

Treball/projecte de fi de màster de recerca

An LFG approach to the syntax of quotation

Eugenio Martin Vigo

Màster: Màster Universitari en Lingüística i Teòrica Aplicada

Edició: 2011-2012

Director/ora: Àlex Alsina

Any de defensa: 2012

Col·lecció: Treballs i projectes de fi de màster de recerca

Programa oficial de postgrau

“Comunicació lingüística i mediació multilingüe”

UNIVERSITAT POMPEU FABRA

An LFG approach to the syntax of quotation

Thesis for the achievement of the degree of MA in
Theoretical and Applied Linguistics

Author:

EUGENIO M. VIGO

Supervisor:

PROF. DR. ALEX ALSINA

September 2012

Acknowledgments

This work would have been impossible without the invaluable help of so many people that have been “just there” during my undergraduate courses (in which I started reflecting about this topic), during my MA courses and also during the time I have been writing this work.

Names are many, lines are scarce to name them all. Thanks to my family for supporting me from the not-so-big distance separating Pamplona and Barcelona. Thanks to Prof. Dr. Alex Alsina for accepting to supervise this work and for the insightful and interesting discussions we had. Thanks to Prof. Dr. Manuel Casado, who ignited my interest in quotation almost three years ago. Thanks to Juan Antonio and Miquel for the great times at the university. Thanks to my classmates for the funny multilingual-multicultural mess this year has been. Thanks to the *Department de Traducció i Ciències del Llenguatge*, including our helpful and awesomely effective secretaries, for the great opportunities you have given me during this intense course.

Finally and of course, most importantly, many many thanks to you, Silvia. Without you, I would have surrendered long ago.

Contents

1	Introduction	7
2	Defining quotation	11
2.1	Previous definitions of quotation	11
2.1.1	The lexicalist view	11
2.1.2	The compositional view	13
2.2	A syntactic definition of quotation	15
2.2.1	Grammatical opacity and quotation	16
2.2.2	Testing our hypotheses	18
2.2.3	Citation vs. quotation	21
2.3	Conclusions	24
3	The syntax of quotation	25
3.1	Previous theories	25
3.1.1	Work based on MP	25
3.1.2	Bonami and Godard’s (2008) work	27
3.2	A proposal based on LFG	29
3.2.1	Quoted f-structures as PRED values	30
3.2.2	C-structure/f-structure mapping	32
3.2.3	“Lexical” features of quotes	37
3.2.4	Summary	38
4	Conclusions	41

Chapter 1

Introduction

Quotation is omnipresent. It is in the news, in literature, in scientific papers, in our daily speech. There it is to help us to refer to what someone else has said: *John F. Kennedy said: “Ich bin ein Berliner”*. There it is to tell children or the learner of some language how to say something correctly: *No, don’t say “he do”; “he does” is correct*. Its uses are manifold and nobody is unaware of its existence, but the question is: what do we know about this so helpful linguistic device?

We know quite a bit of its semantic properties. The tradition starting with the seminal paper by Tarski (1933/1983) led to a huge amount of work on the semantics of quotation, of which one can cite major milestones like the ones by Quine (1981) or Davidson (1979); more recent work includes the book by Cappelen and Lepore (2007). Thanks to them and others, we know how quotation creates a special kind of semantic interpretation, namely the fact that something quoted refers to the linguistic expression itself, thus leading to special properties unquoted segments do not have (e.g., in some sense, quotes are *opaque*; Quine, 1981).

However, we know quite little of the formal syntax of quotation and this is what this thesis is about. There is previous work on the topic, but I consider it quite unsatisfactory, as it is my intention to show here. The main issue to resolve is the following: why is it possible to insert an arbitrary linguistic segment in another one via quotation, even if the quoted segment is in another language or it is an ungrammatical utterance? Take for example the following cases:

- (1) Silvia said: “Alba is dead”
- (2) “abu baba him-baba babá” is a sentence in Linguisticspaperese
- (3) “it is possible how?” is bad English

The problem is not trivial at all. Examples (1)–(3) show sentences where some part of an utterance seems to be completely free of any grammatical restriction, but not affecting the grammaticality of the whole, unlike the usual case in which for a sentence to be perfectly grammatical, all of its parts must be so too. A basic principle of any formal generative¹ grammar states that syntax is compositional, therefore the status of the parts of a sentence is expected to affect the status of the whole, save for the case of quotation. This ought to be explained.

Moreover, there is another problem, related to the previous one. In any quotational sentence, we face the uncommon fact of having *two* grammaticality judgments: the one for the quote and the one for the whole, but only the latter being binding. For example, in (3), we know that the quoted segment is ungrammatical because of our knowledge of English. This ungrammaticality, however, does not render the whole sentence ungrammatical; we know that (3) is in fact grammatical! More interesting is the case of (2), in which we do not know, and possibly it is not even relevant, whether the quote is grammatical or not. What is indeed relevant to us is that the whole (2) is grammatical.

One might believe that these are issues more related to philosophy of language than syntax; I do not deny its philosophical importance, namely what is ungrammaticality really about and whether it is possible to have two judgments for the very same expression. But quotation is also a phenomenon that is intrinsically syntactic! For example, quotation triggers effects in agreement that in fact show that there is a quote in a sentence. Imagine you *listen* (not read) two utterances like these:

- (4) the dogs is a noun phrase
- (5) the dogs are a noun phrase

Even though *the dogs* is not enclosed by quotation marks in (4), the sentence is interpreted as ‘the expression *the dogs* is a noun phrase’. On the other hand, (5), even though it is a completely false (yet grammatical) sentence, means that ‘some animals we consider to be dogs are a noun phrase’. The fact that there is no agreement between the subject and the verb in (4) is effectively the key difference between that example and (5). For some reason to be explained, quotes behave like singular NPs, the quotational interpretation being impossible in the presence of a plural verb.

¹I consider that the term *generative grammar* encompasses any theoretical framework that considers that the faculty of language is based on some computational syntactic component, no matter whether derivational (e.g. Chomskyan generative grammar) or non-derivational (e.g. LFG and HPSG, mainly).

However, a sentence like (6) is ambiguous:

(6) the dogs and the cats are noun phrases

Sentence (6) can be interpreted in at least two ways:² either as the absurd ‘some animals that we consider to be dogs and some others that we consider cats are noun phrases’ or ‘the expression *the dogs* and the expression *the cats* are noun phrases’.

A proper theory of the syntax of quotation must be able to answer these puzzles, namely it must answer the following questions, which are indeed critical to this endeavor:

1. Why is it possible to quote any linguistic expression, be it grammatical or not, in the same or a different language, without affecting the containing structure’s grammaticality?
2. How is that possible?
3. How do we account for the ambiguous structures and what do they tell us about quotation?

This thesis for the achievement of the degree of Master of Arts in Theoretical and Applied Linguistics aims to answer those questions as far as my ability to do so allows me. For this, first I introduce a brief account of what is quotation, a topic that leads us into issues more akin to philosophy, logic and semantics, but that appear to me to be of vital importance for delimiting the object of this research. Following that, the syntactic theory itself is presented. This theory, for reasons that I hope will be made clearer in the course of this work, is based in Dalrymple’s (2001) version of the Lexical-Functional Grammar (LFG) generative framework.

Finally, a methodological note might be worth spelling out before entering into the discussion of these topics. Throughout this work, I will use *italics* for referring to unquoted expressions, technical terms or emphasis. An expression in “*double quotation marks and italics*” is used for referring to quoted expressions. “Double quotation marks” are used for scare-quoting or for citing passages from other works, while semantic paraphrases are shown in ‘single quotation marks’.

²I consider the other possibilities to be quite unlikely, namely those in which one of the elements is quoted and the other one is not.

Chapter 2

Defining quotation

In order to understand the syntax of quotation, there seems to be a prior, self-evident and critical step to take, namely, to discuss what quotation is and what it is not. If we do not take our object apart and really understand what we are talking about, we are bound to fail in our task. A cautious and precise definition of our object of study is required before attempting to describe the formal combinatorial properties of said object, i.e. its syntax.

2.1 Previous definitions of quotation

It seems reasonable to first have a brief account of previous work made on the question of what is quotation before defending my own definition of it. In general, there have been two main ways to understand quotation: there is a position we may call *lexicalist* and another one we may call *compositional*. The former tries to explain quotation from a lexical perspective; the latter, on the other hand, from a formal semantic perspective. However, both of them are rooted in definitions of quotation that are semantic by nature.

2.1.1 The lexicalist view

The earliest work on quotation was actually not by linguists, but logicians like Tarski (1933/1983) or Quine (1981) who studied quotation as a special case in which, somehow, meaning appears to be suspended.³ The problem they had to solve was essentially the one illustrated by the following sentences (adapted from Quine, 1981, p. 26):

³In particular, Tarski (1933/1983) will take advantage of this property of quotation for his proposed solution of the *Liar's paradox*. Of course, such a topic, even though fascinating, is outside the scope of this research.

