1

Introduction

1.1 The issue

In this book I investigate the relationship between syntax and context. In particular, I propose that in the syntactic representation of the sentence a syntactic *layer*—i.e., a sequence of positions functionally related—is especially devoted to play such a role at the interface. I identify this set of positions with the Complementizer-layer and argue that the temporal—and arguably spatial as well—coordinates of the speaker are represented in its left-most projection, which I dub here *C-speaker*.

It is widely recognized that the meaning of a sentence requires a 'context' to be computed. This is a very general phenomenon and in particular it concerns the items called *indexicals*, i.e., 'linguistic expressions whose meaning remains stable while their reference shifts from utterance to utterance'.[1] Pronouns such as *I* and *you* are the prototypical indexical items. Other examples include demonstratives, such as *this* or *that*, temporal and spatial locutions, such as *this room,* or *last month*, *here, now, yesterday, tomorrow*, etc. All these items can be assigned a reference only if we know who is talking, when, and where.

The literature in philosophy and in semantics about these issues is very rich and a discussion of its content would be far beyond the

---

[1] Cappelen and Lepore (2002: 271). Recently a definition of *indexical* such as the one by Cappelen and Lepore has been challenged by scholars like Schlenker (2003), who argue that indexicals can indeed change their meaning—i.e., adopting Kaplan's (1989) terminology they can be *monsters*—at least in certain languages. I will not consider this issue here, because it seems to me it would lie outside a discussion concerning the *syntax* of indexicality, but I will keep to a view according to which there are no *monsters*, following Kaplan's tradition. See also Higginbotham (2003).

scope of this book. There are however some relevant considerations that can be discussed with respect to indexicality even taking a purely syntactic perspective, like the one taken in this book.

There is an obvious question that we might ask as soon as we consider indexical phenomena: how does the syntax interact with the context? The first naive answer might be: it does not, there is no real interaction. The syntax computes structures and interactions among *constituents*, and then, *after*—metaphorically, not necessarily temporally—the syntactic computation has been done, the structure is interpreted and only at that point does the context come into play. Under this perspective, the context would not be a necessary component of syntax.

In this light, *this book* is a noun phrase and does not differ from any other noun phrase, as far as the rules of grammar are concerned. The fact that it includes a demonstrative does not concern the syntax, but some other module of language.

Even if this perspective might certainly be the right one, there are indexical components in language that cannot be as easily put aside, and *outside* syntax. What I mean is not that a *specific* context should be represented on the syntax. The main thesis of this book is that in the syntax there is a position—or better to say a *layer*, i.e., a set of contiguous positions—devoted to the interface between the syntax and the extra-sentential context, whatever it might be.

Note that the studies developed so far on this issue were already on this path. Rizzi's (1997) seminal work on split-Comp already implicitly shows that the left periphery of the clause is projected out of functional items which typically play a discourse role: *Topic*, signalling old information, *Focus*, signalling new information, and the Complementizer positions named *Force* and *Fin(ite)*, also playing a role in the contextual interface.

In this book I propose a precise hypothesis to this extent, claiming that the left-most peripheral position is specifically devoted to the representation of the speaker's temporal and spatial coordinates.

In order to exemplify, consider for instance the temporal interpretation of sentences. *John ate a sandwich* is a well-formed sentence in English, as far as its grammar is concerned, and it includes an

indexical item, namely a past tense. The eating event is located in the past with respect to the utterance event, i.e., with respect to the speaker's temporal location. The same holds for a present tense: *John is eating an apple*, or for a future: *John will eat an apple*. In this case, indexicality becomes part of verbal morphology. This however might be considered just an accident, due to the peculiar morpho-syntactic structure of the language in question. In certain languages, such as Navajo, Chinese, and Haitian Creole, for instance, there are no morphemes devoted to the expression of tense, and still, the sentences are interpreted as past, present, or future, much as they are in Italian and English.[2]

The crucial point relevant to this discussion is constituted by the temporal location of the embedded eventuality in sentences such as example (1):

(1) John said Mary is pregnant

In languages such as English and Italian, this sentence has a peculiar interpretation—see Abush (1997)—called *Double Access Reading*. In order for the sentence to be felicitous, the pregnancy of Mary must hold both at the time John spoke about it and at the time the speaker utters the sentence. In Chapter 2 I will consider this phenomenon, with further details. Here, I would like to point out that in sentence (1) both the main verbal form and the embedded one are interpreted with respect to the context as defined by the speaker's temporal location—i.e., *both* verbal forms have an index-ical component.

