressions of lengths m and n respectively, $\alpha * \beta$ is the expression of length $m + n$ whose first m symbols are those of α in the order in which they occur in α and whose last n symbols are those of β in the order in which they occur in β . (Thus $*$ is associative.) We will sometimes omit asterisks if (we believe that) no confusion will result.

Now assume that L is a language containing two symbols, l and r , and possibly others. Suppose that (1) for any expression α of L , $[l]*\alpha*[r]$ denotes

α . (Thus l and r work like left and right quotation marks.) Suppose further that there is an expression ϑ of L such that (2) for any expressions $\alpha, \beta, \gamma, \delta$ of L , if α denotes γ and β denotes δ , then $\alpha * \vartheta * \beta$ denotes $\gamma * \delta$. (Thus, like, e.g., ‘followed by’, ϑ “denotes concatenation.” Note that ‘appended to’ switches order, but ‘followed by’ does not.)

Theorem *Some expression of L denotes two different expressions of L .*

Proof. Let $\alpha = [l] * [l] * [r] * \vartheta * [l] * [r] * [r]$, $\beta = [l] * [r] * \vartheta * [l] * [r]$, and $\gamma = [l] * [r]$. By (1), α denotes β . By (1), $[l] * [l] * [r]$ denotes $[l]$ and $[l] * [r] * [r]$ denotes $[r]$. Then by (2), α denotes γ . But $\beta \neq \gamma$. ■

An English version of α is:

‘ ‘ followed by ‘ ’ ’

which, as we see, ambiguously denotes the two-symbol expression consisting of the left and right quotes in that order and also denotes a longer expression containing oddly placed quotes, spaces, and certain letters of the English alphabet.

I suppose that by now it is hard to resist the suggestion that our puzzles arise from the circumstance that when a left quote is followed by two or more right quotes, it may not be determined which of those right quotes its mate is. Quotation marks differ in this respect from parentheses: any sequence of left and right parentheses is well formed in at most one way. (A left parenthesis and a right parenthesis to its right are mates if (recursively) no unmated parentheses lie between them; a string of parentheses is well formed iff every parenthesis in it has a mate.) In a slogan: “Quotes don’t know their mates.”

Nevertheless, we should look at the question whether the ambiguity arises from the phrase ‘appended to’. Would our problems disappear if we were to abolish ‘ β appended to α ’ and to substitute for it ‘the result of appending β to α ’ instead? (Of course our English versions of the semantical paradoxes would thereby become rather more complex—but that is of small concern.) ‘The result of appending β to α ’ resembles function terms such as ‘ $f(x, y)$ ’ in mathematics more closely than does ‘ β appended to α ’; could this deviation from a syntax resembling that of formal languages be responsible for our difficulties?

The suspicion that it’s not the use of ‘appended to’ that has made trouble is easily confirmed if we consider:

the result of appending ‘c’ to the result of appending ‘b’ to ‘a’

Does this phrase denote the improbable:

ac’ to the result of appending ‘b

or does it denote the more likely:

abc ?

Would parentheses help? Suppose that instead of writing ‘ α appended to β ’ we were always to write ‘ $\text{app}(\alpha, \beta)$ ’ instead? But that really wouldn’t avail. We would still have:

$\text{app}(\text{‘a’}, \text{‘b’}, \text{‘c’})$

to deal with, which may at first appear ill-formed, but need not be considered so. However, if it is not taken as ill-formed, it must be regarded as ambiguous, for on one way of parsing it, it denotes:

ca’, ‘b

and on another, it denotes:²

b’, ‘ca

I want now to discuss some proposals for dealing with the difficulty that modify or enhance quotation, or replace it with other devices. Eventually I shall offer one that seems to be more or less satisfactory. But first the others.

The first suggestion that comes to mind is to draw a link connecting a left quote with its right mate. We might do this by means of an arch, thus:


‘a’ appended to ‘b’

Or we might enclose in a box the material we wish to quote. Another suggestion is that a pair of quotation marks should be regarded as a discontinuous, two-part, symbol, like the letter ‘Ы’ of the Cyrillic alphabet, or ‘i’ with its dot. In a language in which a pair of quotation marks is a single symbol, it would be no more possible for an analogue of our nonsense string κ to exist than for ‘*lylophonex*’ to be a word of English.