- (7) “Cicero” has six letters
- (8) “Marcus Tullius” has six letters

While sentence (7) is true, (8) is untrue, even though Cicero and Marcus Tullius are both the same person. The challenge here is to explain how come replacing the quote of an expression S by the quote of a synonymous expression S' does not guarantee that the truth-value is preserved.

The approach taken by these authors is to consider quotes to be proper names of expressions that refer to the expression contained in their interior (Tarski, 1933/1983, p. 156; Quine, 1981, p. 23). This is reasonable given the fact that these quotes appear in positions that are reserved to NPs. According to this, the reference of “Cicero” is the proper name *Cicero*, while the reference of “Marcus Tullius” is the proper name *Marcus Tullius*, therefore replacing one by the other does not guarantee the truth-value to be kept, because the reference we are taking is actually different in both cases. However, even though these authors defend that quotes are proper names, both Tarski (1933/1983, p. 160) and Quine (1981, p. 26) appear to be aware that quotes do have certain internal structure, namely that they are composed by quotation marks and the expression quoted in their interior.

This point of view was adapted by Rey-Debove (1997) in her linguistic account of quotation or the phenomenon she identifies as *autonymy* in general (French *autonymie*, i.e. the property of a sign to refer to itself). According to her, a quote consists in a linguistic expression whose *meaning* refers to another expression. For example, if the meaning of *dog* is ‘domestic mammal member of the Canidae family’, then the meaning of “dog” is ‘the expression *dog*’. Notice though the subtle but critical difference between this proposal and the one implied by Tarski (1933/1983) or Quine (1981): by no means do they imply that the quoted expression is the meaning of the quote, but instead its logical reference. Both logicians use the term *denotation* to refer to the relationship between a quote and the sign it quotes and nowhere do they consider the quote to be an atomic expression. On the contrary, as I already mentioned above, both logicians consider that there is certain (even though simple) structure to quotes. However, Rey-Debove (1997) or others like De Brabanter (2005) take the quote to be a sign or expression in itself, akin to the concept of linguistic sign as stated by De Saussure (1916/1984), i.e. a phonetic form with an atomic conceptual meaning that is considered to be knowledge shared by all speakers of some language.

Despite the differences, it appears to me that a common trait underlies these authors’ theses. The effort of stating quotes as some sort of special expression, similar to a lexical proper name comes from the effort of explaining the nature and properties of a related yet different phenomenon:

metalanguage. Metalanguage is the way to describe linguistic entities of an object-language (Carnap, 1956), and it is self-evident that quotation is a device that is useful for such purpose, as the previous examples (7) and (8) illustrate. When describing not only a linguistic entity, but any object, we presuppose that that said object is totally or partially unknown to the addressee, that being exactly the reason why we are describing such an object in the first place. The peculiarity of this in language is that we use languages to describe other languages and sometimes, as in the aforementioned examples, it is the same language which acts both as metalanguage and object-language, English in this case. In order to explain why quotation allows us to describe parts of some language, it is reasonable to establish that quoting some expression consists in somehow “mentioning” but not actually “using” it (Quine, 1981), while still using some expression that resembles it and also syntactically behaves like an NP. In consequence, this motivates the choice of considering quotes to be atomic values that encompass in one singular term an object that is going to be described.

The immediate consequence of these postulates is that, if we take this approach in our research on the syntax of quotation, we are bound to state that quotes are lexical items that are inserted exactly like any noun or verb. In fact, Rey-Debove (1997) proposes the existence of some sort of metalexicon (French *métalexique*) and some research on delocutives also shares this view (De Brabanter, 2005; Casado Velarde, 2010). However, as I will show, this thesis is bound to fail for logical (§2.1.2) and also syntactic reasons (§2.2).

2.1.2 The compositional view

The other view is the one I call *compositional*, as it places quotation at an interpretational level. Examples of this approach are the ones by Davidson (1979) and Cappelen and Lepore (2007), but it is also implied in most of the syntactic work around quotation, as we will discuss later in this work.

For Davidson (1979), quotation is just deixis to a certain expression, such that the meaning of (9a) is (9b):

- (9) a. “dog” is a noun
- b. ‘Dog. The expression of which this is a token is a noun’

Davidson (1979) takes this route to explain quotation for several reasons. For instance, he is fully aware of the shortcomings of the lexicalist view of quotation. In the first place, there is the problem of defining what it really means to mention an expression without using it: in some sense, quoting an expression is a way to use that expression. However, his main point is the following:

But there is a further, and I think, decisive objection, which is that on this theory we cannot give a satisfactory account of the conditions under which an arbitrary sentence containing a quotation is true. In an adequate theory, every sentence is construed as owing its truth or falsity to how it is built from a finite stock of parts by repeated application of a finite number of modes of combination. There are, of course, an infinite number of quotation mark names, since every expression has its own quotation mark name, and there are an infinite number of expressions. But on the theory of quotation we are considering, quotation mark names have no significant structure. It follows that a theory of truth could not be made to cover generally sentences containing quotations. (Davidson, 1979, p. 31)

The advantage of Davidson's (1979) proposal is self-evident. Unlike the lexicalist view, his approach captures our intuitions about reported speech being related somehow to quotation. Take for example a sentence like (10):

- (10) she said: "I'm a dancer"

The lexicalist view fails for the reasons exposed by Davidson (1979): such a hypothesis would entail the quote "*I'm a dancer*" to be some kind of lexical item inserted by the grammar of English into that sentence's structure, that being impossible as it would in turn imply that the lexicon is of infinite size. There is no requirement implied in Davidson's (1979) approach that the lexicon should include every possible quote in every possible language and every possible ungrammatical expression. His view, however, simply would state that a sentence like (10) has the following logical form:

- (11) 'Using words of which these are a token, she said that she was a dancer'

Davidson (1979, p. 39) also explains cases of *mixed quotation* like (12a), for which he proposes a logical form like (12b):

- (12) a. Quine says that quotation "...has an anomalous feature"
b. 'Quine says, using words of which these are a token, that quotation has an anomalous feature'

By identifying quotation with demonstration (i.e. the *demonstrative theory of quotation*) and placing it at the level of compositional meaning, we gain another advantage: this allows us to unify metalinguistic quotation and direct speech as the same phenomenon and under the same explanation. This

corresponds to the rough data shown in §1, namely that both in metalinguistic quotation and in direct speech we are always able to quote expressions that are ungrammatical or in a different language.

The key point that distinguishes compositional theories from lexicalist ones like those discussed in §2.1.1 is that the former ones consider quotation to be a mechanism that operates on a propositional level, while the latter consider it to operate on a lexical level. However, compositional views on quotation are not always Davidson-like demonstrative theories. An example of this is the theory proposed by Cappelen and Lepore (2007): according to them, quotation is a functor that is applied to a certain quotable expression e and triggers a quotational interpretation, such that $quote(quote(e))$ quotes $quote(e)$. Most work assuming the compositional view has been devoted to the study of mixed quotation (Cappelen and Lepore, 2007; Recanati, 2001; Stainton, 1999) and to determining whether quotation is some kind of demonstrative or just some abstract semantic operator.

In summary, under this view, quotation is a linguistic phenomenon that has to be defined from a semantic point of view, based on the analysis of truth-conditions of quotational sentences. This leads to discussions like whether mixed quotation can be considered part of the same phenomenon as direct quotation or not, etc. As I will show in §2.2, I believe there is another way to define quotation, much more consistent with the linguistic data that are available.

2.2 A syntactic definition of quotation

The common trait of the theories exposed so far is that they define quotation from a semantic point of view, taking either metalinguistic discourse or reported speech as their starting point, including different forms of mixed quotation. This leads to the consequence that the expressions that are enclosed by quotation marks in sentences (13)–(16) should all be considered to be quotes:

- (13) “dogs are to” is ill-formed
- (14) David said: “these are diamond dogs”
- (15) David said that “these are diamond dogs”
- (16) Silvia said that quotation “hasn’t any anomalous feature”

There is an intuition that we ought not to neglect: in all those sentences, the apparent quotes are expressions that effectively refer to some other expression *as expressions*. This happens even in (16), where we would state

that “*hasn’t any anomalous feature*” is a quote exactly because we are reproducing someone’s expression by literally showing it as it is. It is because of this semantic point of view that we would not hesitate in considering the previous examples as belonging to the very same phenomenon of quotation.

