THE PREDICATE-ARGUMENT RELATION AND THE STRUCTURE OF
DISCOURSE IN SPOKEN FRISIAN

Eric Hoekstra

Abstract. This article provides an analysis of some fragments of discourse taken from the
Corpus Spoken Frisian. It is claimed that the predicate-argument relation provides structure
to such discourse fragments, provided that the argument position of the predicate can be
occupied not just by a NP or CP but also by a list of clauses and phrases. The scope of the
predicate argument relation is thus extended from intrasentential structure to discourse
structure. In addition, attention is drawn to two syntactic phenomena which are present in
lists of clauses and phrases of spoken Frisian, but absent from written Frisian:
complementiser repetition around a phrase and the use of fan ('like') to introduce direct
speech.

1. Introduction

1.1. Spoken and written language
Spoken language displays a system of text coherence which is to some extent different
from text coherence in written language. To illustrate, everybody can understand a written
text such as is encountered in, say, a newspaper or a novel. Everybody can understand a
spoken conversation between, say, two secretaries. But if the spoken conversation is
written out as a transcript, then it is much harder to understand than a 'normal' written text.
This goes to show, as is well-known, that spoken language written down is not the same as
written language.

Some of the differences between written language and transcripted spoken
language are well-known. Written language features clear breaks between sentences, which
can be gleaned from the orthography. Transcripts of spoken language, in contrast, feature
much less orthography than written language. This is partly due to the choice of
transcription protocol. But even if the transcription protocol allows for the use of dots and

1 I would like to thank Sybren Dyk, Janne Bondi Johannessen, Willem Visser, Mark de
Vries and two anonymous reviewers for comments or discussion.
comma's, it is not always clear where a spoken sentence ends and the next one begins. Thus uncertainty (on the part of the linguist) about beginning and end of a sentence, or even a phrase, seems to be a feature of spoken, but not of written language. This uncertainty, however, does not cause any impediment to understanding spoken language.

Another well-known difference between spoken and written language involves the fact that spoken languages feature lots of repetitions, false starts and phrases and clauses which are left unfinished. Crucially, this does not make spoken language less intelligible than written language. But when we read transcripted spoken language, full of false starts and unfinished clauses and phrases, then we do have trouble understanding it. Again, this goes to show that spoken language obeys laws which are different from 'normal' written language.

How do people construct a coherent interpretation from the 'messiness' of spoken language? Syntax and semantics tell us how sentences are analysed and how they receive an interpretation. The sentence and the phrase have received lots of attention from grammarians. In comparison, less attention has been given to the syntax and semantics of texts.

1.2. Spoken and written Frisian

The purpose of this short article is to draw attention to and provide an analysis of some features of spoken Frisian. Frisian is a minority language spoken in the north of The Netherlands. An overview of the grammar of written Frisian can be found in Tiersma (1999). The linguistic study of Frisian is closely connected to the study of Dutch, for the simple reason that Dutch is the official language in the part of The Netherlands where Frisian is spoken. Thus researchers of Frisian often have a background in Dutch linguistics. Furthermore, official languages like Dutch or English have been well researched by a large number of linguists. It should be borne in mind that researchers of minority languages are few, and governmental backing, which focuses on the standard language, is scanty.

Linguistic interest has focused on the syntax and semantics of the sentence (of written language), at least if we restrict our attention to Dutch and Frisian. J. Hoekstra (1986-1993) occasionally has observations about spoken Frisian in his language column, which appeared in a Frisian newspaper. As for Dutch, J. de Vries (2001) has a semi-scientific catalogue of observations about spoken language; Jansen (1981) analyses some
constructions which are characteristic of spoken language; and W. de Vries (1910-1911, 1911-1912) contains an excellent, albeit chaotic, overview of various phenomena found in spoken language and in dialects. In general, even when spoken language data are used, syntactic attention has focused on the sound, the word and the sentence, not on the discourse.