These proposals have a common defect. Were we to adopt one of them, we could no longer regard expressions as finite sequences of symbols. Were we to introduce arches, boxes, or something similar into the language, or to regard pairs of quotation marks as single but discontinuous symbols, between whose two parts entire expressions could occur, we would violate a requirement of *sequentiality*: expressions must be codable as functions from a finite initial segment of the natural numbers into the set of symbols of the language, i.e., as finite sequences of symbols. In any case we will see how

²Michael Kremer pointed out to me that Anil Gupta made a similar observation in (Gupta, 1982), pp. 184–5.

to link the expressions that will turn out to be our substitute for quotation marks without the use of arches, etc.

We might dispense with quotation marks altogether, *italicizing* instead those expressions that are to be referred to. Italicization, of course, has an evident drawback: iterating it is, at best, difficult. What angle with the abscissa shall a capital “I” make after it has been italicized n times? $90/2^n$ degrees?

There is a less noticeable, but more important, difficulty: it would seem that an italicized symbol would have to be regarded as other than the roman version of that symbol, and a doubly italicized symbol as different from both the singly italicized and the roman versions. Were we to allow, then, that the m - and n -fold italicizations of any symbol of a language L are different symbols of L if $m \neq n$, then L would have to be infinite. We desire a solution that does not require the number of symbols in L to be infinite. We may call this requirement the *finiteness* requirement.

Boldfacing obviously satisfies the finiteness requirement no better than italicization.

We have several times displayed expressions, that is, surrounded them with white space, or written them after colons. Although we shall continue to do so when quotation might be impractical or confusing, it is clear that as a general device for quotation, displaying offers difficulties akin to those of italicization, boldfacing, and other ways of changing the style of type.

If underlining is not to fall afoul of the requirements of sequentiality and finiteness, it would appear that we to have to regard the result of underlining an expression α as the result of repeatedly attaching a single symbol, say ‘_’, after (or before) each of the constituent symbols of α . Thus if we wish to regard expressions as consisting of symbols in a linear order, we would seem to have to take the double underlining of ‘cat’ as the nine-character expression ‘c_a_t_’. (‘_’ contains two expressions; ‘_’, one.) But then, of course, expressions like the three-character ‘__’ are ambiguous as between a non-underlined ‘__’, a ‘_’ whose first ‘_’ is underlined, a ‘_’ whose second ‘_’ is underlined, and a doubly underlined ‘_’.

In conversation, Ruth Marcus proposed the use of an infinite sequence of ever bigger quotation marks. The suggestion obviously violates the finiteness requirement if each of the infinitely many quotation marks is to be a single symbol; but I confess that it was only after she made it that the finiteness requirement struck me as one that needed to be imposed, and our suggestion will actually turn out to be not so very different from Marcus’s.

It seems extremely plausible that a language containing a fixed, finite number of pairs of quotation marks, e.g.:

‘ ’ { } [] () [[]]

would suffer from ambiguities of the sort we have noticed, but I have no rigorous proof. If we were to order the different pairs of marks and stipulate that no pair earlier in the ordering may surround either member of any later pair, then we would evidently violate another requirement, perhaps the most important one of all, that of *universality*: that for every expression α of the language, no matter how ill-formed, there should exist a (well-formed) expression of the language obtained by enclosing α in a pair of quotation marks. The requirement of universality rules out such devices as Quine's of spelling³ as cures for our woes, for we wish to insist that the quotation of an expression α actually contain α as a subexpression.

Three familiar features of the ordinary use of quotation marks may be remarked upon.

1) The distinction between left and right quotation marks seems to be of little significance. As those of us who use crude fonts know, one can almost always get by with: ' or: ". In any case our original statement of Ernst's problem mentioned left and right quotes, and abolishing the distinction would certainly not get us out of our jam.

2) According to Collins' *Authors' & Printers' Dictionary*,⁴

quotations within quotations to have only single quotation marks within the double. "The more conspicuous mark to the more inclusive quotation" (Henry Bradley). Quotations within the single quotation, to be double-quoted.

