The ASIS enterprise: a view on the construction of a syntactic atlas for the Northern Italian dialects
Paola Benincà and Cecilia Poletto
Università di Padova, Università di Venezia

Abstract:
In this article we intend to illustrate how the ASIS (Atlante Sintattico dell’Italia Settentrionale "Syntactic Atlas of Northern Italy") project has been created and developed. We discuss the theoretical, empirical and practical problems that we encountered working on such an enterprise, and the choices we made in order to solve them. We have created a layered methodology, which has proved useful in gathering more and more detailed data in an interplay between theoretical analysis and field work. As no method is perfect, we outline here also some of the problematic aspects of our project.

Introduction¹

There are several theoretical reasons why we decided to create a dialectal syntactic Atlas. In the generative perspective the external language is the manifestation of the internal language, and investigating the minimal variation, in a certain sense, we look at the product to discover how the “machine” is made. Variation is, from our point of view, a natural state for language; due to the acquisition process, which can vary in minimal ways, each person has a slightly different setting of formal (or lexical) properties of functional elements. Therefore, it is straightforward to assume that there exists a very high number of minimally different languages. Using a metaphor, we could assimilate microvariation to the differences found in the DNA of a family of bacteria, all causing the same disease, but still minimally distinct from one another. Nevertheless, the method of linguistic research forces us to make an abstraction and consider a speaker not just the representative of himself, of a single individual grammar, but of a set of minimally different grammars constituting a dialect. If dialectal variation is the natural way in which language manifests itself, this means that the

¹ Although this article is signed only by the two initiators of the project, we would like to underline that the research group behind our enterprise is made up by many more people, such as the other two co-founders, Richard Kayne and Laura Vanelli, and the supporters and collaborators: Nicoletta Penello, Chiara Polo, Nicola Munaro, Federico Damonte, Jacopo Garzonio, Andrea Padovan, Diego Pescarini, Sabrina Rasom, Marinela Sotiri, Laura Sgarioto, Barbara Patruno, Federico Ghegin. An invaluable help has come from our colleagues Raffaella Zanuttini, Christina Tortora, Sandra Paoli, Mair Parry, Massimo Vai, Piera Rizzolatti

Nordlyd 33:00-00, © Paola Benincà and Cecilia Poletto 2005
Scandinavian Dialect Syntax 2005
Edited by Øystein Alexander Vangsnes, Kristine Bentzen and Peter Svenonius
choices made inside a single grammar follow a rational path which can reveal its internal structure, but also language variation can itself be conceived as the result of processes that follow logical patterns, and produce only grammars that are submitted to general principles, thus revealing the universal grammar lying behind it. Moreover, dialects are interesting because in general (depending on the extra-linguistic situation) there is no formal pressure towards a “non natural system”, as is often the case for standard languages, which are forced into rules by grammarians trying to design a coherent rational system, but being unaware of the real mechanisms through which language changes and vary.

Dialectology constitutes a privileged observation point for determining language variation, just because it studies minimally different systems. As such, it is as close as possible to a scientific experiment where variables (intended as differences in grammar) are controlled and few independent factors interfere with the study of a single grammatical property. If the sample is wide and detailed enough, we can conclude that where a theoretically possible choice is absent, it is in all probability not possible for principled reasons.

Related to the fact that dialectal variation maintains certain basic features of a grammatical system as a constant and variation seems to be concentrated on specific domains, there is a potentially far-reaching question concerning variation in general: we might find that dialectal variation is different from typological variation in a very deep sense: no “macroparameters” are ever changed in microvariation. Basic choices like constituent order, or inflectional versus agglutinative systems, never appear affected when we compare dialects. The differences found here are always of a subtler kind, so to say; this can shed light on the question whether there are or not clusters of properties (parameters) that (always or frequently) go together; this means that the choice of +A would induce the child to hypothesize that the language also has +B and +C, and −B and −C are either impossible or somehow marked and require strong positive evidence to be triggered.

This kind of theoretical reasons lead us to think that a syntactic dialectal Atlas is a theoretically revealing enterprise.

