

The Role of Inversion and PP-Fronting In Relating Discourse Elements:

some implications for cognitive and computational models of
Natural Language Processing

Mark Vincent LaPolla
The Artificial Intelligence Laboratory
and
The Department of Linguistics
University of Texas at Austin,
Austin, Texas USA
11 April 1986

0. Abstract

¹ This paper will explore and discuss the less obvious ways syntactic structure is used to convey information and how this information could be used by a natural language database system as a heuristic to organize and search a discourse space.

The primary concern of this paper will be to present a general theory of processing which capitalizes on the information provided by such non-SVO word orders as inversion, (wh) clefting and prepositional phrase (PP) fronting.

1. Introduction

English at its simplest is an SVO, Subject Verb Object, language. However it is not limited to SVO order. Clefts, pseudo-clefts, inversion, topicalization, left dislocation and various types of fronting are instances of deviation from SVO order. Non-SVO orders are not exceptional or found only in obscure literary writing. They abound throughout writing and speech. An interesting question is what use do these structures do in English, and how can a cognitive or computational theory use such information?

Non-SVO word order helps the reader (or listener) to construct a discourse representation. It is a heuristic device for creating coherent and cohesive representations of text and for searching existing representations. In other words, it is a device for finding in long or medium term memory the relevant context, or discourse space, in which to embed and interpret the sentence being processed. It is a linguistic device used for changing the discourse focus (Sidner 1978, 1983; Grosz 1978, 1981). It is important to note that no particular discourse representation construction schema is assumed in this claim but rather that this claim holds no matter what type of construction algorithm is used. Also, what is being proposed here is not a linguistic rule for constructing discourse representation but rather a principle (heuristic) for constructing them. That is a principle for organizing and searching discourse representations.

As a sentence is processed a representation is made of it. This representation consists of a label, which is a (syntactic and semantic) parse² of the first constituent encountered, followed by a parse of the whole sentence. Actually, one could have multiple labels that consisted of the NPs in the sentence,

¹This research was supported by the U.S. Army Research Office under contract DAAG29-84-K-0060. Artificial Intelligence Laboratory, Department of Computer Sciences.

I would like to thank Professor Robert F. Simmons for his support, help and criticisms.

²The exact representation of the text will be discussed later. For now English will be used to represent what eventually will be a discourse representation of some sort. Parse is being used here in a general sense to mean "give the structure of". The "structure" of course depends on the linguistic theory used to give the parses.

since it seems that these non-SVO structures are sensitive to NPs. Thus every discourse representation is in some way redundantly specified for at least one constituent.

For example, the organization of the discourse representation for the sentences, "In the forest stood a house" and "In the park, Mary kissed John", are:

Label: house, forest
D.R. : a house stood in the forest

and

Label: Mary, John, park
D.R. : Mary kissed John in the park

Two uses could be made of this system of "labels". The first is to simply use the labels to index the text and to facilitate the search through the text. The second would be to use the labels as nodes in a semantic network. Thus in the discourse fragment, "A house stood in the forest. Outside stood an angel." the label "outside"³ could be related to either the label "house" or "forest" or both via a "location" arc.

Due to the lack of space the former (index) use will be the focus of this paper but note that it is indeed difficult to separate these two uses since it is necessary to locate the material necessary to construct a discourse space, i.e. even if one were only using a very restricted procedure for creating discourse representations one would still need to locate previous discourse items in order to resolve anaphora, and therefore create some sort of structured link from discourse to discourse.

In the sections that follow we will give linguistic motivation for this analysis. Due to the enormity of the area and the limitation of space, only inversion and PP-fronting will be discussed. In the final section we will present the findings of a study done to see if this system of "labels" could be used to speed the construction of discourse representations, specifically the resolution of anaphora.

2. Inversion

In Green's (1980) study of the discourse function of various "classes" of inversion, she assigns each instance of inversion a particular function. This approach, however, is bound to be inadequate. If one tries to compile a list of such functions for various syntactic forms in a language, how does one know if the list is complete? Every time a new function is discovered for a form one must add it to the list. Another problem with this approach is that one never knows if one has been specific enough or general enough. Being too specific in the characterization of such functions creates a very long list and

³Or rather the relation "outside" that implies outside of something.

can miss generalizations. Being too general might hide the correct function of a form in a language. This section will review and criticize a subset of her list of proposed functions.