Presumably also, quotations within the double quotation within the single quotation, to be single-quoted, etc. Strict conformity to standard typographical practice, which I take Collins to be describing, is in any case impossible: if we wish to praise X for his adherence to the rule in writing: He asked, "Why not?", then we can't write:

Kudos to X , who wrote, "He asked, "Why not?""

but must, in accordance with Collins' rule, write:

Kudos to X , who wrote, "He asked, 'Why not?'"

But how then are we to criticize Y for violating the rule by writing: He asked, 'Why not?' , since we can't justifiably write:

Fie upon Y , who wrote, "He asked, 'Why not?'"

(for why are we praising X but blaming Y ?) and can't adhere to Collins and write either:

³(Quine, 1960), p. 143.

⁴(Collins, 1933), p. 314.

Fie upon Y, who wrote, “He asked, “Why not?””

or:

Fie upon Y, who wrote, ‘He asked, “Why not?” ’

or:

Fie upon Y, who wrote, ‘He asked, ‘Why not?’”

One possibly unfortunate feature of standard usage is that quoting a quoted expression changes it: double quotes become single and single double, as each new pair of outermost double quotes is added.

3) It is commonly thought that the primary function of quotation marks is to produce an expression that refers to another expression, namely the one inside the quotation marks. Perhaps so, but “shudder” quotes, here illustrated, the custom of enclosing titles of works in quotation marks,⁵ and the (substandard) use of quotation for emphasis suggest a weaker view: the primary function of quotation marks is to indicate to the reader that quoted expressions are to be treated in some special manner that is inferable from context (it is to be hoped), but is otherwise unspecified.⁶

We come now to our own proposal, according to which to quote an expression is to enclose it in a certain pair of syntactically complex expressions, called the *q*-marks of that expression; the rest of the proposal is a rule for parsing a given expression α to determine which *q*-marks occurring as subexpressions of α are the mates of which others.

Let ' and ° be two new symbols. $\{j\}$ is the expression consisting of j consecutive occurrences of ' ; $\text{lh}(\{j\}) = j$.

A *q*-mark is an expression $\{j\}^\circ$, where $j \geq 0$. Thus the *q*-marks are °, '°, '°°, '°°, etc.

is the null expression; $\text{lh}(\#) = 0$.

For any expression α , $m(\alpha)$, α 's *q*-mark, is the shortest *q*-mark that is not a subexpression of α . Thus, for example, $m(\#) = m(') = m(s) = \circ$ (s any symbol other than °), $m(\circ) = '^\circ$, $m('^\circ) = ''^\circ$, $m('''^\circ) = ''''^\circ$, $m(\circ')$ = '°°, $m(\circ^\circ) = '^\circ$, $m('') = \circ$.

The (revised) quotation of α , $r(\alpha)$, is defined to be $m(\alpha) * \alpha * m(\alpha)$. The first part of our proposal, then, is to revise the notion of quotation so that the quotation of any expression α is $r(\alpha)$. (In an environmentalist spirit, we might wish to recycle the left and right quotes as ' and °.) It is clear

⁵Exercise. Punctuate: Janacek wrote a quartet called the Kreutzer Sonata.

⁶Some years ago an article appeared in *The New Yorker* (I would dearly love the reference) lamenting the overuse of shudder quotes and proposing the introduction of different kinds of quotation marks to indicate different authorial attitudes: raised circles for (wide-eyed) surprise, etc. Gabriel Segal has observed that the carat or circumflex, found on computer keyboards, resembles an eyebrow: skeptics may take note.

that our new sort of quotation respects the requirements of sequentiality, finiteness, and universality.

To complete our proposal we have to specify a procedure for parsing an expression α containing (zero or more) q -marks. It is simply this: Scan α from left to right, looking for an occurrence of $^\circ$. If one is found, locate, by backtracking, the longest q -mark (occurrence) that ends with that occurrence of $^\circ$. Then continue scanning to the right, looking for another q -mark of the same length. If and when one is found, the two q -marks are mates.⁷ (The subexpression of α consisting of the two q -marks and the expression β lying between them, which may contain other (shorter) q -marks, may end in a string of occurrences of $'$, or may be null, may be taken to denote β , if desired.) Now apply the procedure to the subexpression of α lying to the right of the right q -mark and repeat in like manner until the end of α is reached. Of course there is no guarantee that α will in any sense be well formed; indeed, after a q -mark has been found, nothing at all ensures that scanning the remainder of α will turn up a mate for it.