2. The state of the art in 1990

Several reasons drove our choice of the area of investigation. The first is that Northern Italy is linguistically a homogeneous domain, because the languages spoken in these regions share important grammatical and lexical properties and as such constitute a proper subset inside the domain of
Romance languages that single them out with respect to other Italian dialects. They already had uniform syntactic properties in the medieval period, and shared the same system of verb second and pro drop licensing, which was later lost; they then developed subject clitics and subject clitic inversion in interrogatives in a way that is very similar to the one found in standard French. The parallel diachronic path that main questions follow in French and NIDs is attested for example by a number of constructions, which are still developing, going from cleft sentences to *wh-* in situ.

There were already a number of descriptive studies and texts, mainly reporting lexical, phonological and morphological differences, which also included non systematic observations concerning word order as well, and, although no systematic enterprise on the syntax of these dialects had ever been conceived, there was enough material to provide us with a “launching base” to start with such a project. A systematic study on subject clitics in 30 dialects had already been made by means of a questionnaire by Renzi and Vanelli (1983) and this was the first phenomenon we took into account when preparing our first general questionnaire.

A project of this type requires a number of choices on the methodological side, which in our case have been driven from the theoretical point of view sketched above in the introduction. A theoretical investigation on syntactic variation requires:

- Comparable data (minimal pairs if possible).
- Ungrammatical data.
- Access to optional phenomena.
- Access to interface with phonology and morphology.
- Access to the etymology of a word.
- Access to the interface with semantics and pragmatics.

Comparable and ungrammatical data are required for any generative analysis and the access to the various interfaces and etymology (which often sheds light on the diachronic evolution of the system) are also required in order to be able to sort out whether a given phenomenon belongs to syntax or to some other component of grammar. Syntactic phenomena can be related to phonological phenomena, in the sense that they condition the occurrence of a phonological rule; they can be related to semantics, because some syntactic phenomena only occur with a special semantics. In studying a system, it is not possible to ignore the relation among its component. This type of data are best obtained by a layered methodology of the type we will illustrate in the following section.

3. The layered methodology
One of the first problems the syntactician is confronted with is not only that of gathering comparable data (see reference in footnote 2 on this), but also of discovering new interesting phenomena.

Our first inquiry was conducted by means of a written questionnaire which contained about 100 sentences. They were primarily set for testing subject clitics in different syntactic domains; we considered negative, interrogative, exclamative and relative clauses, main and auxiliary verbs and, given that subject clitics can be combined with an overt subject, we inserted sentences with various types of subjects (wh- traces, tonic pronouns, definite and indefinite DPs, and quantifiers, in pre and postverbal position. Note that, although we started with a single phenomenon, our sample was conceived in order to indirectly gather data on the following phenomena as well:

- Sentence typing
- Syntax and morphology of Wh-items
- Negation syntax
- Quantifiers lexicon and syntax
- The system of verbal tense and mood

Therefore, through a questionnaire originally designed to systematically test variation of a single phenomenon, we were able to discover a number of important new phenomena such as the following:

- Sentential particles and clause typing
- Wh- in situ and wh- doubling
- Do-support in main interrogatives
- Different types of pre and postverbal negative markers

Two further general questionnaires of the same type were prepared with the same method; they contained sentences designed to test the following phenomena:

- Object clitics and clitic clusters
- Auxiliary selection
- Modals and modality
- Complementation

Once a number of new phenomena had been discovered, they still had to be investigated on their own; in the first questionnaire there was obviously no systematic research on them, since it was impossible for practical reasons to test all the variables for all the fields of inquiry listed above, which could only be just touched as in a sort of preliminary log. At the second stage of our inquiry we prepared special questionnaires designed to test each phenomenon, trying to find out what its “domain of variation” was, i.e. to determine which factors influence its occurrence. For instance, in the case of do-support discovered in the Lombard dialect of Monno, we found that
the phenomenon is only possible in main interrogative clauses (not in embedded ones), when the wh-items is not the subject (in which case the verb does not move), if the verb is not an auxiliary or a modal. Given that the phenomenon is strikingly similar to English do-support, it was relatively easy to find the factors that influence its occurrence and to underline the differences (first of all, differently from English, no do-support is found emphatic contexts, negative clauses and VP-ellipsis constructions, contexts that in English are considered a preliminary stage towards the development of the 'support status' of the verb: see Benincà and Poletto (2004)). In other cases, as with sentential particles, the domain of variation was much harder to determine, and required at least three different stages of inquiry with predetermined contexts, as the phenomenon is semantically conditioned (see Poletto and Zanuttini (2003)).