2.1. Use 1

Inversion allows the listener first to identify the object being talked about before assigning information to it, whether new or old. This is most noticeable in the speech of sportscasters (Green; p 584):

- a. Underneath is Smith
- b. High in the air to get the ball was Jim Brady.
- c. Bringing the ball up is Marty Mestemacher.
- d. Back come the Kahoks with the ball.
- e. And sitting down is Kevin Jones.

In this case inversion allows the viewer to single out the player on the TV screen before processing his name. This allows the viewer to first pick out the player, make a non-linguistic representation in memory of that player and then add the (new) information given by the sportscaster: the player's name. If inversion were not used the viewer would have to first store the name given by the sportscaster, identify the player, construct a representation of that play and his actions, retrieve the player's name and then assign that information to the representation created by the description of the player's actions. This is a much more difficult and time consuming procedure. In this instance, inversion helps to cut down the amount of processing necessary to construct a representation.

Therefore one could hypothesize that upon hearing/reading the first few elements of a sentence, the listener follows three steps:

1. if the sentence is SVO (and does not have any overt signals to search for a previously mentioned item of information,) construct a representation of the sentence and add it to the local discourse space.
2. Else search memory for the last mention of the item under construction and add the "new" information, i.e. what is in the predicate, to that local discourse space. Pointers are left pointing to both discourse spaces.
3. Otherwise, construct an entirely new discourse space.

Actually it is unclear where the information should be deposited. For example, a house has been robbed. The police investigate. They ask questions concerning the robbery. Then the father of the household, when asked who they think could have done it, says: "As I said before, that boy John is a thief."

The police who do not have any previously knowledge of John add this information to the present discourse space. They probably note that the person who said it has reasons for

bringing it to the listener's attention that he had said this before. His daughter, however, who knows John but did not know her father's opinion of him, adds to her mental representation labeled "John" this information. She also adds this utterance to the current discourse space, i.e. "robbery of house". The man's wife on the other hand already knew about her husband's opinion. It would seem redundant for her to add this information to her mental representation labeled "John". What she probably does is call up the mental representation "John" and leave a pointer pointing to it from the current discourse space "robbery of house", thereby connecting the two representations. So it would seem that things are not as cut and dried as one would suppose.

2.2. Use 2

The second use for inversion which Green cites is its ability to connect pieces of discourse together. This is used frequently in journalism. One can link and expand a previously mentioned proposition by anaphorically referencing it in the grammatical subject slot, thereby smoothly linking the new information in the predicate to the old, previously mentioned information. She also notes that this connective function is used in literary texts or expository prose. She finds that inversion provides a concise form in which news writers can begin a sentence with old information.

In the analysis being developed in this section, the only difference between the sportscasters' speech and this "connective" function is that the "connecting" in the sportscasters' speech is done to an image rather than a (previously mentioned) linguistic concept. In the sportscaster case, the viewer has to create a nonlinguistic representation in his consciousness, i.e. his discourse space. Then when the name of the player is mentioned the viewer assigns the information to the representation. The "connective function" case differs only in the type of representation built.

2.3. Use 3

In her third function, Green expounds upon the notion that what is inverted is not necessarily new information. She shows that inversion can be used to set a scene for an event or as a means to locate actors in a story, e.g. "Outside stood an Angel", "In a little white house lived two rabbits."

Notice how similar this "function" is to the sportscasters' speech (and the newspaper examples). The sportscaster speech uses inversion to identify the player so that the listener can more easily identify the (new) information, i.e. his name, with the player. The scene-setting, and literary connecting, function of inversion identifies a locale in which to place the actors, the characters. From a processing point of view these are the same things. Even Green notes the similarities between the journalist function and the literary connecting function.

13(b) Sprawled in the foreground is George

Price.⁴

"Ex. 13b, which is part of the description of an accompanying picture, is very much like the news examples: it identifies a piece of the picture, an individual, by locating him with reference to --more specifically, in the foreground of--something taken to be already salient, the picture as a whole." (p. 588; underlining added)

In all these cases a discourse representation can be created that uses as its label the first constituent of the inversion. In the sportscasters' speech the representation of the sentence(s) would be linked to the image of the player and the (information) "living of the two rabbits" would be assigned to the discourse representation labeled "white house". This last assignment might seem strange but suppose that the house was previously identified as being in a large forest:

Discourse 1.

In a large forest stood a house.
In the house lived two white rabbits.

