**Empty quotation**

**ROY SORENSEN**

In formal linguistics and computer science, the empty string plays a role akin to zero. Sometimes it is denoted by $\varepsilon$ or $\lambda$. In programming languages such as Perl an empty string is represented by the absence of text between two quotation marks (Wall and Schwartz, 1991).

Informally, we can subtract our way to the empty string:

1. ‘I’ has one character.
2. ‘ ’ has no characters.

Experts distinguish the empty string from a string in which all the characters have their values set to zero and from illegal strings or undefined strings: ‘The empty string differs from the null string in that the empty string is a defined string of no characters whereas the null string is an undefined string’ (Glass 1983: 410).

Sentence (3) is well-defined and true. Yet all theories of quotation have trouble modeling our ability to mention the empty string. Several theorists preclude the empty string as ungrammatical just as René Descartes precluded the vacuum as nonsense.

Under Donald Davidson’s (1979) demonstrative theory, quotation marks point to the enclosed material. If there is no demonstratum, sentence (3) is not true.

Under Gottlob Frege’s identity theory of quotation (much refined by Corey Washington (1992) and Paul Saka (2007: ch. 7)), the speaker uses the expression between the quotation marks to refer to itself. But in (3) there is no such expression. The speaker has nothing to refer with and nothing to refer to.

Some theorists count the quotation marks as part of the quotation. For instance, Alfred Tarski and W. V. Quine characterize quotation marks as the beginning and ending of a proper name (consisting of quotation marks).
marks plus the filling). According to this name model, (1)–(3) are each false because they undercount the number of characters.

The empty string differs from a string of blanks (Sorensen 1999). Although blanks are formed by omission, they are still characters. As in the previous sequence, the empty string can be reached by subtraction from strings of blanks:

(4) ‘ ’ has two blanks.
(5) ‘ ’ has one blank.
(6) ‘ ’ has no blanks.

Once again, all three sentences are true.

Mario Gomez-Torrente (2001: 130) complains that the demonstrative theory of quotation predicts that we can remove the enquoted expression just as we can remove a picture from its frame:

(7) ‘Socrates’ has eight letters.
(8) ‘ ’ has eight letters.

Paul Saka (2006: 456) agrees with Gomez-Torrente’s intuition that (8) is ungrammatical. Saka takes the further step of claiming that the identity theory of quotation has the advantage of predicting that (8) is ill-formed.

My attitude is the reverse of Saka’s. If the demonstrative theory predicts that (8) is grammatical and the identity theory predicts (8) is ungrammatical, then that pair of predictions would favour the demonstrative theory.

I take the intuition that (8) is ungrammatical to be partly due to (8) being obviously false. The bizarreness dissipates when we consider true statements using empty quotation such as (3) and (6). Each is a representation of an absence rather than an absence of representation.

There are common-sense objections to counting zero as a number. We surmount them by showing how well 0 can play the role of the number and by showing how entrenched this treatment is within mathematics.

The same method prevails against common-sense objections to counting the empty string as something that can be mentioned with quotation marks. Just as mathematicians routinely introduce zero when setting up a number system, formal linguists routinely introduce the empty string. Special comments are needed because the empty string labours in surprising ways. In particular, ‘ ’ is a substring of every string. This makes ‘ ’ an identity element with respect to concatenation. In other words, juxtaposing any string with ‘ ’ yields the string itself (Partee et al. 1990: 434).

Students rightly think twice before accepting oddities such as the empty set. But given that {} is indispensable to their best scientific theories, they should follow the example of W. V. Quine. Quine began as a nominalist. He accepted set theory because mathematics is indispensable to physics and set theory is indispensable to mathematics. Those willing to accept sets
sometimes baulk at the idea of a set that has an absence of members. But considerations of economy compelled Quine’s resignation to the singularities of the empty set.

The demonstrative theory of quotation can be revised to accommodate the empty string. The amendments should be guided by recent work on the perception of absences.

The demonstrative theory works most simply for perceptible demonstrata (as opposed to remembered demonstrata, deferred ostension, and reference to a demonstratum’s type). Happily for us modellers of empty quotation, there are many common-sense examples of perceiving absences. We hear pauses, feel holes, and see shadows. This makes it plausible that a student can see the empty string on a blackboard just as he can see a crack in a wall.

We perceive absences by virtue of how they contrast with their surroundings (Sorensen 2008). This explains why it is easier to see the empty string in the context of other sentences.

Davidson claims that quotation marks point to the shape of the enquoted inscription. Anomalously for him, the empty string lacks a shape. However, we have independent reason for believing that Davidson was being too narrow. As Manuel Garcia-Carpintero (1994) notes, sometimes we are interested in sound rather than shape. Witness the remarks ‘“Aristophanes” rhymes with “cacophonies”’ and ‘“Mellifluous” is mellifluous’.

On other occasions, we are interested in the word type rather than in the particular instance. The etymologist writes on the blackboard: ‘Neologism’ originated around 1800. What he wrote could be true even though the word token originated in 2007.

To reflect the generality of quotation, Garcia-Carpintero suggests that quotations are a species of ‘ostensive signs’:

Let us say that an ostensive sign consists of a demonstrative (‘this’), a category (‘sound’) – which may be merely implicit in the context – 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 ... The difference between an ostensive sign and an expression ... lies in that 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 explained to him. (1994: 258)

Given that absences can be perceptible tokens, Garcia-Carpentino’s account can be extended to empty quotation. We often use absences to refer to their types. A carpenter who wishes his assistant to bore ten
uniformly sized holes may bore the first hole as a sample: ‘Repeat that nine times!’

*Types* of absences are not absences. A remark such as

(9) ‘ ’ is a sequent in Gerhard Gentzen’s logical calculus.

does not concern the particular empty string in (9). Instead, (9) refers to a string type by pointing out a token of that type.

The phenomenon of deferred ostension provides further flexibility. W. V. Quine’s (1968: 40) illustration features a motorist who refers to his car’s gasoline by pointing at the fuel gauge. This would extend to absences. Nanook points to a tiny air hole in the snow and warns ‘That is a bear’. Here an absence is used to refer to a polar bear slumbering beneath the snow. Do not fall in! An absence can also be used to refer to another absence (as opposed to an absence type). A motorist can refer to his car’s lack of oil by pointing to the absence of oil on his dipstick.

Just as there are number systems that lack zero, there are formal languages that lack the empty string. When modelling some mathematical reasoning, we may prefer a system that lacks zero. The same holds when modelling a language.

Contemporary linguists treat natural languages as disguised formal languages. The task is to identify the underlying type of grammar. Given the expressive capacity of English (or any other natural language), linguists should postulate a grammar that includes the empty string. After all, linguists use English to discuss English. Therefore, they must have the capacity to mention the empty string. Since linguists are committed to empty quotation, we should prefer theories of quotation that do not preclude empty quotation as incoherent. Even greater favour should be bestowed on those theories that show how it is possible.¹

---

Dartmouth College
Hanover, NH 03755, USA
Roy.Sorensen@Dartmouth.edu

References


¹ Acknowledgements: Words cannot fully express my gratitude to Manuel Garcia-Carpintero, Ernest LePore, Paul Saka and Mario Gomez-Torrente.