We conjecture that the new kind of quotation will avoid the difficulties to which the old is subject, and will offer some evidence for this conjecture.

We shall need to observe that if we define $\langle \alpha, \beta \rangle$ as the expression $r(\alpha) * \beta$, then the ordered pair theorem holds for the operation $\alpha, \beta \mapsto \langle \alpha, \beta \rangle$.

Ordered Pair Theorem *If $\langle \alpha, \beta \rangle = \langle \gamma, \delta \rangle$ then $\alpha = \gamma$ and $\beta = \delta$.*

Proof. Suppose that $\langle \alpha, \beta \rangle = \langle \gamma, \delta \rangle$, i.e., that $\{j\}^\circ \alpha \{j\}^\circ \beta = \{k\}^\circ \gamma \{k\}^\circ \delta$, where $\{j\}^\circ$ and $\{k\}^\circ$ are the shortest q -marks that are not subexpressions of α and γ , respectively. Since the expressions $\{j\}^\circ \alpha \{j\}^\circ \beta$ and $\{k\}^\circ \gamma \{k\}^\circ \delta$ are identical, $j = k$ and thus $\alpha \{j\}^\circ \beta = \gamma \{j\}^\circ \delta$.

Now let $\zeta = \alpha \{j\}^\circ \beta$ and $\eta = \gamma \{j\}^\circ \delta$, $l = \text{lh}(\alpha)$, and $m = \text{lh}(\gamma)$. Observe that $\zeta(l+j) = ^\circ$, $\zeta(l+i) = '$ for all $i < j$, and $\eta(m+i) = '$ for all $i < j$. Suppose (for *reductio*) that $l < m$. (See Figure 28.) Then $l+j < m+j$, and $\zeta(l+j) \neq \eta(m+i)$ for all $i < j$. Since $\zeta = \eta$, $l+j < m$, and so $l+i < m$ for all $i < j$. Since $\zeta = \eta$, $\eta(l+j) = ^\circ$ and $\eta(l+i) = '$ for all $i < j$, and thus $\{j\}^\circ$ is a subexpression of γ . But since $j = k$, $\{k\}^\circ$ is a subexpression of γ , which contradicts the definition of k .

Similarly, not: $m < l$. Thus $\text{lh}(\alpha) = l = m = \text{lh}(\gamma)$, and since $\zeta = \eta$, $\alpha = \gamma$, $\{j\}^\circ \beta = \{j\}^\circ \delta$, and $\beta = \delta$.⁸ ■

⁷Our scanning procedure will thus correctly parse the expression $^\circ s'^\circ$, which might be taken to be an ill-formed quotation of s , as the (well-formed) quotation of the (ill-formed) expression s' .

⁸The validity of the ordered pair theorem depends upon the paired senses of the definition of ordered pair and of q -mark. Were we to reverse that of the latter, so that the q -marks became $^\circ$, $^\circ'$, $^\circ''$, etc., but leave the wording of the definitions of m and r , and ordered pair unchanged, the theorem would fail, as the expression $^\circ \circ \circ \circ \gamma$, γ arbitrary, shows.