To exemplify, a large corpus has been completed for Dutch, the Corpus Spoken Dutch (‘Corpus Gesproken Nederlands’). Research has tended to use this corpus not so much to investigate phenomena involving the discourse, but to investigate intrasentential phenomena using the spoken language data as a more suitable source of data than their own native speakers’ intuitions. For example, Coussé, Gillis and Kloots (2007) use the corpus data to study vowel reduction. Oosterhof (to appear) uses the corpus to offset contrasting linguists’ judgments on genericity against what is actually found in reality (the corpus). Thus, although studies based on the Corpus Spoken Dutch shed light on intrasentential phenomena, particularly when linguists argue about the data, such studies have not (yet) focused on the phenomena which are the subject of this paper.

My primary aim is to describe some of the syntactic differences between written Frisian and transcripted spoken Frisian. I presented an inventory of such differences in Hoekstra (2005); here I develop some ideas which I briefly touched upon there. Study of the differences between spoken and written language should ideally help us in answering the more abstract question: how does the interpretative component of the human language faculty function. Below I will analyse some specific phrases and clauses from the Corpus Spoken Frisian, and I will argue that the predicate-argument relation is particularly relevant for assigning structure to discourse. In addition, some characteristic features of spoken Frisian, as opposed to written Frisian, will be presented, involving the repetition of complementisers and the use of the marker of direct speech *fan* (“like”). Below, I will first discuss some methodological aspects.

The corpus which is used here is the Corpus Spoken Frisian. It can be consulted in the Frisian Academy, but the expectation is that it will be available through the internet by 2009. The corpus contains transcriptions of recordings. All in all, the corpus contains about 50,000 tokens. The recorded material involves both dialogue and monologue. For this paper, we have restricted our attention to dialogues.

I will use the term discourse fragment to denote a part of a discourse that is cited to illustrate some phenomenon. It is usually a unit larger than the example sentence. I will use
the term discourse chunk to denote a part of a discourse that functions as an argument in a predicate-argument relation. I will show that a discourse chunk may consist of a list of clauses and / or phrases.

I will present a non-quantitative exploration of how syntax, more specifically, the predicate-argument relation, helps to organise phrases and clauses of spoken language into a larger whole. A quantitative exploration can be made once the corpus is large enough to meet the demand of significance. Furthermore, qualitative exploration must precede quantitative investigation so that one has an idea of which aspects of spoken language are suitable for further inquiry (cf. Tacq 1997:21-22).

2. The predicate-argument relation

The predicate-argument relation provides structure to phrases: a predicate and its argument(s) are joined to together to the exclusion of other material. Many grammatical approaches have an elementary notion of predicate argument structure. For example, generative grammar studied predicate argument structure since its inception (Chomsky 1965, Gruber 1965). In later versions of generative grammar, Theta Theory (Chomsky 1981:36, reformulated for chains 335) was introduced to deal with the relation between thematic roles and argument positions. Theta Theory contains the Theta-Criterion, which is formulated as follows:

Theta-Criterion
Each argument bears one and only one theta-role and each theta-role is assigned to one and only one argument.

For example, a transitive verb like kiss has two participants, as is clear from an example like: Mary kisses her child. The thematic role of 'kissee' (the one being kissed) is assigned to the direct object position, which is occupied by the NP 'her child'. The thematic role of 'kisser' is assigned to the subject position, which is occupied by the NP 'Mary'. Theta Theory specifies which theta role is assigned to which position. Theta Theory also requires that every thematic role which can be assigned must be assigned, and that an argument must have exactly one theta role.
In Chomsky (1981) and subsequent work, Theta Theory is discussed only within the context of intra-sentential grammar. Chomsky discusses cases of NPs and CPs (sentences) which are coindexed with an argument position. It will be shown, however, that a list of phrases and clauses may be coindexed with an argument position, too. Thus the predicate-argument relation not only provides structure to a phrase or clause, but it may also provide structure to a list of clauses and phrases. This holds true regardless of the specific theoretical framework which is adopted, as long as it incorporates the notion of predicate-argument structure.

3. Analysis of a specific example

3.1. Data

Issues of phonology, like mistakes of pronunciation, naturally fall outside the scope of this investigation. The following fragment of discourse will be the subject of syntactic analysis.

