

Quoting Nothing

Luigi Pavone

Università di Verona, Italy

Abstract This paper argues that the practice of employing bare pairs of quotation marks to represent the empty string in formal linguistics and computer science is well-founded in the implicit conventions governing the use of quotation marks in natural language. In the framework of the Inscriptional Theory of Quotation (ITQ), it is argued that sentences containing empty quotation (i.e., empty quotational sentences) are grammatical and meaningful. Furthermore, the notion of an empty string is employed in the analysis of reported speech to provide a unified account that identifies mixed quotation as the primary form of reported speech.

Keywords Empty string. Empty quotation. Mixed quotation. Ostensive definition. Reported speech.

Summary 1 Introduction. – 2 Inscriptional Theory of Quotation (ITQ). – 2.1 The Semantics and Pragmatics of Empty Quotation. – 2.2 Grammaticality of Empty Quotation. – 3 Empty Quotation and Reported Speech. – 4 Conclusion.



Edizioni
Ca' Foscari

Peer review

Submitted 2024-07-25
Accepted 2024-11-08
Published 2024-12-11

Open access

© 2024 Pavone |  4.0



Citation Pavone, Luigi (2024). "Quoting Nothing". *JoLMA*, 5(2), 387-398.

DOI 10.30687/JoLma/2723-9640/2024/02/004

387

1 Introduction

The empty string is a fundamental concept in computer science and formal linguistics. It is defined as a string of length zero, meaning it contains no characters. In certain instances, it is represented by the symbol ‘ ϵ ’, that is, by a referring expression with a type-(e) semantic value. In most programming languages, including Python, Java, Swift, and others, the empty string is typically represented using quotation marks without any content in between. This is illustrated in (1) below. (1-a) is an example from Python language, wherein the empty string appears at the end of the line, to ensure that there is no extra content after the exclamation mark. In the follow-up (1-b), the empty string is used to represent missing data.

- 1-a. `print('Hello' + '!' + '')`
- 1-b. `email = ""`

Such examples illustrate limiting cases of *pure quotation* (or *metalinguistic quotation*) as opposed to *attributive quotation*. In contrast to attributive quotation, pure quotation is employed not to report someone’s speech, but rather to talk about linguistic expressions, e.g., to ascribe them syntactic, orthographic or semantic proprieties, as illustrated in (2a)-(2-c). Sentences (2-d) and (2-e) appear, at first glance, to be pure quotational sentences about zero-length linguistic expressions. In fact, if we assume that they are grammatically correct and meaningful, then they appear to be true and false, respectively.

- 2-a. ‘Ho’ is a Chinese proper name
- 2-b. ‘Ho’ has two characters
- 2-c. ‘Ho’ denotes Ho
- 2-d. “ has no characters
- 2-e. “ has eight letters

The use of empty quotation in programming languages is motivated by certain practical considerations. For example, the use of quotation marks with nothing between them maintains consistency with the representation of non-empty strings, as strings with content are typically enclosed in quotation marks. Extending this convention to the empty string produces a uniform syntax. The empty string can be defined with respect to a given string s as follows: “ = $s[i..j]$ with $i > j$, where $s[i..j]$ is a generic sub-string of s that starts at position i and ends at position j (Gusfield 1997, 3-4). For every string s , s concatenated with “ yields s : $s\text{“} = s$, $\text{“}s = s$ (Partee, Meulen, Wall 1990, 434). Consequently, the empty string is a sub-expression of every expression. Given that it plays the role of a neutral element within a linguistic system, it can be said that it exemplifies any linguistic

determinable (e.g., character length, syntactic category, etc.) to a degree of 0. As a result, the empty string can be considered compositionally neutral with respect to both the syntax and semantics of the expressions in which it occurs.

In certain contexts, it is convenient to consider the empty string as a singleton containing an empty tuple of characters: $\{\langle \rangle\}$. To illustrate, suppose we want to obtain the name 'HO' from 'H', 'O', and '' by concatenation. This can be achieved through operations on sets of indexed characters, as follows: $\{\langle 0, H \rangle\} \cup \{\langle 1, O \rangle\} \cup \{\langle 2, \langle \rangle \rangle\}$. In programming languages empty quotation is employed in various scenarios, such as handling user input, initializing variables, building strings, and so forth. As will be demonstrated in Section 3, it plays a role in the analysis of ordinary mixed quotation.