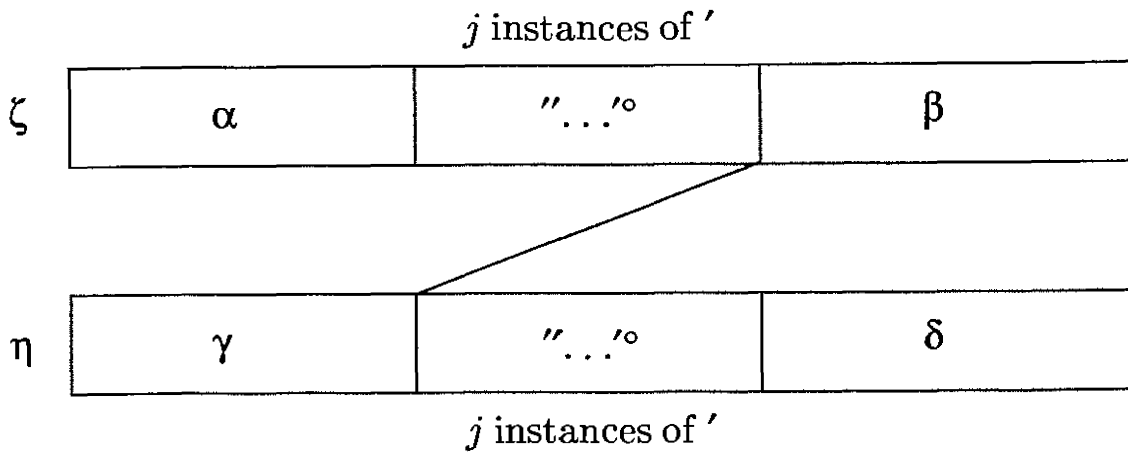


Figure 28.1: Since $\zeta = \eta$, if α is shorter than γ , then the displayed occurrence of $^{\circ}$ in ζ lies in γ , and $'' \dots ''^{\circ}$ is a subexpression of γ , impossible.

Corollary *If $r(\alpha) = r(\beta)$, then $\alpha = \beta$.*

Proof. If $r(\alpha) = r(\beta)$, then $\langle \alpha, \# \rangle = r(\alpha) * \# = r(\alpha) = r(\beta) = r(\beta) * \# = \langle \beta, \# \rangle$, whence by the ordered pair theorem, $\alpha = \beta$. ■

The ordered pair theorem does not hold for the obvious analogue of ordinary quotation: Let l and r be symbols, $q(\alpha) = [l] * \alpha * [r]$, and $(\alpha, \beta) = q(\alpha) * \beta$. Then, when a and b are any symbols, and $\alpha = a, \beta = lrlbr, \gamma = arl, \delta = lbr, (\alpha, \beta) = q(\alpha) * \beta = larlrlbr = q(\gamma) * \delta = (\gamma, \delta)$, but $\alpha \neq \gamma$ (and $\beta \neq \delta$). Even more simply, let $\alpha = l, \beta = r, \gamma = lr, \delta = \#$; then $q(\alpha) * \beta = q(\gamma) * \delta = llrr$.

We now show that ambiguities like the ones we have been noticing arise if a device like (ordinary) quotation is added in the natural way to logical languages of a familiar sort, but that they can be proved not to arise if we similarly adjoin our new sort of quotation.

The logical languages we have in mind are (first-order) languages written in Polish notation. There are no parentheses in these languages, each operator is written to the left of its operands, and every symbol has a degree: The degree of an n -place predicate or function sign is n , that of a variable or individual constant is 0, and that of the existential quantifier \exists , the arrow \rightarrow , and the equals sign $=$ is 2. The rules of term and formula composition are the expected ones, e.g.: If φ and ψ are formulas, so is $\rightarrow \varphi\psi$. Following Shoenfield,⁹ we call an expression that is either a term or a formula a *designator*.

Now, let L be such a language containing at least one function or predicate

⁹(Shoenfield, 1967), esp. pp. 14-16.

symbol f of degree 2. (= will do.) Let l and r be symbols not in L . (l need not be distinct from r .) The natural way to extend L by using l and r as the left and right quote is to say: Let the symbols of L_{bad} be those of L together with l and r ; where α is any expression at all of L_{bad} (including expressions containing l and r), the expression $[l] * \alpha * [r]$ is a term of L_{bad} ; the remaining rules of term and formula composition are as they are in L .

L_{bad} , so defined, is bad. It lacks unique readability, which requires that no designator of L_{bad} can be parsed in more than one way. More precisely, L_{bad} lacks unique readability if there is some designator σ such that for some symbol s of degree n , and two distinct n -tuples $(\sigma_1, \dots, \sigma_n)$ and (τ_1, \dots, τ_n) of designators, $\sigma = s\sigma_1 \dots \sigma_n = s\tau_1 \dots \tau_n$ (or there are different expressions α, β such that $[l] * \alpha * [r] = [l] * \beta * [r]$ —but that is obviously not the case). And there are many such designators in L_{bad} . For let a and b be any symbols of L_{bad} and consider the expression $flarlr lbr$, i.e., $[f] * [l] * [a] * [r] * [l] * [r] * [l] * [b] * [r]$. Since a is an expression (of L_{bad}) (recall that we identify a symbol with the expression of length one whose sole value is that symbol), lar is a term; since rlb is an expression, $lrlbr$ is also a term; and therefore since f is a symbol of degree 2, $flarlr lbr$ is a designator. But arl is also an expression and therefore $larlr$ is a term, and b is an expression and lbr is a term. Thus $flarlr lbr$ has the two distinct parsings:

$$f \ lar \ lrlbr$$

and:

$$f \ larlr \ lbr,$$

i.e., there are terms $\alpha, \beta, \gamma, \delta$, viz., $lar, lrlbr, larlr, lbr$, such that $f * \alpha * \beta = f * \gamma * \delta$ even though $\alpha \neq \gamma$ (and $\beta \neq \delta$); thus L_{bad} lacks unique readability.

If L contains the equals sign, interpreted as usual, and $[l] * \alpha * [r]$ is taken as denoting α in L_{bad} , then there are even expressions of L_{bad} true on one parsing and false on another, e.g. $= larlr larlr$.

It is unique readability, in some clear but as yet undefined sense of the term, that expression μ shows natural languages containing the usual sort of quotation to lack. Expressions κ and μ also show that we would violate universality if we were stipulate that every left quotation mark is to be mated by the nearest right quotation mark to its right.

Let us now show that if we extend L to a language L_{good} by similarly adding to L our new kind of quotation, then L_{good} enjoys unique readability.

Thus let $'$ and $^\circ$ be two symbols not in L . ($'$ and $^\circ$ do not have any degree.)

We now take the symbols of L_{good} to be those of L together with $'$ and $^\circ$. Where α is any expression at all of L_{good} (including expressions containing $'$ and $^\circ$), the expression $r(\alpha)$, defined as above, is a term of L_{good} ; the

remaining rules of term and formula composition are as they are in L . For L_{good} to enjoy unique readability, each designator of L_{good} must be a designator in exactly one way, that is, that if $r(\alpha) = r(\beta)$, then $\alpha = \beta$, and if s is a symbol of degree n , $\sigma_1, \dots, \sigma_n, \tau_1, \dots, \tau_n$ are designators, σ is a designator, and $\sigma = s\sigma_1 \dots \sigma_n = s\tau_1 \dots \tau_n$, then $\sigma_i = \tau_i$ for all $i, 1 \leq i \leq n$.

Theorem L_{good} enjoys unique readability.

Proof. Let σ be a designator. σ either begins with ' or ° or begins with a symbol of some degree. If the former, then $\sigma = r(\alpha)$ for some expression α ; but then if also $\sigma = r(\beta)$, $\alpha = \beta$, by the corollary to the ordered pair theorem. Thus we may assume that σ begins with a symbol s of degree n and that for some designators $\sigma_1, \dots, \sigma_n, \sigma = s\sigma_1 \dots \sigma_n$. To prove the theorem, it suffices to suppose that $\sigma = s\tau_1 \dots \tau_n$, with each τ_i a designator, and show that $\sigma_i = \tau_i$ for all $i, 1 \leq i \leq n$. Since $\sigma = s\sigma_1 \dots \sigma_n = s\tau_1 \dots \tau_n$, $\sigma_1 \dots \sigma_n$ (i.e., $\sigma_1 * \dots * \sigma_n$) = $\tau_1 \dots \tau_n$. We conclude the proof by showing that, more generally, if for all $i, 1 \leq i \leq m$, σ_i is a designator, for all $i, 1 \leq i \leq n$, τ_i is a designator, and $\sigma_1 \dots \sigma_m = \tau_1 \dots \tau_n$, then $m = n$ and for all $1 \leq i \leq m, \sigma_i = \tau_i$. We proceed by induction on the length of $\sigma_1 \dots \sigma_m$.

Case 1. σ_1 begins with ° or '. Then τ_1 begins with the same symbol and for some expressions α and γ , $\sigma_1 = r(\alpha)$ and $\tau_1 = r(\gamma)$. By the ordered pair theorem, $\sigma_1 = \tau_1$, and $\sigma_2 \dots \sigma_m = \tau_2 \dots \tau_n$. Since $\sigma_2 \dots \sigma_m$ is shorter than $\sigma_1 \dots \sigma_m$, by the induction hypothesis, $m - 1 = n - 1$, whence $m = n$, and for all $2 \leq i \leq m, \sigma_i = \tau_i$. Thus for all $1 \leq i \leq m, \sigma_i = \tau_i$.