However, the main problem with this is actually an old one in linguistics: one meaning might certainly be expressed by more than one syntactic form, the classic example being the active/passive voice alternation:

- (17) Silvia saw Anna
- (18) Anna was seen by Silvia

As it is really well-known, both (17) and (18) convey exactly the same meaning, despite showing different syntactic configurations. Exactly the same happens with the following examples, in the case of reporting someone else’s expressions:

- (19) Silvia said: “quotation hasn’t any anomalous feature”
- (20) Silvia said that “quotation hasn’t any anomalous feature”
- (21) Silvia said that quotation hasn’t “any anomalous feature”
- (22) Silvia said that quotation hasn’t any anomalous feature

All sentences in (19)–(22) are true if and only if Silvia said that quotation has not any anomalous feature. From the point of view of meaning, they are equivalent, but are they from a syntactic point of view?

2.2.1 Grammatical opacity and quotation

As already stated in the introduction of this work, there are a couple of properties quotation has that are syntactic in nature, not semantic. Let us consider the (1)–(3) paradigm again, copied here again and expanded for convenience:

- (23) Silvia said: “Alba is dead”
- (24) Silvia said: “T’estimo”
- (25) Eliseus said: “people is good”
- (26) “abu baba him-baba babá” is a sentence in Linguisticspaperese
- (27) “it is possible how?” is bad English

The whole literature on quotation agrees in taking (23)–(27) as cases of so-called direct quotation (Bonami and Godard, 2008), regardless of the fact that the quotes in (26) and (27) are being used to metalinguistically refer to

some expression while the one in (23) is used for reporting speech directly. As already described before, it is interesting to note that direct quotes, no matter if used for direct speech or metalinguistic purposes, always allow quoting expressions that are ungrammatical (e.g. (25) and (27)), in a foreign language (e.g. Catalan in (24)) or even in an invented one (e.g. (26))!

In summary, we might state that direct quotation allows quoting expressions that are unknown to the grammar of the language that is being used to quote them. This means that quotes are *grammatically opaque*, a property that I formally define as follows:

- (28) *Grammatical opacity*
 For two expressions S, Q , such that $Q \subset S$, $g(S) = \top$, Q is opaque for a grammar G without quotation iff $g_G(S) \vee g_G(Q) = \top$, where $g_G(X)$ is a function taking an expression X and returning \top if X is grammatical according to G or \perp if ungrammatical.

This formalization basically states that some subsegment in a sentence is said to be grammatically opaque when that subsegment does not make the whole sentence ungrammatical. In this definition, it is assumed that expressions are modelled as *lists* of elements, such that if an expression A is a subsegment of an expression B , then $A \subset B$. For example:

- (29) $[_S [_Q \text{ Silvia is}] \text{ a girl}]$

In (29), Q is a subsegment of S . Notice that the notion of subsegment does not depend on the notion of syntactic constituent, but only on the bare linear order of the elements constituting a sentence.

Additionally, in the definition for grammatical opacity, it is assumed that the grammar G of a language L evaluates the grammaticality of some sentence X by making use of a function $g_G(X)$. The main intuition behind this is that there is one grammatical judgment for some expression X ,⁴ but given that there are multiple languages, there have to be different “versions” of this function for each language in the world. In any case, this is not meant to be a statement about the nature of the human faculty of language, but a formalization useful to the purposes of this research.

The fact is that *only quotes are grammatically opaque*. The reason is self-evident: unquoted elements cannot be opaque because, if they were, then language would be chaotic and unparsable; any possible ungrammatical expression would be allowed by the grammar of the language, this being equivalent to denying the existence of grammar itself. However, in order

⁴I am abstracting away the fact that some expression might be judged differently by different speakers.

to make it possible to quote any possible expression, e.g. to explain why some expression is ungrammatical, it is necessary to allow quotes to violate the grammar of the language. In some sense, quotes are “islands” where grammar is not enforced in order to make metalinguistic and direct reported speech always possible in any context.

In simpler terms, I propose that there is a biunivocal relationship between quotation and grammatical opacity, i.e. that some expression Q is opaque in some other expression S if and only if Q is a quote in S . Formally:

- (30) *Quotation–Opacity Hypothesis (QOH)*
 For Q, S expressions, such that $Q \subset S$, and G a grammar such that $g_G(S) = \top$; $opaque(Q, S, G)$ iff $quote(Q, S)$.

In (30) it is assumed that there is a predicate function $opaque(X, Y, Z)$ such that it returns \top if the expression X is grammatically opaque in Y for a grammar Z , and that there is a predicate function $quote(Q, S)$ such that it returns \top if Q is a quote in S .⁵ Both functions return \perp if their respective conditions are not met.

2.2.2 Testing our hypotheses

Let us now test how this theoretical apparatus helps us distinguishing quotes from unquoted expressions. Take for example cases like these, in which quotation marks have been stripped off in order to simulate an acoustic stimulus:

- (31) men are too stupid
 (32) men are too stupid is a sentence

QOH essentially claims that if some subsegment of a sentence can be replaced by an ungrammatical one, then the original subsegment was a quote, as also, presumably, the replacing subsegment. Replacing *too stupid* by the obviously ungrammatical *too after* in both (31) and (32) results in (33) and (34), respectively:

- (33) * men are too after
 (34) men are too after is a sentence

Sentence (34) is untrue, but perfectly grammatical if *men are too after* is considered to be a quote. This shows that at least *too stupid* is quoted in

⁵This is the correct definition for such a predicate. Defining it as $quote(Q)$ is equivalent to embracing the lexicalist theory of quotation, where quotes are quotes on their own and not in the context of some sentence; thus the two parameters required by the function.

(32), as it is grammatically opaque in that specific syntactic context. On the other hand, the same replacement in (31) yields the ungrammatical (33).

The problem is now to determine which elements surrounding *too stupid* in (32) (repeated below as (35)) are part of the quote. Let us now try replacing *too stupid is a sentence* by *too after*:

(35) men are too stupid is a sentence

(36) *men are too after

Again, this yields an ungrammatical sentence, thus *too stupid is a sentence* is not the quote we are looking for; the only possibility left now is, obviously, to extend our search leftwards. The next step is, then, to try replacing the subsegment *are too stupid* of (37) (= (32)) by *too after*. Of course, this yields the grammatical (38):

(37) men are too stupid is a sentence

(38) men too after is a sentence

Although we have replaced only *are too stupid* by *too after*, this only means that *are too stupid* is part of an opaque subsegment in (37). Similarly to the situation in (34) when testing for the opacity of *too stupid*, there is still the possibility that *are too stupid* is just a part of the opaque subsegment, not the whole of it. Therefore, we have to test whether *men* is also part of the quote. The answer, which the reader surely already knew by the intuition that in (38) the whole *men too after* is interpreted as a quote, is that *men* is indeed part of the *men are too stupid* quote in (32). This is shown below by the fact that replacing the whole *men are too stupid* subsegment of said example by *too after* yields an untrue yet perfectly grammatical sentence:

(39) too after is a sentence

The reason why the opaque subsegment of (32) is the whole subject is that quotes can only be located at NP positions, the only type of constituent where opacity can be achieved. To show this, the segment *el hombre es malo*, which is ungrammatical for the English grammar, is inserted in the following examples replacing various positions of the sentence (40). This yields (41a)–(41g):

(40) Silvia is writing something on a white sheet of paper very quickly

(41) a. Silvia is writing [_{NP} el hombre es malo] on a white sheet of paper very quickly

- b. Silvia is writing something [_{PP} on [_{NP} el hombre es malo]] very quickly
‘Silvia is writing something on the phrase “el hombre es malo” very quickly’
- c. *Silvia is writing something on a(n) [_{AP} el hombre es malo] sheet of paper very quickly
- d. *Silvia is [_{VP} el hombre es malo]
- e. *Silvia [_{IP} el hombre es malo]
- f. *Silvia is writing something [_{PP} el hombre es malo] very quickly
- g. *Silvia is writing something on a white sheet of paper [_{AdvP} el hombre es malo]

As seen above, it seems that quotes can only be inserted in NP positions without rendering the whole sentence ungrammatical. An explanation for this is that for an element to be grammatically opaque, it must be a semantically atomic value, i.e. an element that is not decomposable in smaller semantic units and, *therefore*, it is a zero-place predicate. The reason for identifying non-decomposability with lack of any arguments is that if an element asks for some arguments, then it will have a reference and be communicatively useful if and only if all of its arguments are saturated; in other words, a non-zero-place predicate is only meaningful when composed and therefore, it is actually decomposable, namely in the part usually annotated as a lambda function and in the argument it takes. Of all syntactic categories, only proper names comply with the property of being semantically atomic: common nouns have their extralinguistic reference as their argument; adjectives and adverbs have the head they modify as their argument; and the syntactic arguments of verbs are also their semantic ones (Dalrymple, 2001). Given that a grammatically opaque segment is such that does not affect the grammaticality of the sentence containing it, it is impossible that it could take any argument; if a function takes arguments, it is implied that the function knows how to “process” and yield some result from the arguments that have been passed to it. A grammatically opaque segment cannot perform this essentially because it is a segment that is ignored by the grammar of the language, therefore it remains unparsed (even if it could in fact be parsable) and, consequently, it cannot take any arguments as it would not be able to handle them in any way. In conclusion, only zero-place predicates can be opaque and the only zero-place predicate there is is proper names. Given that the category of proper names is N, then the syntactic position of quotes has to be NP.