(1) Wy wolle ek sa sa mei sa weinich mooglik jild wat dwaan yn ider gefal
do we want also so so with as little possible money something do in every case
datsto net at it jild der net mear komt datsto stopje moatst.
that-you not if the money there not more comes that you stop must
'tWe also want that, to do something with as little money as possible and in
any case we don’t want to have to stop if the money doesn’t come anymore.'²

This discourse fragment can be syntactically analysed into simple clauses and into whatever phrases may be left over once the clauses have been set apart. A clause or phrase can occasionally be incomplete as it is left unfinished (cf. (2d) below). Syntactic analysis of (1) thus yields the following syntactic phrases and clauses:

(2)

² Dutchisms in Frisian like "toen" foar written Frisian "doe" are coded in the transcription to indicate that a Dutchism is involved. I have omitted these codes in the examples in order to facilitate understanding. Furthermore, Dutchisms are not relevant to the subject of this paper.
I will attempt to make explicit the syntactic constituents which this discourse consists of and how they fit into a larger whole. While analysing these phrases and clauses, I will point out some syntactic peculiarities of spoken language which appear in this discourse fragment.

3.2. Analysis: thematic role assignment to a list

The verb to want (Frisian: wolle) selects a complement which receives the thematic role of theme. Which discourse chunk receives that thematic role of want? The answer is: the thematic role of theme is assigned to the discourse chunk represented by the phrases and clauses (2b-2f). Together they express what it is that is wanted. The repetition of so in (2a) seems to mark the speaker’s uncertainty about how to phrase his or her thoughts.

Theories of grammar such as Chomsky (1981) and subsequent work discuss
predicate-argument structure in the context of intrasentential grammar. Thus, the empirical
data base has been subject to an idealisation, that is to say, the data base has been
narrowed. However, this narrowing may be unwarranted, seeing that the predicate-
argument relation also serves to provide structure to a list of sentences and phrases. The
key observation concerning (2) can thus be worded as in (3):

(3) The theme theta-role of 'want' is assigned to a list of phrases and clauses (a
discourse chunk).

To make this easier to the understanding, we can 'translate' the spoken Frisian of (2) into
written Frisian as in (4):

(4) Wy wolle ek mei sa weinich mooglik jild wat dwaan en wy wolle yn ider gefal
    we want also with as little possible money something do and we want in any case
    net, at it jild der net mear komt, datst stopje moatst.
    not if the money there not more comes that-you stop must
    'We also want to do something with as little money as possible and in any case we
don’t want to have to stop if the money doesn’t come anymore.'

In this ‘translation’ we repeated the theta-assigning predicate want, which indicates that
there are two recipients for the theme theta role of this predicate. Furthermore, there seems
to be an implicit coordination of elements in (1), which is made explicit by the use of the
coordinating conjunction.

It thus happens to be the case that a chunk of apparently loose phrases and clauses
receives an interpretation by virtue of it being assigned one thematic role as a whole. In
written language, what is assigned a thematic role is a neat subordinated clause or a
conjunction of subordinated clauses. This need not be the case in spoken language. We can
now be more explicit about the substructure of (1). The direct object of want (wolle) is a
list of two subordinated clauses, as indicated below:

(5) Wy wolle ek sa sa { [2b: mei sa weinich mooglik jild wat dwaan] [2c-2f: yn ider
gefal datsto net at it jild der net mear komt datsto stopje moatst] }
The outer brackets, made curly to highlight them, indicate the edges of the list which is assigned the theme theta-rol by the verb *want*. The inner square brackets represent the constituents of the list. The list may also be termed an implicit coordination as there is no explicit conjunction used to string these sentences together.\(^3\)

List structure turns out to be relevant for the analysis of spoken language, as pointed out for English by Lerner (1991). The subject of Lerner's research, however, is turn-taking, more specifically, turn-taking where the two speakers together produce a single sentence or discourse fragment. He notices that a change of turn may occur under various conditions, one of which is before the last element of a list. His examples include *if-then* constructions and verbs of saying. Such turn-taking involves recognition and use by the turn-taking speaker of the structure that is projected. This phenomenon may be taken as support for the idea that there is such a thing, alongside intrasentential structure, as discourse structure, which helps to interpret spoken language. In our view, it is the predicate-argument relation that is of vital importance in providing discourse with structure.