This use of inversion does not seem to signal an extensive search of memory but rather seems to create a more local chain of association: there is a sense in which the second sentence is an elaboration of the stored information about "a house". Compare the short discourse structure above with

Discourse 2.

In a large forest stood a house.
Two white rabbits lived in the house.

Intuitively, this discourse seems harder to process than the previous one. Finally compare these sentences with:

Discourse 3.

A house stood in a large forest.
In the house lived two white rabbits.

This last discourse seems as easy to process as Discourse 1 (D1). The theory under development here accounts for this.

Discourse 2 (D2) is harder to process than D1 because when processing D2 one must store the concept "two white rabbits lived..", in some manner, and then search for a previous mention of "house" in which to embed the information. (Notice it is not the inversion that makes "the house" anaphoric but rather the use of the definite article.⁵)

⁴Green's numbering. I will continue to use Green's numbering for her examples. I will use a more coherent numbering system for my examples.

⁵I would still like to maintain that inversion is used as a signal of anaphora.

Inversion, in this case, makes the discourse easier to process.

Discourse 3 is as easy to process as Discourse 1 because the inverted element, "a house", becomes the label and the "connecting phrase", the label, of the representation of the second sentence is "in the house". An interesting observation is that Discourse 5 seems harder to process than D1, as expected, but easier than D2. The explanation for the latter observation is not at first obvious.

Discourse 5.

A house stood in a large forest.
Two white rabbits lived in the house.

In the case of the above discourse (D5), "a house" is the label of the last representation built. So even though the connecting phrase "in the house" is not in initial position, which accounts for why this discourse is harder to process than D1, there is a "top level" item, i.e. label, "a house", to which the connecting phrase "in the house" can connect. However, in D2 not only is the connecting phrase "in the house" buried in the second sentence, i.e. not in initial position, the item to which it must connect is also buried. Extending this reasoning the theory would predict that D3 would be easiest to process, D1 and D4 the next hardest and D2 the hardest. This claim is a strong claim about the internal structure of discourse representations and could be falsified with psychological experimentation.

The intuitive sense in which a sentence is harder or easier to process is perhaps also related to the idea that the subject is an external argument which participates in a predicate relationship with the entire VP and not just the verb. In this sense the object(s) of the verb are more "deeply embedded" in the sentential predicate than the subject. Thus inversion, PP-fronting, etc., can be viewed as moving an embedded, or internal argument, to a more external position, e.g. adjunct position.

Not also that those verbs which appear in inverted sentences seem to be ergative verbs. That is the deep structure of the sentence "Outside the house stood an angle" is probably [_S e [_{VP} stood an angle] [_{PP} outside the house]⁶ (Heim 1985, personal communication). This might help to explain the greater "availability" of fronted material.

An important point to note is that Green does not consider the scene setting function and the literary connecting function to be the same thing. In the scene setting examples the inverted element is completely new information, whereas in the literary connecting function this does not have to be the case. This is an important point for the theory in this paper as well. Crucially, the claim of this theory is not that the inverted element is old information but that it is the

important element with respect to embedding of information. When it is new information it sets up a context in which new information can be embedded, including the information in its own predicate. When it is old information it serves to find the correct context in which to embed the information in the predicate.

In all of the above cases, inversion is used to locate and identify an (old) entity, an event in the sportscaster speech, a (previous) location, or an image, and give more (new) information about that entity, or create a context in which to embed information.

3. PP Fronting

PP-fronting is used to provide a continuity, a cohesion, in the text. It provides a useful progression of labels to which to attach the accompanying information. For example, an article by Lawrence (1985) opens with a fronted prepositional phrase which provides a time setting, or relation on the (narrative) time line, for the activity in the sentences which follows⁷:

Discourse 6.

Long before I was tall enough to ride on the big coaster myself, I spent many pleasant hours persuading my reluctant father to accompany me. (p. 4)

The PP also provides a way to link up the topic (theme) of the article to the opening statement of the article. The theme is the "new adult" Amusement Parks. The article initial prepositional phrase picks out a particular item within an amusement park and associates the remembrances of the writer to it.

The next sentence also has a fronted PP. This PP also links the next sentence to the article's main topic:

Discourse 7.

As an aficionado of amusement parks, I was overjoyed when our whole family finally flew to California to tackle Walt Disney's extravaganza. (p. 4)

The next paragraph starts out with yet another fronted temporal prepositional phrase, moving the time setting up to the present:

Discourse 8.