It is easy to see that the hypothesis that quotes can only be NPs is a corollary of QOH: if some expression is a quote if and only if it is opaque and only NPs are opaque, then quotes are NPs. This idea corresponds to the intuition held by the proponents of the lexicalist view of quotation: quotes are like proper names, but in the sense that they occupy NP positions. On the other hand, the view of quotation proposed here is compositional: it is based on the *syntactic* property of grammatical opacity. However, the main difference between an approach like this and, e.g. Cappelen and Lepore's (2007), is that the compositional level in which quotation is analyzed here is syntax and not semantics. In other words, QOH is a syntactic definition of quotation which successfully predicts where quotes can be inserted and where they cannot.

2.2.3 Citation vs. quotation

As stated before, quotes can only be inserted in NP positions. However, what about cases like (42)–(44)?

- (42) the president said that “the economy is doing fine”
- (43) the president said that the economy “is doing fine”
- (44) the president said that the economy is doing “fine”

These and similar constructions are cases of what Davidson (1979) calls *mixed quotation*, subsequently defended by Cappelen and Lepore (2007). Mixed quotation consists in constructions that apparently introduce a partial quote intertwined with the quoting sentence's syntactic structure. For example, in (42), the addressee knows that *the economy is doing fine* was uttered by the president. On the other hand, in (43), it is only guaranteed that *is doing fine* was uttered by the president, but not necessarily *the economy is doing fine*; in fact, he may have said *our economic plan is doing fine*.

It is completely naïve to deny the reported speech character of such examples. It is not a redundant mechanism: so-called mixed quotation is in fact a legitimate and well-defined device to introduce someone else's speech in one's own discourse. However, this does not imply that quotation, as understood in this work, is actually operative in “mixed quotation”. In fact, it is not, as these simple opacity tests show, where the respective apparent quote in each a. case is replaced in b. by the apparent quote “*la economía va bien*”:

- (45) a. the president said that “the economy is doing fine”
b. * the president said that “*la economía va bien*”

- (46) a. the president said that the economy “is doing fine”
 b. * the president said that the economy “la economía va bien”
- (47) a. the president said that the economy is “doing fine”
 b. * the president said that the economy is “la economía va bien”

If in these cases of “mixed quotation” quotation was really operative, then all a. cases in (45)–(47) would be grammatical, because of grammatical opacity. However, there is no trace of opacity in any of these cases. It is impossible to “quote” ungrammatical or foreign expressions in a “mixed quotation” structure, exactly because it consists in quoting some expression integrating it into the sentence’s syntactic structure! Opacity is, thus, impossible to achieve here per definition.

The problem as I see it is that we ought to distinguish the discursive function of language that allows speakers to refer to other speakers’ utterances, namely *citation*, from one specific grammatical mechanism that is sometimes used for citation, namely *quotation*. There is no necessary one-to-one correlation between citation and quotation: metalinguistic quotation does not really reproduce a previous utterance, while indirect speech is a form of citation where quotation is undoubtedly not present at all. In fact, “mixed quotation” is just a pragmatically-enriched form of indirect speech, where quotation marks introduce some additional authority to the reproduced segment.

A similar problem occurs with cases like the following ones:

- (48) “John”, Mary said, “it’s raining outside!”

This kind of construction is what Bonami and Godard (2008) call Incidental Quotative Clause (IQC) constructions. In these sentences, there is an apparently quoted subsegment to which an incise headed by a quotational verb is added, e.g. the segment *Mary said* in the example above. According to these authors, these incises behave like incidental adjuncts that are adjoined to a sentence, the only special property of IQC constructions being that the sentence bears some features required by the quotational verb (cf. §3.1.2 for the syntactic discussion). However, despite of the initial appearances, it is easy to see that the elements graphically enclosed by quotation marks in (48) are not actually quoted, not at least under our view of quotation; replacing the apparent quotes of (48) by ungrammatical expressions yields the ungrammatical (49):

- (49) * “Run you”, Mary said, “una cucaracha grün”

Therefore, there is evidence to conclude that the apparently quoted subsegments in IQC constructions are not grammatically opaque and, in con-

clusion, that they do not fall under the requirements here presented for an expression to be quoted.

Another case of citation that has to be discarded is Recanati's (2001) so-called "open quotation", of which (50) is an example (taken from Recanati, 2001, p. 649):

- (50) Stop that, John! "Nobody likes me", "I'm miserable" ... Don't you think you exaggerate a bit?

It is true that we can replace the apparent quotes in (50) by foreign or even ungrammatical expressions, e.g. *niemand mag mich* and *I be miserable*, respectively:

- (51) Stop that, John! "Niemand mag mich", "I be miserable" ... Don't you exaggerate a bit?

The problem with this is that such segments are not opaque either: they are autonomous segments, meaning they are not part of any bigger expression and, therefore, they do not conform to the context the definition of grammatical opacity (28) requires for it to be applicable to some expression. A case like (51) may be qualified as nonsense or a discourse that exhibits both code-switching and improper use of the English language, but not as ungrammatical as a whole! Syntax, as it is well known, deals with the configurational properties of the minimal meaningful unit of expression, namely the sentence,⁶ not with the configuration of discourse. It is much more reasonable to classify "open quotation" as another kind of context-dependent citation.

Discarding such cases of citation from our theory of quotation allows us to solve an obvious problem spotted by Cappelen and Lepore (2007, p. 139), namely what they call the "syntactic chameleonism" of quotation. Their attempt to solve it is by introducing the existence of a quoted QXP for every possible XP available in a language, but such idea is defective by design: if any possible XP can be quoted, then this theory incorrectly predicts all ungrammatical expressions as grammatical, as the offending segment could be considered a QXP. The problem becomes unsolvable if we accept that quotes are opaque only sometimes; it is contradictory and illogical. There seems to be no other way than to accept that quotation is a matter of syntax,

⁶According to De Saussure (1916/1984) and the structuralist tradition, the minimal meaningful unit of expression is the lexeme, as it is the smallest unit that has some meaning associated to it. The problem with this is that even an utterance like *dogs* is actually a sentence, namely one composed by just one noun. This is much clearer in languages that have morphological case marking, where an utterance like that may receive any case, its final interpretation being different depending on the discourse context.

one device among many others to achieve reference to a discourse different than one's own.

2.3 Conclusions

In summary, I propose a syntactic definition of quotation over the semantically based ones that have been defended by other authors in previous research. This definition of quotation is stated in terms of the previously defined property of grammatical opacity, such that an element Q in a sentence S is a quote if and only if Q is grammatically opaque in S . This led us to discover that quotes can only occupy NP positions, given that grammatical opacity only arises in that kind of constituent. Finally, this led us to discard “mixed quotation” and IQC constructions from our object of study, mainly because the segments that are apparently quoted in such constructions do not comply with the definitions proposed here. This leaves us with the so-called direct quotation as the only object of study in this work.

Chapter 3

The syntax of quotation

We have reached the most important chapter of this work. In this chapter, I will defend my LFG-based theory of the syntax of quotation, after briefly reviewing the very scarce literature I have found on this topic, most of it done under the Minimalist Program of Chomskyan generative grammar (MP; Chomsky, 1995). As will be shown, the flaws of the work previously done on this topic are enough to motivate the need for a new theory to explain the data that has been exposed throughout this work so far.

3.1 Previous theories

As already mentioned in §1, only very few theories about syntactic phenomena related to quotation have been proposed in the past, most of them developed following Chomskyan standard generative grammar, especially in terms of the MP. Two major works done from this perspective are Collins and Branigan's (1997) and Suñer's (2000). On the other hand, as far as I am aware, the only theory that has been developed in a different framework is Bonami and Godard's (2008).

3.1.1 Work based on MP

Let us start this review by briefly discussing the contributions made by researchers using the MP framework.

The paper by Collins and Branigan (1997) attempts to explain the so-called “quotative inversion” in English, namely the phenomenon consisting in the optional inversion of subject and verb in constructions like these (adapted from Collins and Branigan, 1997, pp. 1–2):

- (52) a. “Don’t turn back!”, warned Marcel

- (53) b. “Don’t turn back!” , Marcel warned
 a. “What on earth” , asked Sylvie, “shall we do with the dean?”
 b. “What on earth” , Sylvie asked, “shall we do with the dean?”