\(^3\) A reviewer suggests that a list of clauses and phrases without explicit conjunction is not principally different from a list of clauses with an explicit conjunction. That seems too restrictive to me, as the relation between two elements of a list need not be the same as the relation between two explicitly coordinated elements. To illustrate, consider an example that is presented in the text as (8), here as (i):

(i) Miskien \{ [datst mei in hammer ] [datst it dêrmei dwaan kinst. ] \}
    "maybe that-you with a hammer that-you it there-with do can"
    'Maybe you can do it with a hammer.'

The relation between the two bracketed list elements does not involve coordination. Adding a conjunction yields a decidedly odd result:

(ii) Miskien \{ [datst mei in hammer ] en [datst it dêrmei dwaan kinst. ] \}
    "maybe that-you with a hammer that-you it there-with do can"
    'Maybe (you) with a hammer and you can do it with that.'

Consider also the text example (7), in which the relation between the two list elements feels intuitively like apposition. Thus, the conclusion is that the relation between two elements of a list is much broader than just coordination.
3.3. Other cases involving thematic role assignment to a list

(2) is not an isolated case in spoken language. Spoken language regularly contains a list of phrases and clauses which as a whole receive a thematic role from a verb. To illustrate, consider the sentences below:

(6) No ha ik ek ja. ik wit noch wol { [dat ik dat eh] [toen wy noch lyts wienen en sa toen hiest de skoallefakânsje] [dat wie earst wer hiel leuk] and so then had-you the school-holidays that was first again very nice [giest op fakânsje] [en dan op it lêst dan hiest sa echt fan wat moat ik go-you on holidays and then at the last had-you so really like what should I no dwaan] [koest dy noch ferfele] } now do could-you still bore 'Well, I also had that, yes, I still remember, when we were young and all that, then one had school holidays. That was at first very nice again. One went away on holiday vacation. And then at last one really felt like what should I do now. One could still get bored.'

(7) Mar der binne ek boeken dan tinksto fan { [ik moat eefkes al al dy ferhalen but there are also books then think-you like I must quickly all all those stories trochnimme of sa] [Kin dat net by de moandei stean.] } peruse or so can that not with the Monday stand 'But there are also books, then you think like I must quickly go through all those stories. Can’t that be part of Monday?'

In (6), the list of clauses, indicated by curly brackets, receives the theme theta role of the verb witte, which translates as remember. The first element of that list is left unfinished (“dat ik dat”). The third list element (“dat wie earst hiel leuk”) begins with the pronoun “dat”, which refers back to the school-holidays, which were introduced in the second list element; the six clauses that are listed together make up what the speaker remembers.

In (7), the verb to think assigns its thematic role of theme to a list of two clauses. Note, incidentally, that such lists are regularly introduced in Frisian by the preposition fan,
literally 'of', which in this quotative use is similar to English *like*. On the syntactic analysis of *like* (*fan*) in Frisian and Dutch, see Hoekstra (2007, in preparation).4

3.4. Lists containing phrases and incomplete clauses

The examples discussed so far were used to illustrate the observation that a list of clauses as a whole may receive a thematic role from the verb. Another observation, characteristic of spoken language and banned from written language, is that the complementiser may be repeated, in case part of the chunk has a syntactic function in another part of the chunk. Some examples are given below:5

(8) Miskien { [datst mei in hammer] [datst it dêrmei dwaan kinst] }  
\textit{maybe that-you with a hammer that-you it there-with do can}  
'Maybe you can do it with a hammer.'

(9) Wy wolle ... { [datsto net] [at it jild der net mear komt] [datsto stopje moatst.] }  
\textit{we want ... that-you not if the money there not more comes that you stop must}  
'We don’t want to have to stop if the money doesn’t come anymore.'