Case 2. The first symbol s of σ_1 is of degree d . Then the first symbol of τ_1 is also s , and for some designators $\pi_1, \dots, \pi_d, \rho_1, \dots, \rho_d, \sigma_1 = s\pi_1 \dots \pi_d$ and $\tau_1 = s\rho_1 \dots \rho_d$. Thus $s\pi_1 \dots \pi_d\sigma_2 \dots \sigma_m = s\rho_1 \dots \rho_d\tau_2 \dots \tau_n$, and $\pi_1 \dots \pi_d\sigma_2 \dots \sigma_m = \rho_1 \dots \rho_d\tau_2 \dots \tau_n$. Since $\pi_1 \dots \pi_d\sigma_2 \dots \sigma_m$ is shorter than $\sigma_1 \dots \sigma_m$, by the induction hypothesis, $d + m - 1 = d + n - 1$, whence $m = n$, and for all $i, 1 \leq i \leq d, \pi_i = \rho_i$, and thus $\sigma_1 = s\pi_1 \dots \pi_d = s\rho_1 \dots \rho_d = \tau_1$, and for all $i, 2 \leq i \leq m, \sigma_i = \tau_i$. ■

Returning now to our own language, we want to examine what happens to the expressions involved in Ernst's paradox on our proposal. κ becomes:

b° appended to $^\circ a$,

λ remains:

ab ,

and μ becomes either:

$'^\circ b^\circ$ appended to $^\circ a'^\circ$

or:

${}^{\circ}b^{\circ}$ appended to ${}^{\circ}a^{\circ}$.

In the former case μ denotes κ ; in the latter, λ . In neither case is the denotation of μ ambiguous.

And what of analogues of ${}^{\circ}\text{app}('a', 'b', 'c')^{\circ}$? Well, according to our scanning procedure, $'{}^{\circ}\text{app}({}^{\circ}a^{\circ}, {}^{\circ}b^{\circ}, {}^{\circ}c^{\circ})'$ is ill-formed, as it contains three operands, instead of the required two.

$''{}^{\circ}\text{app}({}^{\circ}a^{\circ}, {}^{\circ}b^{\circ}, {}^{\circ}c^{\circ})''$

and

$''{}^{\circ}\text{app}({}^{\circ}a^{\circ}, {}^{\circ}b'^{\circ}, {}^{\circ}c^{\circ})''$,

however, are both well-formed, denoting the oddities $'{}^{\circ}b^{\circ}$, ${}^{\circ}ca'^{\circ}$ and $'{}^{\circ}ca^{\circ}$, ${}^{\circ}b'^{\circ}$, respectively.

It would be a pity if our new kind of quotation prevented us from imitating the calculation involving ${}^{\circ}$ 'appended to its own quotation' appended to its own quotation ${}^{\circ}$ which shows there to be an expression that denotes itself. But it does not. For, taking ${}^{\circ}$ quotation ${}^{\circ}$ to mean our new sort of quotation, we have that

$$\begin{aligned} & {}^{\circ}\text{appended to its own quotation}^{\circ} \text{ appended to its own quotation} \\ = & {}^{\circ}\text{appended to its own quotation}^{\circ} \text{ appended to the quotation of } {}^{\circ}\text{appended to its own quotation}^{\circ} \\ = & {}^{\circ}\text{appended to its own quotation}^{\circ} \text{ appended to } {}^{\circ}{}^{\circ}\text{appended to its own quotation}^{\circ}{}^{\circ} \\ = & {}^{\circ}{}^{\circ}\text{appended to its own quotation}^{\circ} \text{ appended to its own quotation}^{\circ}{}^{\circ}. \end{aligned}$$

As usual, the last expression both denotes and is identical with the first, and we have found an expression of the desired kind.

We conclude with the refrain:

According to W. Quine,
Whose views on quotation are fine,
 ${}^{\circ}\text{Boston}^{\circ}$ names Boston,
And $'{}^{\circ}{}^{\circ}\text{Boston}^{\circ}{}^{\circ}$ names ${}^{\circ}\text{Boston}^{\circ}$,
But 9 doesn't designate 9.