The reader will quickly identify (53) as a case of an IQC construction, as shown in §2.2.3; the apparent quotes in that example are not real quotes according to my hypotheses. The cases in (52) instead can be considered to be real direct quotation examples, as they appear to show grammatical opacity:

- (54) a. “¡No te gires!” , warned Marcel
 b. “¡No te gires!” , Marcel warned

The main problem with their research paper is that the authors purposely define no position for quotes in cases like the ones in (54). For example, this is the analysis Collins and Branigan (1997, p. 11) give for cases (55)–(56):

- (55) “When on earth will the fishing begin today?” , [_{CP} O_i [_{Ag_rP} asked Harry t_i]]
 (56) [_{CP} O_i [_{Ag_rP} asked Harry t_i]]: “When on earth will the fishing begin today?”

According to them, the quote is not moved from anywhere, but merged directly into its final Spell-out position, which they leave undefined in order to explain cases of IQC constructions like (53a). However, according to the postulates of MP, the strong [DP–] feature of *ask* has to be checked somehow, namely by an object; therefore, an operator *O* is postulated to fill that position and also, for theory-internal reasons, to occupy C in order to explain the inverted VS order: the special inverting [+quote] C selects a T node that has a weak EPP feature, thus the subject is raised covertly, as opposed to “normal” conditions (Collins and Branigan, 1997). Other details are also taken in account in their paper, but I think this suffices for getting a more-or-less basic idea of Collins and Branigan’s (1997) proposal. As the reader can notice, this leaves us with some (suboptimal) explanation of so-called “quotative inversion”, but it leaves us without answering one of the most important questions about the syntax of quotation, namely where quotes are placed.

Suñer (2000) follows the tradition started by Collins and Branigan (1997), with some slight improvements, but still approaching the issue from the problem of “quotative inversion”, even in IQC constructions, a construction that, according to the discussion in §2.2, does not include any quote. In fact, her

analysis does not differ much from Collins and Branigan’s (1997), also positing a QOp operator that essentially behaves the same way as theirs, with only small differences in the movement chains and landing sites that Suñer (2000) proposes for it. However, her proposal has a major difference compared to Collins and Branigan’s (1997): she proposes that the quote is inserted into the positions occupied by the QOp and the traces left by its movement (Suñer, 2000, pp. 543–544), this “linkage”, as she calls it, occurring after the derivation takes place. This idea is slightly more acceptable than Collins and Branigan’s (1997) lack of definition on the position of quotes, but it is still unsatisfactory because it is also aimed to explain IQC constructions. Given the arguments against considering constructions of such kind as quotational ones, Suñer’s (2000) hypotheses have to be discarded as well.

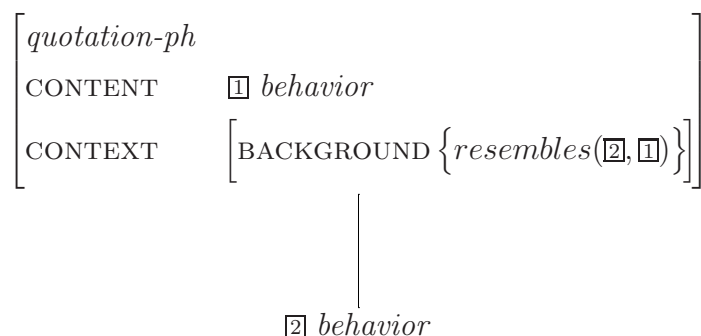
The real problem behind these two proposals seems to be the adopted framework itself: MP or Chomskyan generative grammar in general. One of the main assumptions made by Chomsky (1995) is that syntax works as a derivation that uniformly takes a set of lexical items and returns a Logical Form (LF) that is interpretable by the so-called Conceptual-Intentional interface.⁷ Assuming this derivation to be uniform has two consequences that make quotation impossible to really explain in this framework: it requires establishing that the same rules apply at all points of the derivation and it requires, therefore, everything, from checking of lexical features to semantic properties, to be encoded in terms of constituent structure, which according to MP responds to universal principles of language. However, quotes are grammatically opaque, therefore they may insert a subsegment that is ungrammatical without affecting the grammaticality of the whole; letting ungrammatical or foreign quotes to be directly inserted by Merge into the derivation would introduce features that the grammar of the language would not be able to parse and, therefore, they would not be possible to check and be deleted/erased either. That is the reason why both Collins and Branigan (1997) and Suñer (2000) place the quote outside of any distinguishable position in the constituent structure of the sentence, but propose that a non-observable and theory-internal operator takes the place of the quote.

3.1.2 Bonami and Godard’s (2008) work

Finally, there is an interesting research paper which I would not like to close this section without discussing, at least briefly: the proposal by Bonami and Godard (2008) on French direct quotation, based on HPSG.

⁷I am ignoring the derivation from Spell-out to the Phonetic Form (PF) as it is clearly irrelevant for my purposes

These authors, following Clark and Gerrig (1990), consider quotes to be demonstrations of behaviors or events, and given the lexicalist orientation of HPSG, the distinction between referring to some behavior and just being some linguistic behavior (i.e. Quine’s (1981) concepts of mention and use, respectively) has to be codified in a way that quotational verbs like Fr. *dire* ‘to say’ could select for. This leads Bonami and Godard (2008, p. 364) to propose the following unary rule for what they call *quotation phrases* or *quotation-ph*, shown at Figure 3.1.



Without entering into details, it is easy to see how Figure 3.1 manages to explain grammatical opacity, by “hiding” the content of the quote under the *resembles*(X, Y) relation.

⁸Related to this, there is also no explanation as for why quotes occupy NP positions, but this is a minor point, partially explained by the fact that the authors establish the position of quotes at the subcategorization features of the head verb. However, this is a detail that cannot be ignored in a general theory of the syntax of quotation like mine.

consequence of their model. Let us ignore the former and only concentrate in the latter; for this, take the following sentence as an example of an IQC construction (adapted from Bonami and Godard, 2008, p. 370):

(57) “the president”, said Mary, “has already arrived”

According to Bonami and Godard (2008), (57) is to be analyzed as a sentence that is a quotation-ph as a whole to which an IQC incise *said Mary* has been adjoined to the segment *the president has already arrived*, which they qualify as a quotation-ph for semantic reasons. However, as they do notice, IQC constructions cannot include foreign “quotes”; therefore, they introduce a special IQC rule that selects only behaviors that are in the same language of the IQC. This construct is quite unnecessary if it is assumed that there is no quotation in a sentence like (57): as already said, pure adjunction to an autonomous sentence can explain this kind of data, the citational meaning being a semantic problem, not a syntactic one. Moreover, their model is surprisingly in contradiction to their statement that Recanati’s (2001) “open quotation” constructions are independent clauses and therefore not quoted (Bonami and Godard, 2008, p. 361). Given that they allow a head clause to be a quotation-ph in their analysis of IQC constructions, then “open quotation” ones could be analyzed as quotation-ph’s to which no incise is adjoined.

In summary, Bonami and Godard (2008) serves as a source of inspiration for any approach to this topic that is made within any unification-based theoretical framework of syntax like both HPSG and LFG are. This proposal shows a creative, yet not perfect, way to somehow block unification in order to explain grammatical opacity: by categorizing quotes as references to instances of behavior and letting the grammar know how to select for them. Their hypothesis fails in giving a complete picture of quotation and makes the analysis of IQC constructions too convoluted, while my approach of treating them as non-quotational cases of citation seems to be much more economical and clarifying.

3.2 A proposal based on LFG

In this last section of this chapter, I will present my hypothesis about how quotes are inserted into sentences. Let us recapitulate the main properties that quotes have in terms of syntax before starting to discuss how to model their behavior in LFG. As has already been mentioned, quotes are grammatically opaque, meaning that the ungrammaticality of a quote does not make

the sentence that contains it ungrammatical (cf. (28) and (30)). Also, we also know that they seem to occupy positions that are reserved to NPs.

First, the f-structure of quoted material will be presented. Second, I will propose the c-structure/f-structure mapping that is required for explaining the properties of this construction. Finally, some evidence from Romance languages will be discussed in order to expand the cross-linguistic validity of my hypothesis.