We had two types of special questionnaires, either concentrated on a single construction (interrogatives, imperatives, exclamatives, negatives), or on phenomena that are only found in some dialects.

Special questionnaire do not only serve to determine the domain of variation of a single phenomenon, but also the variation of this phenomenon according to the dialect and to semantic differences related to different structures.

As mentioned above, the first general questionnaire was a written one, and was preliminarily tested with a number of linguists, who served as first informants and provided us with feedback, helping us to refine the sentences (often a lexical choice is not the right one, for instance, or what is a transitive verb in a dialect is intransitive in another, so that you'll not get an object clitic through that sentence, etc.). Some constructions are for instance not at all present in several dialects, therefore they were discarded, because this rendered comparability of data less complete. In other cases sentences were potentially ambiguous, and they were modified according to the comments we received.

The special questionnaires have often been performed orally, especially those in which a special semantic or pragmatic context was supposed to be involved. Written questionnaires are perfectly adequate to investigate obligatory phenomena, but optional phenomena do not always surface in written questionnaires, or they may surface in a non systematic way, so that drawing empirical generalizations on their distribution is reduced unless optionality is resolved.

2 On the advantages and the problems connected to our choice see Cornips and Poletto (2004).
In many cases more than one stage of field work was required, as the ungrammatical sentences to test evidently depend on the theoretical hypotheses that are made about a phenomenon. Therefore, there has been a constant feedback between data and analysis, which has often required interviewing the same informant several times with more and more detailed contexts.

At this stage, the best informants were selected to serve as actual collaborators: they helped us to find new differences that were not included in our questionnaires; to define special semantic contexts in which a syntactic phenomenon occurs (this happened, for example, with sentential particles); to interpret data from other informants in order to solve ambiguities or apparent contradictions. The role of these selected informants has been crucial to discover new phenomena, as informants who are bilingual can easily find out whether some constructions also exist in standard Italian or not. This aspect has been particularly important when we decided to widen our area of investigation and approach the central and southern Italian dialects. While both of us and many of the members of the group are native speakers of one or another northern dialect, we lacked the basic intuitions when approaching central and southern dialectal systems.

In the next section we concentrate on some problems we had in revising the general questionnaire and use it for Southern Italian dialects.

4. Good luck and bad luck

As in other enterprises, good luck and bad luck have been crucial ingredients of the ASIS project as well. In this section we examine a couple of concrete cases we have been dealing with when working on the first phase of the project on Southern Italian dialects. In the first case, the questionnaire resulted deficient, because the questions we had prepared were of no use in order to show the phenomenon we were looking for; in the second case, on the contrary, we were lucky enough to find an unexpected construction worth investigating.

The general questionnaire for Southern dialects was prepared after a survey of the traditional and generative literature on these areas; a number of phenomena were selected to be tested and sentences were prepared in which the relevant phenomena were supposed to occur. The phenomena selected are the following:

- Prepositional accusatives
- Information Focus in Comp/Aux to C
- Complementation
• Verbal system: future tense, and mood
• Auxiliaries and modals
• Clitic clusters
• Quantifiers and negation
• Sentence typing

Prepositional accusative is reported in the literature for the southern area: here is an example:

(1) Salutami a suaru
    Greet-me to your sister
    Greet your sister for me

Given that it is always necessary to start with hypotheses concerning the factors that rule the distribution of each phenomenon, we decided to select the following factors, on the basis of the discussion found in the literature concerning the same phenomenon in Spanish:
• +/- animate/human
• +/- bare quantifier
• +/- quantified NP
• +/- definite (specific)