Note that this consideration is not trivial, in that it cannot be simply said, as a general rule, that the embedded verbal form is inter-preted *as if it were in isolation*, i.e., simply with respect to the speaker's temporal location. Consider, for instance, the following case:

(2) John said Mary ate an apple

---

[2] On Navajo and Chinese, see Smith (1997, 2007) and Smith and Erbaugh (2005). On Haitian Creole aspect and tense, see Giorgi and Pianesi (2001a). The temporal interpre-tation is considered as derivative with respect to the aspectual one. For a comparative discussion see Giorgi (2008).

In sentence (2) the eating must precede the saying. That is, by means of (2) the speaker reports the following sentence by John:

(3)   'Mary ate an apple'

Whereas she cannot be reporting the following one:[3]

(4)   'Mary will eat an apple'

Even if the eating event located in the future by John could lie in the past with respect to the speaker, it is not possible to report (4) by means of (2). For instance, if John utters sentence (4) on 3 December and Mary indeed eats the apple on the 4th, I still cannot report John's saying on the 5th by means of (2). The embedded event must be temporally located with respect to the main one. It cannot be interpreted as if it were a past form in isolation.

This is true with respect to sentence (1) as well: the embedded eventuality must hold *now*, *and* it must hold at the time John said it. The indexical interpretation of the present verbal form is not enough to yield the correct interpretation. Therefore, in a language like English, on one hand it is necessary to hypothesize that the embedded verbal form has, or at least can have, an indexical component as in example (1) above, on the other this is not sufficient to obtain the correct interpretation of an embedded form.

The interpretation assigned to sentence (1) in English, and Italian as well, as in many other languages, however, is not a universal fact. The same sentence, with an embedded present tense, in a language such as Romanian—Russian, Japanese, Chinese, etc.—does not have this interpretation. Consider for instance the following Romanian examples:[4]

---

[3] In this case, the speaker has to use, both in English and Italian, the so-called future-in-the-past:

i.   John said Mary **would eat** an apple

I will discuss this verbal form in Chapter 4, section 7.

(5) Maria e insarcinata.
Maria is(PRES IND) pregnant

(6) (Acum 2 ani) Gianni a spus ca Maria e insarcinata.
Two years ago John said that Maria is(PRES IND) pregnant

The present tense is the form used in main sentences to express simultaneity with the utterance time. But in Romanian, the equivalent of sentence (1), i.e., (6), has the same meaning as sentence (7) in English:

(7) (two years ago) John said Mary was pregnant

In sentence (7), as in the Romanian example in (6) above, Mary's pregnancy holds at the time of the saying, but, contrasting with (1), the pregnancy does not have to hold at utterance time. This is shown by the fact it is possible to add the temporal specification *two years ago*, which is totally incompatible with an embedded present tense in English:

(8) *Two years ago, John said that Mary is pregnant

Both in English and Romanian a present tense in a main clause is interpreted indexically, but in an embedded clause the indexical component disappears in Romanian, whereas it is retained in English. Why is there such a cross-linguistic difference? How is it possible to capture it?

Let's go back to the naive hypothesis given above, the one according to which the syntax and the context do not talk to each other, i.e., the indexical component comes into play in a module of language separated from the syntax. To account for the phenomena just presented, we would be compelled however to endow this non-syntactic, indexical, module with syntactic notions, at least with notions such as *main* and *embedded clause*. For instance the indexical module of English would contain the following rule:[5]

(9) If the verbal form of the embedded clause is a present tense, then the embedded eventuality must be located with respect to the speaker's temporal location, as it is in isolation.

---

[5] Note that as I briefly mentioned above, rule (9) is necessary, but is not enough to account for the interpretation of the embedded verbal form in English, given that *is* must also be interpreted as expressing simultaneity with the main event of saying.

Whereas in Romanian and Japanese rule (9) would be absent, and replaced by the following:

(10)   If a present verbal form appears in a main clause, then the eventuality must be located with respect to the speaker's temporal location. If the present tense is associated with a verbal form appearing in an embedded clause, then the eventuality is not located with respect to the speaker's temporal location.

Rules such as (9) and (10) are necessary, otherwise the presence or lack of the Double Access interpretation associated with (1) could not be captured.

It is possible in this respect to observe two important facts. The first is that the existence of rule (9) vs. (10) seems totally unjustified and arbitrary. Why should languages differ and why should they differ in exactly that way?

As for the second fact, as I remarked above, in these rules it is necessary to refer to syntactic notions, such as *main* and *embedded* clause. It therefore appears impossible to account for the interactions of sentences with the context *a priori*, without resorting to a syntactic 'level'. But, then, if a syntactic analysis is necessary, *exactly* what is the role of syntax with respect to phenomena such as the temporal interpretation of sentences?[6]

And so we are back to the original question, which can be addressed at this point with the basic remark that there is after all a relationship between syntax and indexicality and that *how* this relationship is established is an interesting issue of investigation. This is the topic of this book.