3.2.1 Quoted f-structures as PRED values

Intuitively, what is needed in order to model grammatical opacity in this framework is to isolate the quote’s f-structure in a way that unification and any “access” to its values is blocked, e.g. any access due to inside-out or outside-in functional uncertainty. The most reasonable way to achieve this is to take advantage of the fact that quotes are expressions whose meaning refers to some expression; from an LFG perspective, this means that the quote’s PRED value is the f-structure of the quoted material. This way, the f-structure of (58) is the one shown at Figure 3.2 (features irrelevant for the discussion are omitted, following standard practice in LFG):

(58) he said: “John ate an apple”

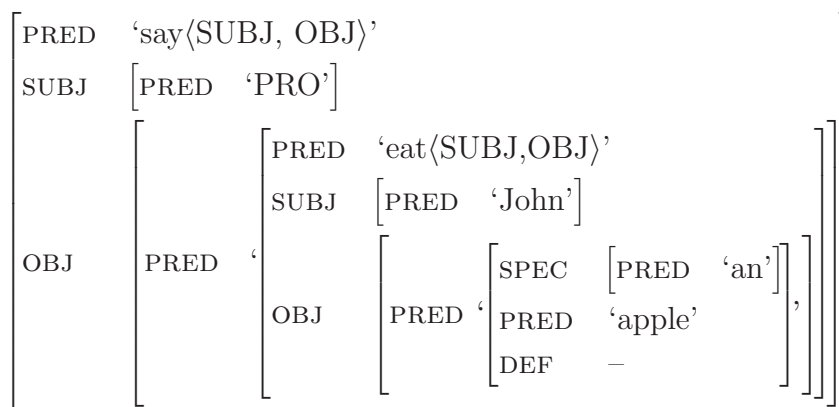


Figure 3.2: f-structure of (58)

As shown in Figure 3.2, the f-structure of the quote is the PRED of the OBJ. Even though it is obvious that there is an internal structure to the quoted f-structure, PREDs are semantic forms (Dalrymple, 2001) and as such their values are atomic from a syntactic perspective. In fact, the quoted f-structure is shown as a full-fledged f-structure just because of conventional

reasons: other notational devices could be considered as long as it is made clear that the quote's PRED value is an expression in itself, not a lexical item.

For ungrammatical or foreign quotes like (59) or (60), the f-structures are the ones shown in Figures 3.3 and 3.4, respectively.

(59) he said: “John ate an apples”

(60) he said: “Juan comió una manzana”

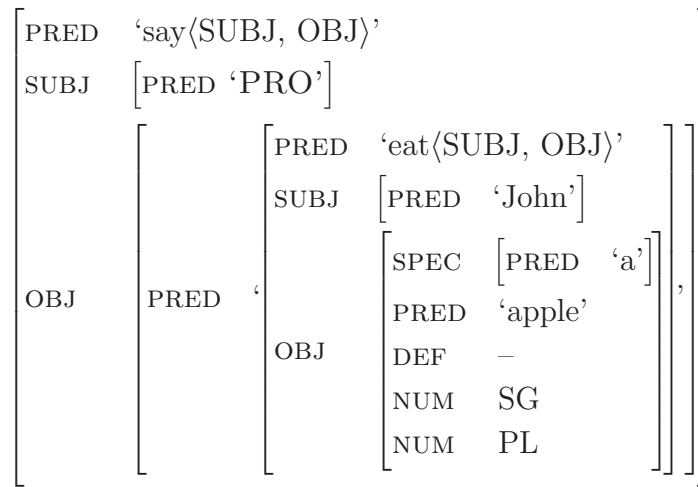


Figure 3.3: f-structure of (59)

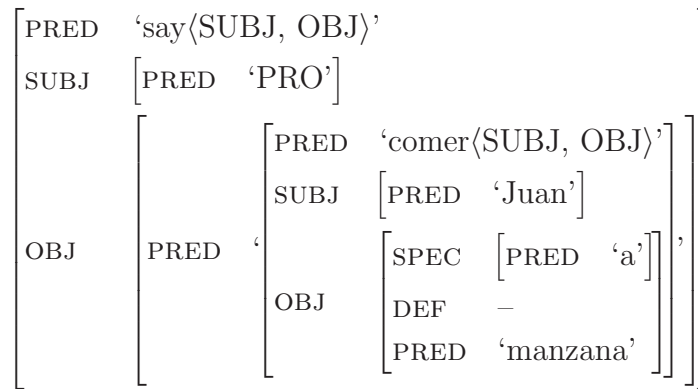


Figure 3.4: f-structure of (60)

In both cases, the quoted f-structures are unknown to the grammar of English, but they are nevertheless expressions and as such they do have f-

structures (an illformed one in the case of (59)). However, the illformedness or unparsability of these f-structures is irrelevant precisely because they are the value of PRED and, therefore, treated as an atomic value.

Up to this point, this hypothesis correctly predicts that quotes cannot be placed in positions where a PRED with argument list is provided, i.e. positions occupied by verbs. If quotes could be placed at VP/IP, the resulting f-structure would lack any argument list to check Coherence and Completeness of governable functions against, thus failing to comply with both f-structure constraints altogether. However, we still need a way to block non-NP positions that are associated to PRED values like adjoined APs, AdvPs or PPs. Given that this last problem is reduced to a problem of constituent types, the place to solve it is very probably at c-structure; this is done in the next subsection.

3.2.2 C-structure/f-structure mapping

In §2.2.2 it was shown that quotes can only be placed in NP positions. The problem of restricting quotes to a certain constituent seems to be one of purely configurational nature, not functional in any case, even though the hypothesis proposed in the previous subsection already ruled out certain positions as impossible, namely positions dedicated to verbs and auxiliaries that usually provide PRED values with argument lists.

At first sight, quotes seem to be placed in NPs. The following annotated phrase structure rule for DPs would suffice to yield f-structures like the ones proposed earlier, at least in English and probably most configurational languages:

$$(61) \quad \text{Provisional phrase structure rule for quotes}$$

$$\text{DP} \quad \rightarrow \quad \left(\begin{array}{c} \text{D} \\ (\uparrow\text{SPEC}) = \downarrow \end{array} \right) \quad \text{NP} \quad (\uparrow\text{PRED}) = \downarrow$$

The advantage of accepting here the DP-Hypothesis (Abney, 1987), here assuming Dalrymple's (2001) version, is mainly that it allows annotating NPs in general, regardless of their position. For this, I assume DEF and NUM to be percolated from the determiner's lexical entry. Otherwise, several rules should be proposed in order to deal with all positions NPs could occupy, namely specifier of IP, complement of VP, etc. From a theoretical point of view, that would be obviously less economical and less desirable than a single rule like (61).

Therefore, the c-structure of a case like (62) is the one shown at Figure 3.5.

(62) “Johnny” was written there

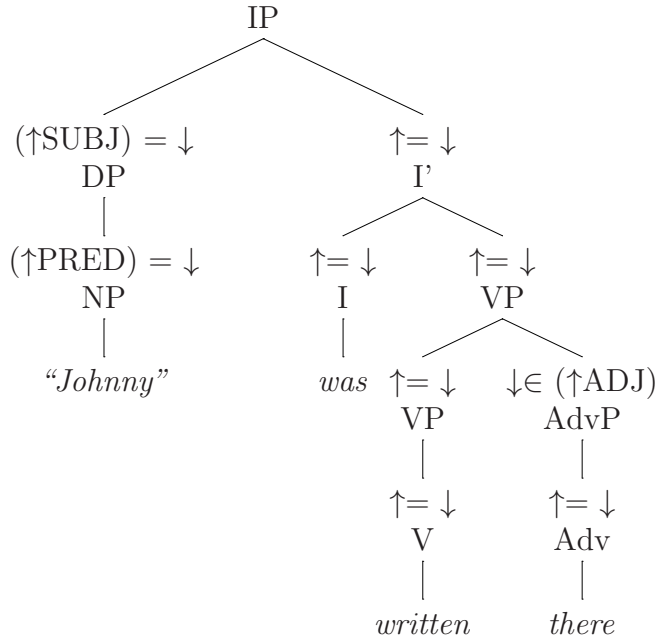


Figure 3.5: c-structure of (62)

The problem with this is that there is no categorial information stated for the quoted content: the quote “Johnny” is directly assumed to be the NP itself. Therefore, let us expand our phrase structure rules as shown below:

(63) *Provisional phrase structure rules for quotes (2)*

- a.
$$\text{DP} \rightarrow \left(\begin{array}{c} \text{D} \\ (\uparrow\text{SPEC}) = \downarrow \end{array} \right) \text{NP} \quad \uparrow = \downarrow$$
- b.
$$\text{NP} \rightarrow \begin{array}{c} \text{XP} \\ (\uparrow\text{PRED} = \downarrow) \end{array}$$

If (63) is accepted, then the prediction is made that quotes can be modified by determiners. This seems to be true for (64):

(64) every “Juan camina mal” was painted quickly

The f-structure and c-structure of (64) are shown in Figures 3.6 and 3.7, respectively.