Prepositional accusatives are also reported in the literature concerning southern Italy and have been analyzed in detail by Ledgeway (2003) for Neapolitan. In fact in that area they occur quite often in the sample, but in other areas, as for instance Sardinia, they seem to be quite rare. Many of the general questionnaires for the Sardinian area did not report the phenomenon, which was only attested in very few sentences, all containing a kinship noun. The first hypothesis we could make on the basis of available data was that the phenomenon is not present in some dialects, and therefore exclude areas like Sardinia from the special questionnaire investigating the prepositional accusative in detail. After an oral interview with a native speaker coming from Baunei (in the North-Eastern part of the island), we found out that prepositional accusative indeed exists. The reason why it was represented only by a couple of examples in our data is that in Sardinian prepositional accusative is sensitive to variables that have not been considered in the first general survey, so the factors listed above, based on the literature on Spanish and Neapolitan, do not trigger prepositional accusative, which in Sardinian is not sensitive to the [+human+] or [+specific]
features, but to the presence of a relational thematic role, as the following example shows:

(2)  appu idiu (a) calleddu  
(1) have seen (to) puppet  
 I have seen a puppet

In (2) the presence of the preposition induces the reading in which the dog belongs to someone, so there is an implicit relational theta role only when the preposition is there. That this is correct is shown by the fact that intrinsically relational DPs like kinship nouns require the P

(3)  appu idiu *(a) mamma  
(1) have seen (to) mum  
 I have seen my mother

This shows that before drawing the conclusion that Sardinian does not have prepositional accusative on the basis of the first survey, it is better to check with an informant in order to be sure that the phenomenon does not appear simply because it obeys different requirements. Therefore, when creating a questionnaire we have to remember that the variables selected are not necessarily valid for all the dialects.

In other cases, investigating a given phenomenon provides us with new and interesting clues about its analysis or about general topics discussed in the syntactic literature. The general questionnaire on Southern Italian dialects contained several examples with negative quantifiers in various argument positions and negative adverbials like mai ‘never’ and più ‘anymore’: in the dialect of Lizzano (a Salentino dialect in the province of Taranto) postverbal negative quantifiers and more generally negative polarity items do not only trigger the presence of a preverbal negative marker (as is generally the case for negative concord languages) but also cooccur with a locative clitic nci, which in the sample is always present, (therefore, we can hypothesize that it is obligatory) when the quantifier is postverbal but is not when the quantifier is preverbal.

(4)  a Non nci spicciava mai di turmeri  
Not loc.cl stopped never of sleeping  
“She continued sleeping”  
b ca nun nc’è statu vistu niscunu  
that not loc.cl has been seen nobody  
“that nobody was seen”
c  non nc’eti anima viva  
Not loc.cl is living soul  
“There is not a living soul”

d  Non nci mi passa mancu pi la capu.  
Not loc.cl me goes –not.even through the head  
“I do not even think about it”

e  no’ nci l’hogghiu vistu chiù  
not loc.cl it have seen anymore  
“I never saw it again”

(5)  Nu sacciu ci è arrivatu  
Not know who is arrived  
‘I do know who has arrived’

\( Nci \) does not occur in simple negative clauses, therefore it is not required by the negative marker itself, except in two cases:

(6)a  no nci mi ne ste scappu  
Not loc.cl me cl stay go  
“I am not going away”

No nci l’ avu spittata  
Not loc.cl her have waited  
“We did not wait for her”

In both cases there seems to be an implicit element triggering the presence of the clitic: the first example was the translation of the standard Italian postverbal negative marker \( mica \), which is a presuppositional marker in its nature. This dialect (as in general the whole Southern area) do not use an overt postverbal negative marker, but seems to render the presuppositional value with the simple presence of the clitic. We can hypothesize that the structure indeed contains an empty postverbal negative marker similar to standard Italian \( mica \). The second example has to be interpreted as suggested by the Italian stimulus sentence “we did not wait for her \( anymore \)”. Notice that (4c) includes the negative polarity item “living soul” and thus suggests that it is not only N-words that require the presence of the locative clitic. Therefore, the phenomenon of the locative clitic insertion is interesting for the following theoretical reasons: a) it seems to treat negative polarity items and N-words as a unique class; therefore this might be an argument in favour of these two types of elements belonging to the same class (contrary to the analysis that treats N-words like universal quantifiers) b) it seems to suggest that negative
concord is a syntactic phenomenon, not just a semantic one, c) once its analysis is precise enough it might be used as a test to verify whether a given structure contains a null negative element.