The issue is relevant both with respect to the general topic of the architecture of syntax and from the point of view of a cross-linguistic analysis. I will largely consider the first aspect, with some reference to cross-linguistic differences illustrated by a comparison between English and Italian, and occasionally Chinese and Romanian.

---

[6] In Abush's (1997) terminology—see also Schlenker (2004)—one might say that in English the embedded present tense is *de re*, whereas in Romanian it is not. Independently of a discussion about the *pro* and *contra* of this proposal, let me point out that this idea would not help in clarifying the issue concerning the relationships between syntax and indexicality. See also Chapter 2 for further details.

1.2 **The proposal**

The proposal I argue for in this book is the following:

(11) There is a syntactic position in the left-most periphery of the clause, and precisely in the Complementizer-layer, that encodes the temporal— and presumably spatial as well—coordinates of the speaker.

This position contains the *features* identifying the utterance event and exhibits different properties according to its syntactic environment— as is usually the case with syntactic phenomena—so that it is possible to account for intra- and cross-linguistic differences. I will illustrate theoretical and empirical arguments to this end.

Similar proposals have already emerged occasionally in the literature, in particular, but not exclusively, in the literature on temporal phenomena, for example Giorgi and Pianesi (2001a), Bianchi (2003, 2006), and Sigurðsson (2004, 2007). My proposal is a development of the one by Giorgi and Pianesi—even if there are several important differences—and differs in relevant ways from Bianchi's and Sigurdsson's, as will be discussed later in the book.

Perhaps the main distinguishing feature of the account I propose here concerns the relevance attributed to interpretive facts. As in Giorgi and Pianesi (1997) and subsequent work, I claim that the temporal interpretation of a sentence—and perhaps other interpretive aspects as well—is read off the syntax. Therefore, a difference in the temporal interpretation of the sentence *is* a difference in the syntactic structure. This hypothesis, though it may obviously be wrong, seems to me to have a strong heuristic power. Moreover, it is clearly the simplest possible starting point: there is no *deus ex machina* creating one particular interpretation instead of another. All we have is a syntactic structure and a context. The interpretation of the sentence arises from these components and nothing else.

In this introduction, and in this book, I am mostly talking about the *temporal* interpretation of clauses and about the necessity of hypothesizing the presence of the speaker's *temporal* location in the syntax—precisely in the left-most position of the C-layer. The natural questions therefore might be: are there other phenomena, besides

those connected to the temporal interpretation, which might be relevant to the purpose of the hypothesis developed here? The answer is 'yes', and I briefly consider the distribution of long-distance anaphors in a tenseless language such as Chinese, in Chapter 4. I will show that the hypothesis that in Italian and English takes care of the temporal interpretation of embedded clauses, in Chinese contributes to explaining the distribution of anaphoric items such as *ziji* (self).[7] Besides this brief remark, however, I only consider questions relating to the temporal interpretation. The matter of whether the present hypothesis might be relevant in other domains as well is left for further research.

The other issue, namely, whether the *spatial* location of the speaker is represented in the C-layer, together with the *temporal* one—*hic et nunc*—is also left vague. I think there is evidence that this is indeed the case, as might be expected. For instance, in Halkomelem Salish, as analysed by Ritter and Wiltschko (2004, 2005, 2008), it seems that the temporal interpretation is based on the speaker's *spatial* coordinate and not on her temporal one. The relevance of the speaker's spatial location also emerges in the analysis of some phenomena concerning Free Indirect Discourse, discussed in Chapter 6. In general, however, I limit my remarks to the speaker's temporal location.

1.3 The background

The theoretical background of this work is constituted by the analysis of temporal relations provided in the generative framework of the 1990s, together with the minimalist proposals by Chomsky (1995, 2001, 2005).

I assume, following the seminal intuitions by Zagona (1988) and Stowell (1996), and much in the spirit of Giorgi and Pianesi (1997), that the temporal relations are represented in the syntax. Even if I do not propose here a multi-layered representation like the one

---

[7] See also Giorgi (2006, 2007).

made explicit in Stowell (1996), I follow his, and Zagona's, basic intuition that tenses are relational predicates, as also proposed in Higginbotham (2002, 2004, 2006).[8] According to the proposal adopted here, therefore, tenses are relational predicates represented in T. In languages distinguishing indicative and subjunctive, such as Italian, it is possible to capture the difference between the two by hypothesizing that the subjunctive is not a relational tense, in that it does not express a *tense* at all. The idea I will develop is that the subjunctive realizes an *agreement* relation with the superordinate form, with some interesting exceptions, however, which will prove useful in making the whole proposal more precise.