This paper addresses the question of whether the conventions governing the use of bare quotation marks in programming languages to represent the empty string are in alignment with the conventions that govern the use of quotation marks in natural language. In the framework of the Inscriptural Theory of Quotation (ITQ), which posits that quotations are linguistic expressions with type- $\langle e, t \rangle$ semantic values, it will be demonstrated that sentences containing empty quotations (i.e., empty quotational sentences), such as (2-d) and (2-e), are grammatical and meaningful. It will be argued that the use of bare quotation marks in programming languages is motivated by the implicit rules governing quotation marks in ordinary language.

The paper runs as follows. Section 2 outlines ITQ. In Section 2.1, ITQ is extended to encompass the phenomenon of empty quotation. Section 2.2 addresses the issue of the grammatical acceptability of empty quotational sentences. The argument is presented that sentences such as (2-d) and (2-e) are grammatical. The illusion of their non-grammaticality is explained by showing how, in certain contexts of use, empty quotation predicates may appear to be vacuous (in a sense that will be defined). In Section 3 the notion of empty string is employed to analyze ordinary language, by demonstrating how it enables a unified account of reported discourse, with direct and indirect speech identified as special cases of mixed quotation. Finally, Section 4 presents our conclusions.

2 Inscriptural Theory of Quotation (ITQ)

A number of attempts have been made in the philosophy of language and linguistics to formalize the nature and function of quotation. For an overview of these efforts, see Cappelen and Lepore (2007). The predominant focus has been on developing a unified theory of quotation that can encompass various quotation forms, including pure quotation, direct quotation, indirect quotation, hybrid quotation, and

mixed quotation. Within such a program, most theories concur that a quotation – what Recanati (2001, 649) terms *close quotation* (as opposed to *open quotation*) – is a referential expression of a certain kind, such as a proper name (Tarski [1933] 1956, 159), a definite description (Geach 1957, 82), a demonstrative (Davidson 1979; Cappelen, Lepore 1997), a function (Richard 1986). This perspective can be described as *Referentialism about Quotation* (RQ), which posits that a quotation is a singular term (or serves as a singular term).

In contrast to RQ, ITQ suggests treating quotations as complex predicates that *ostensively* describe classes of inscriptions/utterances (cf. Pavone 2024). Such a predicativist view on quotation is informed by the work of Goodman (1951, 262) and Scheffler (1954). According to ITQ, a quotation as a whole – i.e., a quotation-mark expression consisting of both quotes and what is enclosed between them – describes a class of objects in a replication relation to each other, where replication is understood as a relation of similarity in relevant linguistic features associated with a set of linguistic determinables, $k = \{d_1, \dots, d_n\}$, which is contextually/pragmatically provided.

To illustrate, consider the example provided in (3) below. In certain contexts of use, it can be reasonably assumed that (3-b) is a replica of (3-a), but this is not necessarily the case when the intended linguistic similarity is taken with respect to a different set of relevant linguistic determinables.

- 3-a. CAT
3-b. cat
3-c. CAT

Let us suppose that k includes length in characters, lexical category, font, and no other element. In this case, (3-b) would be a replica of (3-a), as it has the same length and the same font as (3-a), and both are common countable nouns. In contrast, (3-c) is not a replica of (3-a), as it does not replicate the same font as (3-a). However, by assuming an alternative set of relevant features, expanded to include the uppercase/lowercase distinction, and reduced to remove the font, (3-c) counts as a replica of (3-a), while (3-b) does not. It can be said that the conventions governing quotation in natural languages provide quotation-marks expressions with a character, in the customary Kaplanian sense (1989), that is, a rule for determining the conditions of applicability of a quotation predicate in various circumstances of evaluation and contexts of use. The conventional linguistic meaning of a quotational sentence is to be supplemented through a pragmatic process of identifying the relevant notion of replication.