These hypotheses clearly have to be tested with a special questionnaire on this phenomenon. The new questionnaire has to include further cases of negative polarity items, to test whether the clitic is present with all negative polarity items, or some subset of them, it has to contain more cases of preverbal and postverbal N-word and negative adverbs, to test whether the clitic is obligatory with all N-words (and phrases) in all argumental positions; to include sentences with other object clitics, to test whether the locative clitic is located in the same position where it usually serves as a locative marker, it has to list different sentence types, in particular interrogative sentences, to test whether the phenomenon is sensitive to the presence of an operator different from negation.

5. On determining the “domain of variation”

Some comments on how we formulated our hypotheses on the domain of variation of single phenomena are in order. The first thing to do when preparing a questionnaire centered on a single phenomenon is to select the variables according to which the sample sentences have to be prepared. In some sense, preparing a questionnaire of this type means that we have to have a very detailed hypothesis - or alternative hypotheses - on how the phenomenon works, which is evidently not feasible. Often the first examples coming from the general questionnaire and descriptive grammars do not help us to figure out what the domain of variation of the phenomenon might be. However, given the amount of theoretical work that has been done in the syntax of Romance languages (and in other language domains as well) can help us; there are a number of common possible factors that are known to influence the distribution of several phenomena. For instance, if we are investigating negation, it is worthwhile to try with imperatives and other modal forms, as we know that negative elements are sensitive to modality in some languages; the position of negation with respect to the verb (including auxiliaries and modals) and to different adverbial classes will also have to be tested, as well as its possibility to occur in the CP domain.

So, although finding out the variation domain of a given phenomenon might seem a difficult enterprise, we have tools provided by previous and ongoing research. Below is a list of possible variables which are often relevant to a wide number of phenomena.
There are general variables that have to be investigated first when trying to figure out how a phenomenon works. Here I list some of them with examples for each:

- Sentence type

Sentence type is often a variable influencing phenomena located in the higher portion of the clause. The example that follows shows the influence of sentence type on the possibility of inversion of a subject clitic, which is never found in declarative clauses, (unless the enclitic has become a stable inflectional morpheme), while it is obligatory in some varieties in interrogatives:

(7)  
| a | El va casa Padova  
|   | He goes home  
| b | Va-lo casa?  
|   | Goes-he home?  
| c | *Valo casa  

Even phenomena that apparently have nothing to do with the left periphery can be sensitive to sentence type: the agreement pattern with postverbal subjects in relative and embedded interrogative clauses is different with respect to the agreement pattern found declarative clauses:

(8)  
| a | l'rya l pu'sti Monno  
|   | it comes the postman  
| b | l'salta zée le 'fée "CLsg cade giù le foglie  

(9)  
| a | i m a doman'da ngo la fus anda'da la Ma'ria  
|   | they me have asked where she were gone the Mary  
| b | la me'nestra ke la fa la tua mama l ε buna  
|   | the soup that she does the your mother it is good  

While in \(wh\)- contexts the subject clitic is a full agreeing form in person (cf. (9a, b), number and gender, the clitic occurring in declarative clauses is an expletive form (8a,b).

- Main-embedded clauses asymmetry

The distinction between main and embedded clauses is also typically influencing phenomena related to the CP layer. Once again subject clitic inversion is sensitive to this factor as well:
Another phenomenon which is sensitive to the main versus embedded distinction is the presence of sentential particles, which can only be found in main clauses. This fact can also provide a clue for the analysis of these elements, which must be somehow connected to the left periphery even if they sometimes happen to appear in final position.