More than two decades later, I'm still journeying to parks.

These first few examples of word order deviation all have the characteristic of giving the reading a temporal "focus" and order in a series of events that occur over a number of years.

In order for a discourse processor to understand this text, it

⁶This is roughly the structure.

⁷The underlining is mine

would have to have a place to start. The logical starting place would be with the label AMUSEMENT PARK, since this is in the title of the piece. Under this discourse representation label it could build other representations. The first representation that it would build would be about the author since this is the first matrix NP of the first sentence, D6. For the next sentence, D7, it would already know what type of temporal relation to assign to the proposition expressed in the matrix clause. The sentence in the next paragraph is easily processed since it advances the time of the preceding paragraph. Rather than building a representation of items and attaching to this a set of properties, these fronted PPs build an abstract representation of temporal items related by the time of each item.

The general theme of the article is amusement parks. However, since the opening of the article is more a personal recollection rather than expository, the information that needs to be organized is not information about particular objects, i.e. amusement parks, but rather episodes in the author's life. Each episode's temporal relation is specified by the fronted prepositional phrase. In general this is the function of fronted prepositional phrases, the specification of relations. This is a widely used technique; used more than any other non-SVO pattern:

The general hypothesis is that the first thing that one encounters in a sentence is a link to preceding information, either explicit or implied. The link provides the proper context in which to build the new representation. It also provides the means for quickly searching the discourse space.

PP-fronting, like inversion, allows the reader to connect the current phrase, or sentence, being processed to an appropriate, and most likely salient, antecedent. For example Green (1980) comes to the same conclusion about inversions in sportscasters' speech. She notes that sportscasters use inversion when broadcasting play-by-play to identify the player by his action and then name him:

Stealing it and then losing it was Dave Bonko.

Back come the Kahoks with the ball.

And in comes number 51, and that will be Mike Matakitis.

Into the game for the New Trier West is Brenner.

The reason she gives is that this is helpful to the TV viewers, since they don't have scorecards identifying the players. She goes on to say that in this way the viewer can single out the player on the screen before receiving his name. (This also gives the sportscaster time to look up the player's name if need be.) Further more she notes that sportscasters use this inverted style even when the player are well known or their numbers are clearly visible. This observation fits in nicely with the model being built here.

Thus like inversion PP-fronting is used to help link, via labels, i.e. focused material, one discourse representation to another.

4. The Experiment

In the above two sections we briefly motivated and developed an analysis of the organization of discourse representations. Basically the analysis claimed that each discourse representation, no matter how it is represented, i.e. what particular theory or formalism, were indexed via their focused NPs. The analysis also claimed that non-SVO word structure was a signal to search through the labels to locate the structure in which to embed the representation currently being processed.

There are two aspects of this analysis that we will focus on in this section: the creating of labels and the searching of the labels. The more complicated aspects of building and embedding, or relating, the structures to one another will be ignored for the sake of exposition.

A simple experiment was performed to explore the computational usefulness of the proposed labeling system. Three programs were written in Symbolics Prolog. Each program processed a set of twenty-six sentences and created discourse representations. To create the discourse representations the DRS construction algorithm found in Kamp (1986) was used. Added to this were straightforward rules for creating DRSs for locative prepositional phrases. The task for each program was to resolve simple anaphora by searching through the discourse representations for the antecedent. A straightforward feature matching technique was used to do this. If one were trying to resolve the reference for a pronoun and a full NP then only the features of the lexical item, e.g. masculine, singular, was matched. If the reference for a full NP was being resolved then the whole lexical item was searched for.

The first program only constructed discourse representations. It did not construct labels as well. Thus whenever anaphoric resolution was called for by the DRS construction algorithm, this program had to search through the entire data base until a match was found. The second program created labels but they were only searched when the sentences being processed had non-SVO structure. The third program created labels as well but it only searched the labels. That is the heuristic always applied.

Each DRS was a flat list. Each label list was also flat. Before each run of the program the machine was cold booted.

The data was a list of 24 sentences. The last sentence contained the only fronted PP, which referred back to the first sentence. The results of this experiment are discussed in the next section.