The quotation-mark expression occurring in (2-b), for instance, is to be regarded as a complex predicate ostensively describing the following class of objects: $\{x: x \approx_k \setminus \text{Ho} \checkmark\}$. The symbol ' \approx_k ' represents a

replication relation between inscriptions/utterances defined with respect to a certain set k of contextually provided linguistic determinables. The small arrows represent Reichenbach's token-quotation marks (1947, 284). We can define them as a pointing device that the quoter employs to provide the applicability conditions for the quotational predicate by demonstrating a quotational exemplar, i.e., that particular object printed between token-quotation marks that has both perceptual and non-perceptual properties. As linguistically incorporated, they introduce a covert demonstrative pronoun into the language. Consequently, token-quotation marks play a dual role, acting both as a demonstrative and as a means of demonstration. The extension of the quotational predicate is to be construed as follows: the class of inscriptions/utterances that are similar in linguistic form (with respect to k) to the quotational exemplar.

Accordingly, (4-a) below is to be analyzed as (4-b), which asserts that all k -replicas of the quotational exemplar can be used as a verb. However, the truth-conditional value of (4-b) may vary depending on context of use. To illustrate, if the set of the relevant linguistic features in a context c includes the lexical category to which the quotational exemplar belongs (i.e., proper name), $[(4-b)]_c = 0$, as no k -replica that has the same lexical category as the token-quoted item can be used as a verb (no proper name can be used as a verb.). In contrast, $[(4-b)]_{c^*} = 1$, in a context c^* providing a set of purely orthographic determinables. In fact, 'ho' is a verb in the Italian language.

- 4-a. 'Ho' can be used as a verb
 4-b. $\forall x(x \approx_k \searrow \text{Ho} \swarrow \rightarrow x \text{ can be used as a verb})$

To formulate the rules by which a linguistic similarity is specified, the Kaplanian notion of semantic context of use is supplemented by a set k of linguistic determinables. The extension of a quotational predicate is defined as the class of linguistic objects that have (or are) the same determinables (as specified in the context) as the quotational exemplar. A limiting case arises when the process of specifying the notion of replication at stake leads the addressee to identifying the extension of the quotation predicate with the singleton containing the quotational exemplar itself. For instance, there will be contexts of use in which (5-a) below, whose inscriptional analysis is (5-b), is not true.

- 5-a. 'Ho' = 'Ho'
 5-b. $\forall x(x \approx_k \searrow \text{Ho} \swarrow \leftrightarrow x \approx_k \searrow \text{Ho} \swarrow)$

Intuitively, the quoter suggests a range within which the search for the recipient should be conducted, to identify the kind of replication involved for the domain of applicability of the quotation predicate.

To assist in the comprehension of a quotation, the quoter may also provide a range of exemplars, illustrating both positive and negative applications of the quotation predicate (cf. Pavone 2023, 242). We define appropriate contexts of use as those that provide linguistic determinables that are exemplified by the quotational exemplar. It would be inappropriate, for instance, to attempt to define a class of objects that have the same color as an exhibited colorless object. Let us call this appropriateness constraint (AC) on contexts of use for quotation-marks expressions. This may be seen as an instance of Grice’s maxim of relation (1989, 28).

2.1 The Semantics and Pragmatics of Empty Quotation

The theoretical framework previously outlined provides a basis for the semantics and pragmatics of empty quotation. Quotational sentences (2-d)-(2-e), repeated below for convenience, illustrate instances of empty quotation.

- 2-d. “ has no characters
- 2-e. “ has eight letters

In contrast to Gomez-Torrente (2010), who thinks that our pre-existing understanding of quotation does not permit the desired interpretation of (2-d), in terms of a true sentence about the empty string, we maintain that the conventions that elucidate non-empty quotations in natural languages are the same as those that elucidate empty quotation in programming languages.

In accordance with the instructions set forth by ITQ, (2-d) is to be logically rendered as (2-d’), where the quotation predicate has this extension: $\{x: x \approx_k \setminus \checkmark\}$, for some k -replication. This is the class of all k -replicas of the token-quoted item. (2-d’) asserts that all k -replicas of the token-quoted item have no characters. Similarly, command (1-a) can be interpreted as (1-a’), which instructs us to print one replica for each token-quoted item.

- 1-a. `print('Hello' + '!' + '')`
- 1-a'. $\exists x \dots z (x \approx_k \setminus \checkmark \wedge \dots \wedge z \approx_k \setminus \checkmark)$. `PRINT x ... z` (in this order)
- 2-d'. $\forall x (x \approx_k \setminus \checkmark \rightarrow x \text{ has no characters})$

Bare quotation marks are predicates lacking appropriate contexts of use. The use of bare quotation marks in sentences such as (2-d) and (2-e) violates AC, as the minimal blank between quotation marks fails to exemplify linguistic features. However, violation of AC may trigger a pragmatic calculation for a conversational implicature.