The presence of auxiliary verbs can influence various phenomena; an example are the cleft versions of interrogative clauses; the cleft version is obligatory when the verb is in a simple form and optional when the inflected verb is a form of be (as in compound tenses of ergative verbs):

- Presence of auxiliaries

The syntactic or semantic class of the verb can influence phenomena located in the lower portion of the sentence structure. For instance, postverbal subjects are sensitive to the syntactic verbal class:

- Verb classes
This does not seem simply a distinction between unaccusative and unergative verbs: Tortora (1997) shows that the possibility of a postverbal subject cuts across the class of unaccusative verbs: non focussed postverbal subjects are only possible with inherently directed motion verbs. Moreover, inside the class of intransitives the presence of an implicit locative argument also licenses non focussed postverbal subjects

(17) a  %Ze ndà via me mama  (Only focussed subject)  
       Is gone away my mother 

          b  Ga telefonà me mama  (ok without focus)  
              Has phoned my mother 

The type of nominal elements used to test the sentences can also influence some phenomena. Clitic doubling is sensitive to the type of nominal element they double, and there exists an implicational scale among elements that can be doubled, which seems to be valid for both subject and dative doubling:

(18) a  If DPs are doubled in a given dialect, tonic pronouns are also doubled, 

          b  If QPs are doubled, both DPs and tonic pronouns are doubled, 

          c  If variables in wh-contexts as relative, interrogative and cleft structures are doubled, then doubling is always obligatory with all other types of subjects

(19) a  TI te magni sempre Venezia  
       YOU SCL eat always 

          b  *TI magni sempre Venezia  
              YOU eat always 

(20) a  Nane (el) magna Venezia  
       John (SCL) eats 

          b  Nisun (*el) magna nobody (SCL) eats 

(21) a  Nissun (*el) me capis Montesover (Trentino) 
       nobody (SCL) me understands 
       ‘Nobody understands me’ 

          b  El popo *(el) magna el pom Montesover 
              ‘The child SCL eats the apple’ 

(22) a  El fio el mangia l pom Milan 

This list is obviously not exhaustive. In general we can say that depending on the location of the phenomenon in the structure, one hypothesize which factors can possibly influence; however, as seen above, there are some phenomena apparently located at the right border of the sentence structure (such as sentential particles and postverbal subjects) which are sensitive to sentence type of main/embedded clauses. For the moment the ASIS project has completed special questionnaires on the following phenomena:

- Exclamatives
- Interrogatives (*wh*- *in situ*, do-support)
- Imperatives
- Negation
- V2 in main and embedded clauses
- Locative constructions
- Germanic varieties in contact with Romance
- New phenomena for NIDs:
  - Sentential particles
  - Topic-Focus distribution in the left periphery of the clause
  - Low quantifiers and adverbials
  - Clitic doubling and clitic positions
  - Auxiliary selection

6. An example of the layered methodology: the case of *wh*- *in situ*

The layered methodology does not stop at the second step with special questionnaires; the fact that we happen to find new phenomena by looking at others is a circular process which narrows down and focuses our...
perspective on microvariation more and more. An example of this stepwise system is provided by interrogative clauses. In the general questionnaire on NIDs, some cases of *wh- in situ* and some cases of *wh*-doubling were found.

(24) a A-lo fat che?
Has-he done what?
b Sa l’a fat cusè?
What he-has done what?

The special questionnaire on these phenomena was carried out with oral interviews only in the areas where *wh- in situ* and *wh*-doubling were attested. The variables\(^3\) selected for the questionnaire are the following:

- bare whs vs. non bare whs
- d-linked vs. non d-linked non bare whs
- main versus embedded clauses
- auxiliary vs. main verbs
- (presence of negation)

Among these variables, some were indeed relevant: the distinction between bare and non bare *wh*-items was confirmed; *wh- in situ* and *wh*-doubling appeared to be indeed sensitive to the main versus embedded context, as no *wh- in situ* or *wh*-doubling in found in embedded clauses (apart from some well specified cases).

Auxiliary selection did not show any relevance at all for the phenomenon we were investigating, but helped us to further define a new phenomenon in some Lombard dialects. Consider the next two sentences: in the first one the auxiliary *fa* ‘do’ appears before the infinitival form of the verb *magnà* ‘eat’.