The core of the hypothesis argued for in this book is that, when appearing in subordinate clauses, relational tenses—in Italian in particular, and in DAR languages in general—require a special Complementizer, in that the eventuality, besides requiring anchoring to the superordinate event, must be located with respect to the speaker's temporal coordinate. This Complementizer has peculiar syntactic properties, which distinguish it from the Complementizer introducing a subjunctive, i.e., non-relational verbal forms. As I said above, it encodes the speaker's temporal and spatial coordinates, permitting—and requiring—evaluation of the verbal form with respect to the speech event even in embedded clauses.

The presence of the speaker-related features in the left-most position of the C-layer shows up not only indirectly, when analysing Double Access Reading phenomena, but also directly, when considering the distribution of certain first-person verbal forms, such as *credo* (I think), which—under specific syntactic conditions— will be analysed as epistemic heads. These heads will be shown, in certain clauses, to occupy exactly the speaker-related position in the Complementizer-layer I hypothesize.

Technically, the subordinate clauses not requiring the Complementizer endowed with the speaker-related features will simply

---

[8] For further developments of the same proposal, see also Zagona (1994, 1995, and 1999). For a slightly different development, see Guéron (1993, 2004) and Guéron and Hoekstra (1995).

be accounted for by means of applications of *Merge* and *Agree*. Conversely, the clauses endowed with the speaker-related position will also internally merge T in C, much in the spirit of Pesetsky and Torrego (2001), so that the correct interpretation will be triggered.[9]

In the literature about temporal phenomena it is possible to find reference to systematic exceptions to the general framework sketched by the various scholars. For instance, in dependence of a superordinate future verbal form, and in 'special' contexts, such as those found in narration, it is well known that the distribution of tenses is not the one normally expected.

In this book I address several of these 'difficult' issues. In particular, I will consider the distribution of the Italian imperfect indicative, the properties of the so-called future-in-the-past, the dependencies from a main future, and the properties of Free Indirect Discourse—a particular literary style, which looks like a counter-example to every possible generalization in terms of syntactic structure and DAR.

I will show that the framework I propose can account for these facts quite elegantly, and in a natural way. The same general picture is adopted for the simple cases discussed in Chapters 2–4: two possible options are available for a speaker of English or Italian; either the C-layer includes the speaker-related position or it does not. The temporal location of the speaker, however, might be affected by several factors, for instance the presence of certain operators or specific fictional devices. These factors yield the 'anomalies' mentioned above, but the general framework does not need to be altered. In particular, Chapter 6, on Free Indirect Discourse, provides a very strong argument in favour of my hypothesis. The leading idea is that the speaker's coordinates are syntactically represented, and, as such, might be subject to syntactic manipulation.

As I mentioned above, I will briefly sketch an analysis of the speaker's projection in Chinese, a language not exhibiting the same morphological complexity observed in Italian. A final important question must be raised, however: What can be said about languages

---

[9] For further discussion, see also Pesetsky and Torrego (2004a, 2004b, and 2006).

such as Romanian, which are as rich in verbal morphology as Italian, but do not exhibit any DAR? I will leave this issue for further research and offer here only some speculations with respect to it.[10]

1.4 The organization of this book

There are six further chapters in this book. In Chapter 2 I address the differences between indicative subordinate contexts and subjunctive ones, mostly with reference to complement clauses. I show that many interpretive and (purely) syntactic phenomena can be explained by means of a simple hypothesis: in some cases the left-most position of the Complementizer-layer is endowed with the speaker's temporal and spatial coordinates.

In Chapter 3 I consider the distribution of first-person, subjectless verbal forms such as *credo* (I think) in Italian, and show that there is evidence for claiming that the hypothesized speaker-related position is made visible by these items under certain syntactic conditions.

In Chapter 4 I examine the properties of the Italian imperfect—not in general, but only as far as the hypothesis developed in this work is concerned—and of the future-in-the-past. I show that in spite of their peculiar behaviour, their distribution is expected under the present proposal. In this chapter I also show that in a tenseless language such as Chinese, the properties of the distribution of Long Distance Anaphors, which I have discussed elsewhere, can be taken as evidence in favour of the analysis developed here.

In Chapter 5 I discuss the dependencies from a future verbal form and in Chapter 6 Free Indirect Discourse examples. Finally, in Chapter 7, I add a few closing remarks.

---

[10] I proposed some reflections about this point in Giorgi (2007). Note however that the empirical evidence is quite complex, and only native speakers can actually deal with it. This is the main motivation for not addressing the issue in this monograph.