Let us consider the following scenario. It is known that John is a person who lacks scruples, and the speaker wishes to assert that Tom is similarly corrupt, possibly without explicitly committing to that assertion. The speaker might say that Tom has the same scruples as John. This assertion is evidently false, given that John has no scruples. However, the speaker conveys the content that Tom is a person without scruples. Communication seems here to violate the Gricean maxim of relation “for the purpose of getting in a conversational implicature by means of something of the nature of a figure speech” (Grice 1989, 33).

A similar phenomenon we propose occurs with the interpretation of empty quotations, such as in (2-d) and (2-e). Since no linguistic object is exhibited, the quotational predicate extension is literally empty: there is no object that has the same linguistic form as the token-quoted item, which does not exemplify linguistic features. Consequently, sentences (2-d) and (2-e) are trivially true. However, the violation of AC can result in a pragmatic calculation that leads to a conversational implicature.

The desired class of objects, that is, the singleton containing the empty string, can be obtained through a pragmatic process that can be described as follows. The violation of AC prompts the addressee to interpret the token-quoted item as a proxy for what it metonymically represents (by a form of deferred ostension). The token-quoted item (the minimal blank between quotes) can be seen as the output generated by the command to print the empty string. This causal relationship allows for an interpretive shift from effect to cause, that is, from the token-quoted element to the empty string. For this interpretation to take place, it is necessary to ensure that the blank between quotation marks is of a sufficient length, such that it does not exceed the blank that typically separates characters within a word. Therefore, the extension of the quotation predicate is to be construed as the class of replicas of the proxy token-quoted item, that is, the class of replicas of the empty string. Given that the empty string has only one replica, namely the empty string itself, for any linguistic determinable that it exemplifies at a degree of 0, the quotation extension is identical to the empty string singleton.

2.2 Grammaticality of Empty Quotation

Some scholars (e.g. Sorensen 2008) maintain that sentences such as (2-d)-(2-e) are grammatical (and meaningful), while others (e.g., Gomez-Torrente 2001; Saka 2006; 2011) argue that they are not. However, our intuitions are mixed. Those who argue for the grammaticality thesis and those who argue for the ungrammaticality thesis are both committed to explaining respectively the apparent ill-formedness and well-formedness of (2-d)-(2-e). ITQ appears to be able to do this job better than its competitors.

According to Sorensen (2008, 58), (2-e) can appear to be ill-formed due to its obvious falsity, but the judgment of ungrammaticality, he argues, is nullified when true empty quotational sentences, like (2-d), are considered. This explanation seems to be inadequate for two reasons. The assertion that Hitler won World War II or that a circle has four sides is obviously false, yet this does not trigger ungrammaticality judgments. Furthermore, those who reject the grammaticality of empty quotations tend to reject both (2-d) and (2-e). Sorensen should also provide an explanation for the illusion of ill-formedness of (2-d).

Saka (2011, 206) argues that sentences such as (2-d) may appear to be well-formed due to their capacity to convey/communicate a truth, in virtue of a context-induced reading, given that communication of truths does not necessarily require well-formedness. To explain the disagreements on the grammaticality of empty quotational sentences, the author invokes what he calls the speech-only thesis, which posits that language is only speech, and writing is not strictly a language. Consequently, from a linguistic/naturalistic point of view on language, empty quotational sentences appear to be ungrammatical because they cannot be pronounced in such a way to produce a corresponding utterance in a natural spoken language (2011, 216). In contrast, from a logical/philosophical point of view on language, in which any syntactic/semantic system counts as a language, empty quotation can appear to be grammatical.

Saka's approach to mixed intuitions is based on premises regarding the nature of language and disciplinary differences between linguistics and philosophy of language that are highly controversial. An alternative explanation of mixed intuitions that is less committed to general assertions about the nature of language and methodological issues in the language sciences would be preferable. This is what ITQ seems to provide.