(25) a ke fa-l majà?
what does-he eat?
‘What does he eat?’
b ke a-l majà?
what has-he eaten?
‘What has he eaten?’

We discovered in this way that a version of the well-known phenomenon of English do-support is also found in these dialect (as pointed out above), a

---

\(^3\) The last variable is in brackets because it was not tested systematically.
quite striking fact in the Romance domain. The presence of do support was then confirmed through a further test specifically designed to discover its distribution, (our empirical and theoretical results are reported in Benincà and Poletto (2004)).

The special questionnaire on \textit{wh- in situ} and \textit{wh-} doubling provided evidence in favor of the following descriptive generalizations:

- a If \textit{wh- in situ} is found with a single \textit{wh}-item, this \textit{wh}-item corresponds to "what".
- a' If \textit{wh-} doubling is found with a single \textit{wh}-item, this \textit{wh}-item corresponds to "what".
- b If a language allows \textit{wh- in situ} cooccurring with SCLI, the only \textit{wh}-items that can be left \textit{in situ} are those that can become clitics.
- b' If a language allows \textit{wh-} doubling cooccurring with SCLI, the only \textit{wh}-items that can be left \textit{in situ} are those that can become clitics.
- c If a language allows a \textit{wh- in situ} strategy, this is applied to \textit{wh}-phrases only if it applies to \textit{wh}-words.
- c' If a language allows a \textit{wh-} doubling strategy, this is applied to \textit{wh}-phrases only if it applies to \textit{wh}-words.
- d \textit{Wh-} doubling in embedded contexts is possible in the few cases in which the complementizer is not lexicalized.

Notice that it results that \textit{wh- in situ} and \textit{wh-} doubling have a parallel distribution and are subject to the same restrictions. Therefore, whatever the analysis of the two phenomena turns out to be, it has to take into account each of the descriptive generalizations above and the fact that \textit{wh-} doubling and \textit{wh- in situ} share the same properties. If we are able to formulate descriptive generalizations of this type on the basis of comparative dialectal data, they can drive our theoretical research and exclude some analysis which does not account for each of them and for the parallel between the two phenomena.

6. Some practical questions

After having outlined what the methodology we have developed for our syntactic atlas is, we will briefly outline the practical format used. The first small net of informants for Northern Italian dialects was essentially composed of linguists who are also mother tongue speakers of a dialect. Once the general questionnaire was checked, the first selection of non
linguist informants has been often made on the basis of the chance we had to reach native speakers in each area, so no age group was excluded, although data on age, education, the origin of the parents and grandparents were asked in order to control for changes or sociolinguistic variants in our data. The net of informants has been since then enlarged and at the same time layered in a group of collaborators as seen above. The same procedure is being adopted for Southern Italian dialects. Each questionnaire is evaluated during common working sessions of the whole research group, and although it might seem at first sight quite difficult to realize who can be a good informant, we were able to exclude several informants only on the basis of written questionnaires: they were very often non consistent and did not translate what they were asked for (for the features of a good informant, see Cornips and Poletto (2004)). As for the format of the questionnaires themselves, we paid attention not to have too many sentences which were minimally different in a row in order to avoid a well-known psychological adaptation effect, therefore the questionnaires where first prepared according to each phenomenon tested and then the sentences were “scrambled”. Each oral working session with informants was no longer than one hour, due to the structure of the questionnaires and included also presentation of the test and data on the informant.

7. Conclusion

In this article we have presented the way the ASIS project has been set up in the hope that our experience will help other projects to develop. We are very much aware of the fact that the methodological choice we made has influenced the sample we have gathered and that in many cases our work, although stretched throughout a considerable time span, has not been precise enough; we are continuously working to render our methodology more and more refined. The sociolinguistic situation of Italian dialects is very peculiar, dialects are generally quite distinct from the standard language in the perception of native speakers; there is generally no negative bias towards dialects, so that it is very common to find linguists who are native speakers of a dialect and can be used as first informants; this situation has clearly facilitated our investigation, and we know that in other countries the choice we made would turn into a complete failure, especially in those in which dialects are perceived only as a “bad” or incorrect version of the standard language.

References
Ledgeway (2003)