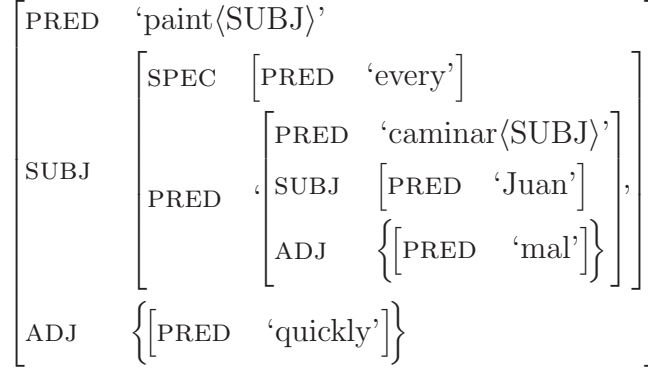


Figure 3.6: f-structure of (64)

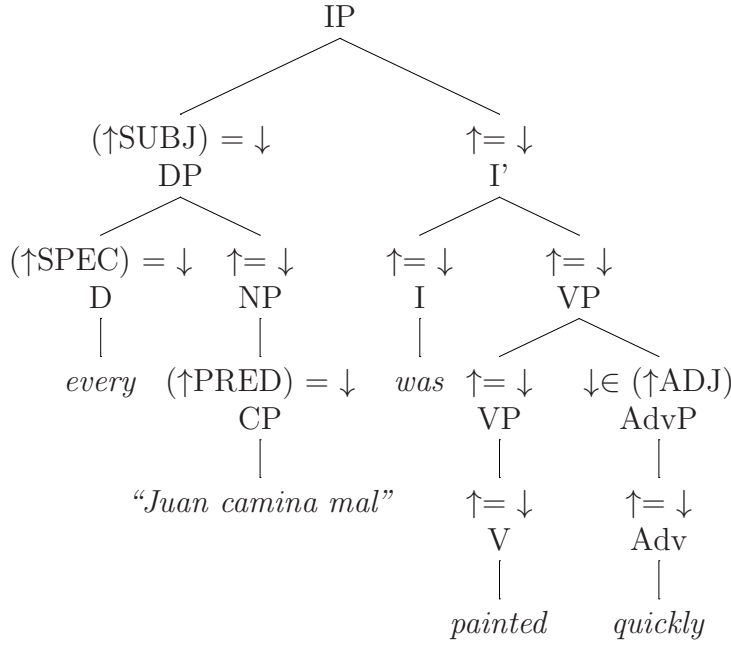


Figure 3.7: c-structure of (64)

However, it is self-evident that (63) does not predict why the following cases are also possible, where the quote is modified by an adjective:

(65) John analyzed the incorrect “an apples”

Such a case calls for a revision of (63), namely establishing that quotes are not placed under NP, but rather N', thus allowing for this kind of modification by adjectives without disrupting the explanation given for all previous cases

of unmodified quotes; in those cases, the NP would just consist of an N' which introduces the quote. My proposal is now shaped like this, based on the analysis of adjective adjuncts by Dalrymple (2001, p. 257):

(66) *Provisional phrase structure rules for quotes (3)*

- $$\begin{array}{lll}
 \text{a.} & \text{DP} & \rightarrow \left(\begin{array}{c} \text{D} \\ (\uparrow\text{SPEC}) = \downarrow \end{array} \right) \text{NP} \quad \uparrow = \downarrow \\
 \text{b.} & \text{NP} & \rightarrow \left(\begin{array}{c} \text{AP}^* \\ \downarrow \in (\uparrow\text{ADJ}) \end{array} \right) \text{N}' \quad \uparrow = \downarrow \\
 \text{c.} & \text{N}' & \rightarrow \text{XP} \\
 & & (\uparrow\text{PRED}) = \downarrow
 \end{array}$$

The f-structure and c-structure of (65) are shown in Figures 3.8 and 3.9 (cf. next page), respectively.

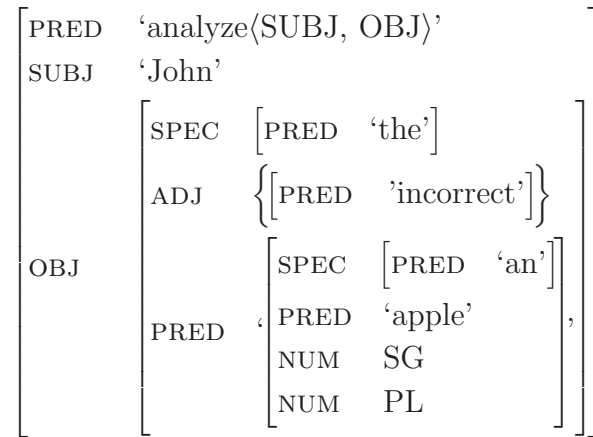


Figure 3.8: f-structure of (65)

There is no problem with proposing (66c) without any further constraint: rule (66c) is only applicable in quotational contexts per definition. There is no way to bypass this rule to include an unquoted arbitrary segment, thus successfully restraining the generation of c-structures only to what is needed to explain quotation.

Finally, a word has to be said about CPs that may appear modifying a quote, as in the following DP:⁹

(67) the “Nixon stinks” that was painted on the wall

⁹Thanks to Prof. Alex Alsina for this example.

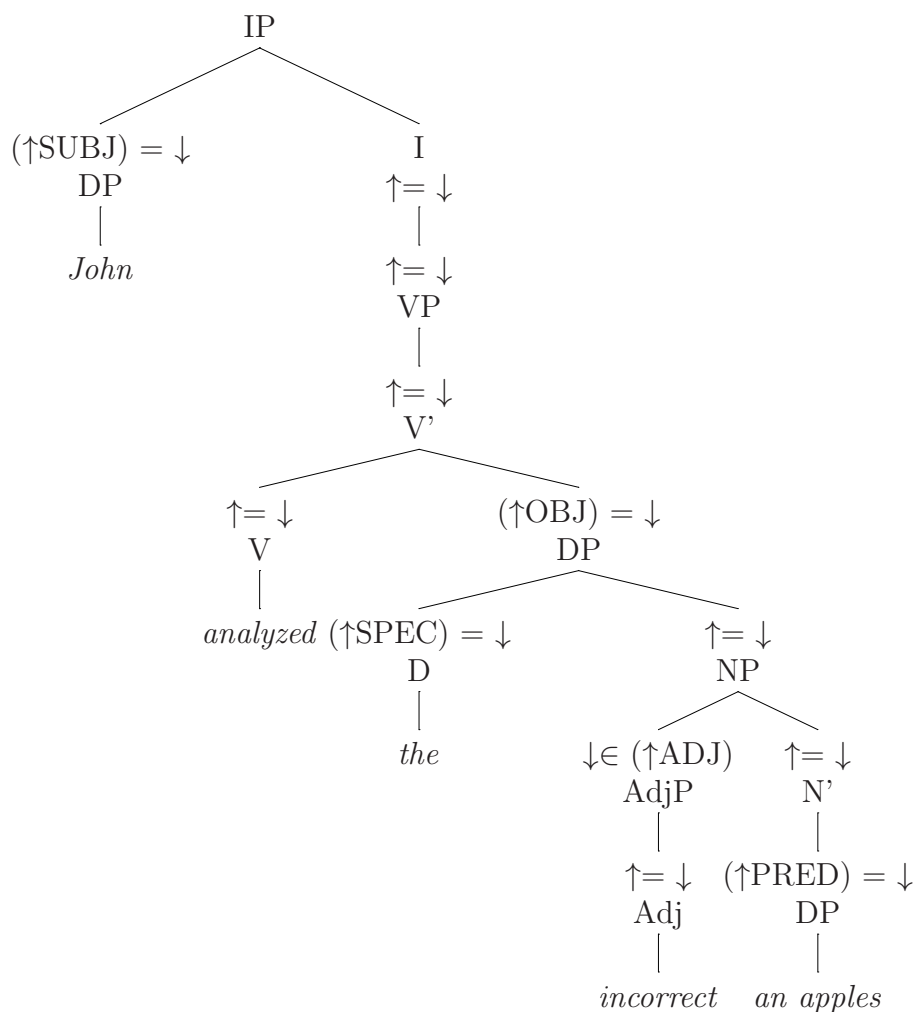


Figure 3.9: c-structure of (65)

In order to account for this kind of modifiers, a further modification of the phrase structure rules proposed so far has to be made, namely adding the rule (68c) for adjoining a CP to N', according to the analysis proposed by Dalrymple (2001, p. 402):

(68) *Provisional phrase structure rules for quotes (4)*

- a. $DP \rightarrow \left(\begin{array}{c} D \\ (\uparrow\text{SPEC}) = \downarrow \end{array} \right) NP$ $\uparrow = \downarrow$
- b. $NP \rightarrow \left(\begin{array}{c} AP^* \\ \downarrow \in (\uparrow\text{ADJ}) \end{array} \right) N'$ $\uparrow = \downarrow$

- c. $N' \rightarrow \left(\begin{array}{c} N' \\ \uparrow = \downarrow \end{array} \right) \quad \begin{array}{c} CP \\ \downarrow \in (\uparrow ADJ) \end{array}$
- d. $N' \rightarrow \begin{array}{c} XP \\ (\uparrow PRED) = \downarrow \end{array}$

Now we have a rough set of rules that takes into account how quotes are inserted into sentences. The reader should notice that of all rules in (68) only (68d) is specific to quotation; rules (68a)–(68c) are actually rules that valid for all English DPs in general.