The underlying concept of ITQ is that quotations are ostensive-ly defined predicates. This implies that quoting requires an associated ostensive act. In defining a quotational predicate, the quoter employs token-quotation marks as a pointing device to describe the class of replicas (in some relevant set of linguistic determinables k) of the quotational exemplar. When a bare pair of quotation marks occurs, the addressee of the quotation may interpret the associated ostensive act, performed by the token-quotation marks, as lacking a *demonstratum*. Under this interpretation, the quotational predicate appears to be vacuous - from which a judgment of ungrammaticality. Alternatively, the associated ostensive act is interpreted as genuinely pointing to something that helps the addressee to find the correct condition of applicability for the quotational predicate. The possibility of these two readings is a candidate for explaining our mixed intuitions about the grammaticality of empty quotations.

3 Empty Quotation and Reported Speech

The concept of empty string is not limited to programming languages. It can also play an important role in the analysis of ordinary language. It is widely accepted that mixed quotation, illustrated by (6-a) below, is a hybrid case of reporting speech that exhibits both direct and indirect verbal forms. However, through the concept of empty string, mixed quotation can be analyzed as a basic reporting verbal form, of which the direct and indirect forms arise as limiting cases. Bare quotation marks can play a role in reported speech analysis similar to their role in programming languages as useful placeholders or to express default values in situations where a string is expected but may be absent or unspecified.

The inscriptional analysis of (6-a) is (6-b). The individual constant 'a' denotes Paul, '≡' is a symbol for the paraphrase relationship between inscriptions/utterances, the capital 'Y' represents the set of all sub-strings of the corresponding inscription y, 'I(z, y)' is a 2-place predicate expressing that z inscribes/utters y.

6-a. Paul said that proper names “are not words in a language”

6-b. $\exists x, y (x \approx_k \text{are not words} \dots \checkmark \wedge x \in Y \wedge y \equiv P \wedge Iay)$

As a whole, (6-b) asserts that there is a replica (with respect to a set of relevant linguistic determinables k) x of the token-quoted items such that x is a sub-string of a paraphrase y of the inscription P , and Paul inscribed y . P is defined as the value of a binary function Φ which takes the unquoted (U) and the quoted (Q) parts of the complement clause of the mixed quotation as its arguments. Φ works as follows: $\Phi(U, Q) = U \frown S$, that is, U concatenated with S , where S is construed as a paraphrase (in the context of the reporter) of x , and x is what the reportee (Paul) inscribed. In other words, $S \in [x]_r$, where $[x]_r$ is defined as the set of all expressions that in the context of the reporter have the same semantic value as x . Conventionally, S can be assumed to be the shortest expression in $[x]_r$.

In this framework, direct and indirect speech arise as limiting cases of mixed quotation when the function Φ takes empty string as one of its two arguments. A covert empty quotation is postulated in indirect speech, illustrated by (7-a), whose inscriptional analysis is (7-b). In (7-a), the complement sentence contains no quoted part. Hence $\Phi(U, \text{'}) = U \frown S$, where $S \in [x]_r$, as above. Given that x is a replica of an empty string, x is devoid of any semantic value, or alternatively, it is assigned a null semantic value. Accordingly, $[x]_r$ is the set of all expressions that in the context of the reporter have a null semantic value. The shortest expression in $[x]_r$ is a string with a length of zero. Therefore, $U \frown S = U \frown \text{'}$, and $P = U$. (7-b) asserts that Paul inscribed an inscription y such that the empty string x is trivially a sub-string

of y and y is a paraphrase of P . In a similar manner, direct speech can be analyzed as a case in which $\Phi(", Q) = "\sim S$.

7-a. Paul said [that] " proper names are not words in a language

7-b. $\exists x, y (x \approx_k \setminus \swarrow \wedge x \in Y \wedge y \equiv P \wedge Iay)$