5. Results

In pilot experiments the DRS list was allowed to have complex structure. In other words, the DRS list was a lists of lists (of lists, etc). The label list on the other hand was flat, i.e. a list of lists. The processing in the case of the complex structure was speeded up by a factor of 3, overall, when the heuristic was applied at all times (in the third program). In the case of the second program only the processing of the non-SVO sentences, the last sentence in the corpus were sped up. However in subsequent experiments the DRS lists were all converted into flat lists. The effect was less dramatic but significant none the less. Below is given two tables of the figures for all 3 runs. The first table is the time each program took for processing the non-SVO sentence, i.e. the last sentence in the corpus and the time it took to process the corpus overall⁸.

Table 1	Time (secs) of Non-SVO Sentence	Time for Entire Corpus
First Program No heuristics	.023	.66
Second Program Heuristic on Non-SVO structure	.0187	.659
Third Program Heuristic at All times.	.019	.54

⁸The times given in table 1 are rounded to the nearest thousandth of a second. However all percentages were computed with original microsecond numbers and then rounded.

Table 2	% of speed Up as compared to First Program
First Program No heuristics	Speed up: 0% on Non-SVO 0% overall
Second Program Heuristic on Non-SVO structure	Speed up: 19% on Non-SVO 1% overall
Third Program Heuristic at All times	Speed up: 16% on Non-SVO 19% overall

It is obvious from the above table that always searching the label list(s) is far superior to searching the larger discourse structures. It is also significant that when the heuristic (searching the label list) is only applied to Non-SVO structures, i.e. the last sentence of the corpus, that the speed up is significant. (The difference between the second and third program with respect to the last sentence, the non-SVO sentence is not significant and is due to machine related factors.) Therefore it seems that it is worthwhile building a separate list of NPs and searching that list for at least resolving anaphora and using it not only for the (linguistically) motivated non-SVO structure search but all the time as well.

6. Conclusion

We have motivated a heuristic device that consists of creating a list of the NPs in each sentence. This was motivated on linguistic grounds for non-SVO word order. It was suggested that this list could facilitate in the construction discourse representations and for resolving anaphora in a computer program. The latter of these two suggestions was investigated. It was found that indeed a significant decrease in processing time occurred.

The first of the two above hypothesis was not empirically investigated. One avenue of interesting research would be to see if the information provided by non-SVO word order could help in the construction of more complicated discourse representations and if such representations would help in areas like Question-Answering.

A second avenue of research would be in psycholinguistics. Basically experiments could be set up to test the hypothesis that non-SVO word order somehow signals a search of the discourse space.

References

- Sidner, C. Focusing and Discourse. *Discourse Process*, Oct-Dec 1983, pp. 107-130.
- Chafe, Wallace. Language and Consciousness. *Language*, 1974, 50, 111-133.
- Chafe, Wallace. Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View. In Charles Li (Ed.), *Subject and Topic*. New York: Academic Press, 1976.
- Davidson, Alice. Peculiar Passives. *Language*, March 1980, 56(1), 42-66.
- Green, Georgia M. Some Wherefores of English Inversions. *Language*, 1980, 56(3), 582-602.
- Grosz, B. . Focusing and Description in Natural Language Dialogues. In Joshi, A. B. L. Webber and I. A. Sag (Eds.), *Elements of Discourse Understanding*. Cambridge: Cambridge University Press, 1981.
- Grosz, B. . Focusing in Dialog. In *Tinlap-2: Theoretical Issues in Natural Language Processing*. New York: ACM and ACL, 1978.
- Haviland, Susan E. and Herbert Clark. What's New? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 512-538.
- Heim, Irene R. *The Semantics of Definite and Indefinite Noun Phrases*. Doctoral dissertation, University of Massachusetts at Amherst, September 1982.
- Kamp, Hans. Situations in Discourse without Time or Questions. *CSLI*, 1986. To appear in *CSLI Stanford*.
- Kamp, Hans. Context, Thought and Communication. In *The Proceedings from the Aristotelian Society*, 1985.
- Lawrence, A. Amusement Parks. *AAA*, 1985, .
- Prince, Ellen F. A Comparison of Wh-Clefts and It-Clefts in Discourse. *Language*, 1978, Vol. 55(4).
- Schank, Roger C. *Dynamic Memory: a theory of reminding and learning in computers and people*. Cambridge University Press, 1982.
- Sidner, C. Levels of Complexity in Discourse for Anaphora Disambiguation and Speech Act Interpretation. In *Proceeding of the Fifth International Joint Conference on Artificial Intelligence*. IJCAI, 1978.