Even though the set of rules (68) explains most of the syntax of quotes, in the following subsection I will present some further refinement that is still required in order to explain some features of quotes that are akin to morphological information of “regular” DPs.

3.2.3 “Lexical” features of quotes

In this work, I have made clear that quotes are terms whose NUM is SG and they force agreement accordingly:

- (69) “dogs” is grammatical
- (70) * “dogs” are grammatical

Moreover, evidence from Spanish (or other Romance languages) shows that quotes not only have a certain number, but also a certain gender; masculine, in the case of Spanish:

- (71) “perras” está mal escrito
 “*perras*” is badly written.masc.sg
 “perras” is badly written’
- (72) * “perras” está mal escrita
 “*perras*” is badly written.fem.sg

The problem is that we need to specify somehow that quotes have their own “lexical” features without actually being lexical items at all. The best solution here appears to make use of a defining equation that assigns a “default” number and gender to the quoted XP. In English, number is all that is needed, so this final proposal would be enough to make sure that quotes are always treated as singular terms and never as plural ones:

- (73) *Phrase structure rules for quotes in English*
- a. $DP \rightarrow \left(\begin{array}{c} D \\ (\uparrow SPEC) = \downarrow \end{array} \right) \quad \begin{array}{c} NP \\ \uparrow = \downarrow \end{array}$

$$\begin{array}{lll}
\text{b.} & \text{NP} & \rightarrow \left(\begin{array}{c} \text{AP}^* \\ \downarrow \in (\uparrow \text{ADJ}) \end{array} \right) \text{N}' \quad \uparrow = \downarrow \\
\text{c.} & \text{N}' & \rightarrow \left(\begin{array}{c} \text{N}' \\ \uparrow = \downarrow \end{array} \right) \text{CP} \quad \downarrow \in (\uparrow \text{ADJ}) \\
\text{d.} & \text{N}' & \rightarrow \begin{array}{c} \text{XP} \\ (\uparrow \text{PRED}) = \downarrow \\ (\uparrow \text{NUM}) = \text{SG} \end{array}
\end{array}$$

This set of rules gives a principled explanation of the difference of interpretation between (4) and (5) (repeated below for convenience) shown in §1.

(74) the dogs is a noun phrase

(75) the dogs are a noun phrase

The quotational interpretation of sentence (74) comes from the fact that the rule (73d) is applied, thus yielding a DP that is singular. On the other hand, (75) is not interpreted as a metalinguistic description of *the dogs* precisely because the “normal” rule for N' is used for it.

Moreover, these rules explain why ambiguities may arise when the quoted material is “singular”, e.g. in the following cases where quotation marks have not been added:

(76) a dog is a noun phrase

This case might be generated by the rules in (73) or by the regular ones for unquoted DPs in English. The only difference would be that two different c-structure/ f-structure pairs would be associated to this very same “surface form”. Again, the very same happens with the case (6) in §1.

Finally, a word has to be said for languages other than English. In Spanish, other Romance languages, German, etc. an additional constraint should be added to (73a), specifying the correct “default” gender for NPs. By this, cases (69)–(72) receive a successful explanation, while also leaving room for differences between languages. C-structure is the locus of syntactic linguistic variation and, therefore, it should not surprise anyone that even bigger differences may appear in other languages that have not been analyzed here.

3.2.4 Summary

In summary, this LFG proposal establishes that quotes are N' constituents that provide the PRED of the construction, thus yielding a way to model the notion of grammatical opacity at the f-structure level, given that PRED

values are semantic forms that are syntactically atomic. At the level of c-structure, I have proposed rules that are mainly valid for English but easily adaptable to other Western languages that also have gender constraints associated to quotes.

Chapter 4

Conclusions

In this work, I have shown that, in order to really understand the syntax of quotes, a revision of the concept of quotation had to be made. Otherwise, the result would have been, as it has been in previous work, a theory that would have attempted to explain irreconcilable data as the very same phenomenon; I am referring to the cases of citation that have been mistaken as cases of quotation.

Quotation has been redefined in terms of grammatical opacity (cf. (28) and (30)), i.e. quotes are subsegments of sentences that may be ungrammatical without rendering the sentences containing them ungrammatical. This leads to the conclusion that only direct quotation is “true” quotation, as it is the one that is able to insert any arbitrary quoted content either for metalinguistic description or faithful direct reported speech.

Syntactically, quotes have been modelled in LFG as f-structures that are used as PRED values. By this means, a satisfactory explanation is given for grammatical opacity as PRED values are semantic forms that are opaque or “non-accessible” to other f-structure functions. At the level of c-structure, I have proposed that quotes are inserted as N’ constituents that dominate an XP, properly annotated for achieving the aforementioned f-structure. Also, it has been proposed that quotes are given a “default” number (or gender), modelled as defining equation.

Of course, much more study is sure to come about this truly interesting and also amazing topic; quotation is the place where syntax can be somehow suspended for purposes as important as language teaching and social communication. It is granted that this proposed model probably does not explain all quotational phenomena, but I hope it can help to trigger future research on it and, hopefully, inspire better and more general theories that make us understand more deeply one of the most intriguing devices of human language.

Bibliography

- Abney, S. P. (1987). *The English noun phrase in its sentential aspect*. PhD dissertation, MIT.
- Bonami, O. and Godard, D. (2005). On the syntax of direct quotation in French. In Müller, S. (Ed.), *Proceedings of the HPSG08 Conference. NICT, Keihanna, Japan*. Available at:
<http://csli-publications.stanford.edu/HPSG/9/hpsg08.pdf>
- Cappelen, H. and Lepore, E. (2007). *Language turned on itself: the semantics and pragmatics of metalinguistic discourse*. Oxford: Oxford University Press.
- Carnap, R. (1956). *Meaning and necessity: a study in semantics and modal logic* (2nd ed.). Chicago: The University of Chicago Press.
- Casado Velarde, M. (2010). Discurso y creación léxica: delocutivos y decitativos en español. *Revista de investigación lingüística*, 13, 65–85.
- Chomsky, N. (1995). *The minimalist program*. Cambridge MA: The MIT Press
- Clark, H. H. and Gerrig, R. J. (1990). Quotations as demonstrations. *Language*, 66(4), 764–805.
- Collins, C. and Branigan, P. (1997). Quotative inversion. *Natural language and linguistic theory*, 15, 1–41.
- Dalrymple, M. (2001). *Lexical-functional grammar*. San Diego CA: Academic Press.
- Davidson, D. (1984). Quotation. *Theory and decision*, 11(1), 27–40.
- De Brabanter, P. (2005). The impact of autonymy on the lexicon. *Word: journal of the International Linguistic Association*, 52(2), 171–200.

- De Saussure, F. (1984). *Cours de linguistique générale*. T. de Mauro, (Ed.). Paris: Payot. (Originally published in 1916).
- Pollard, C. and Sag, I. (1994). *Head-driven phrase structure grammar*. Stanford, Chicago: CSLI Publications and University of Chicago Press.
- Quine, W. V. O. (1981). *Mathematical logic* (rev. ed.). Cambridge MA: Harvard University Press.
- Recanati, F. (2001). Open quotation. *Mind*, 110(439), 637–687.
- Rey-Debove, J. (1997). *Le métalangage* (2nd ed.). Paris: Armand-Colin.
- Suñer, M. (2000). The syntax of direct quotes with special reference to Spanish and English. *Natural language and linguistic theory*, 18, 525–578.
- Stainton, R. (1999). Remarks on the syntax and semantics of mixed quotation. In Murasugi, K. and Stainton, R. (Eds.), *Philosophy and linguistics* (pp. 259–278). Boulder CO: Westview.
- Tarski, A. (1983). The concept of truth in formalized languages. In Tarski, A. *Logic, semantics and metamathematics: papers from 1923 to 1938* (2nd ed.; pp. 152–278). J. Corcoran, (Ed.), J. H. Woodger, (Trans.). Indianapolis: Hackett Publishing. (Originally published in 1933).