4 Conclusion

Some scholars, such as Gomez-Torrent and Saka, argued that empty quotations are not grammatical. This perspective contrasts with the conventional practice in formal linguistics and programming languages of employing empty quotes to represent the empty string. In response to this phenomenon, proponents of the non-grammaticality thesis may invoke the homonymity thesis, postulating that the quotation marks employed in programming languages do not belong to the same type as those used in natural languages. In contrast, this paper has argued that the practice of using bare quotation marks in formal linguistics and programming languages has its roots in the conventions governing the use of quotation in natural languages. According to ITQ, quotation-marks expressions are predicates ostensibly defined by means of the exhibition of a quotational exemplar. The ostensive definition at stake is based on the notion of similarity in linguistic form defined with respect to a set of contextually provided linguistic determinables. The appropriateness constraint, what has been called AC, is violated where a bare pair of quotation marks occurs, as no linguistic item is here token-quoted. The violation of AC prompts the addressee to interpret the minimal blank between the token-quotation marks as a proxy token for the empty string that it metonymically represents. Such a conversational implicature appears to be a well-established and an accepted convention within the domain of programming languages. The notion of empty string is not limited to programming languages. It plays a role in the analysis of ordinary mixed quotation, which has been proposed as the primary reporting form. Direct and indirect reported speech can be regarded as special forms of this.

Acknowledgements

I would like to express my gratitude to Emiliano Boccardi and Filippo Costantini for their helpful feedback on an earlier draft of this paper. The paper also benefited from constructive comments from two anonymous referees of this journal.

Bibliography

- Cappelen, H.; Lepore, E. (1997). "Varieties of Quotation". *Mind*, 106(423), 429-50. <https://doi.org/10.1093/mind/106.423.429>.
- Cappelen, H.; Lepore, E. (2007). *Language Turned on Itself*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199231195.001.0001>.
- Davidson, D. (1979). "Quotation". *Theory and Decision*, 11, 27-40. <https://doi.org/10.1007/bf00126690>.
- Geach, P. (1957). *Mental Acts*. London: Routledge.
- Goodman, N. (1951). *The Structure of Appearance*. 3rd edition. Boston: Reidel.
- Gomez-Torrente, M. (2001). "Quotation Revisited". *Philosophical Studies*, 102, 123-53. <https://doi.org/10.1023/a:1004664020280>.
- Gomez-Torrente, M. (2010). "On Quoting the Empty Expression". *Philosophical Studies*, 148, 439-43. <https://doi.org/10.1007/s11098-009-9336-3>.
- Grice, P. (1989). *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
- Gusfield, D. (1997). *Algorithms On Strings, Trees, And Sequences: Computer Science and Computational Biology*. Cambridge: Cambridge University Press.
- Kaplan, D. (1989). "Demonstratives". Almog, J.; Perry, J.; Wettstein, H. (eds) (1989). *Themes from Kaplan*. Oxford: Oxford University Press, 481-563.
- Partee, B.; Meulen, A.; Wall, R. (1990). *Mathematical Methods in Linguistics*. Dordrecht: Kluwer Academic.
- Pavone, L. (2023). "The Semantics and Pragmatics of Impure Direct/Mixed Quotation". *Intercultural Pragmatics*, 20(3), 239-50. <https://doi.org/10.1515/ip-2023-3002>.
- Pavone, L. (2024). "An Inscriptorial Account for Mixed Quotation". Capone, A.; Perconti, P.; Graci, R. (eds), *Philosophy, Cognition and Pragmatics*. Cham: Springer, 189-99. https://doi.org/10.1007/978-3-031-50109-8_11.
- Recanati, F. (2001). "Open Quotation". *Mind*, 110, 637-87. <https://doi.org/10.1093/mind/110.439.637>.
- Reichenbach, H. (1947). *Elements of Symbolic Logic*. New York: Macmillan.
- Richard, M. (1986). "Quotation, Grammar, and Opacity". *Linguistics and Philosophy*, 9, 383-403.
- Saka, P. (2006). "The Demonstrative and Identity Theories of Quotation". *The Journal of Philosophy*, 103(9), 452-71. <https://doi.org/10.5840/jphil2006103920>.
- Saka, P. (2011). "Quotation and Conceptions of Language". *Dialectica*, 65(2), 205-20. <https://doi.org/10.1111/j.1746-8361.2011.01266.x>.
- Scheffler, I. (1954). "An Inscriptorial Approach to Indirect Quotation". *Analysis*, 14, 83-9.
- Sorensen, R. (2008). "Empty Quotation". *Analysis*, 68, 57-61. <https://doi.org/10.1093/analysis/68.1.57>.
- Tarski, A. [1933] (1956). "The Concept of Truth in Formalized Languages". Tarski, A., *Logic, Semantics, Metamathematics*. Oxford: Oxford University Press, 152-78.