In (8), the predicate *maybe* assigns a thematic role to the chunk marked by curly brackets. The chunk consists of: complementiser + phrase + complementiser + clause with

4 A reviewer objects that the analysis proposed in the text equates functional framing with grammatical dependance. It seems to me that such an equation or reduction is correct for the cases under consideration. Given that there is such a thing as predicate-argument structure, it is hard to escape the conclusion that it helps to structure a loosely connected list of phrases and clauses.

5 This construction is discussed in Hoekstra (1993:201) and (De Vries 2001: 534-535). A slightly different construction is analysed in Jansen (1981: 198-241). Jansen’s examples involve cases which feature not the repetition of a complementiser, but the repetition of a verb, as shown below:

(i) hy is yn eh yn Ljouwert is er berne  
\textit{he is in Ljouwert is he born}  
'He was born in Ljouwert.'

The phrase \textit{in Ljouwert} is sandwiched in between two occurrences of the tensed verb *is*. This construction is also typical of spoken language, and it is absent from written language.
resumptive pronoun. The phrase (PP) has a function in the following clause, which is marked with a resumptive pronoun.

(9), identical to (1) but with the first element of the list left out as indicated by the dots (...), is a more complex case. It consists of: a complementiser + negation, if-clause, complementiser + then-clause. Again we see how elements mashed in between the two complementisers (negation, if-clause) have a function in the following clause. Again, the repetition of the complementiser is illegitimate in written language.

We are focusing on the second element of the list that gets the theme theta role of want. This second list element is itself again a list consisting of three elements. Those three elements are structured by the if-then relation. The then-clause is discontinuous, if we assume that the negation belongs with the then-clause. The relevant part of (1), which is repeated here as (11), has been assigned an outline structure (10):

(10) Scheme: WANT \{ \ldots \} \{ [Compl NEG] [IF \ldots] [Compl \ldots] \}

(11) Wy wolle \{ \ldots \} \{ [datsto net] \\
we want \ldots that-you not if \\
[at it jild der net mear komt] [datsto stopje moatst. \ldots] \}
\{the money there not more comes that you stop must\}
'We also want \ldots that we don’t want to have to stop if the money doesn’t come anymore.'

The first (incomplete) clause and the third clause together make up the then-clause of the if-then structure. The negation of the incomplete first clause is essential for the interpretation: the speaker wants that the addressee should not need to stop in case the money is not forthcoming. Thus spoken discourse in part receives its coherence from predicate-argument structure and from logical relations induced by conjuncts like if-then.

3.5. Theta role assignment to a list by a nominalised adjective

Of course, not only verbs and if-then schemes structure discourse. Consider the following example:
(12) No en it rare \{[at ik op de Gysbert folder] [fan dat sy dy't sy hawwe] [dër’t well and the strange if I on the Gysbert folder of that they which they have which alle aktiviteiten op stean] [dër stiet net iens it kongres fan de Akademy op] \} no all activities on stand there stands not even the congress of the A. on well dat fyn ik wol hiel raar that find I indeed very strange 'Well, and the strange thing is: If I (look) at the Gysbert folder which they have and which has all activities on it, then it doesn’t even have the conference of the Academy on it. Well, I find that very strange indeed.'

(12) illustrates another property of spoken language as opposed to written language. In written language, a predicate like strange assigns its theme theta role to a that-clause, and such clauses show the word order characteristic of embedded sentence in Frisian (and Dutch and German). In declarative main clauses, the tensed verb is in second position, whereas in embedded sentences it follows objects, particles and other obligatory elements. However, (12) does not feature the complementiser that (Frisian: dat). Instead, the nominalised adjective assigns its theme theta role to the list of phrases and clauses, which has been indicated by curly brackets. Note, incidentally, that the predicate assigning the thematic role (strange) is repeated at the end of the list; the list of phrases and clauses is resumed by the topic pronoun dat.

3.6. The predicate-argument relation and the if-then construction

In 3.4. we already saw that the if-then construction structures two clauses. Both the if-part and the then-part of the if-then construction may introduce a list. The if-part of the if-then construction contains a list in the example (12) above, where the first three list elements together make up the if-part. The then-part of an if-then construction may also introduce a list of clauses, just as verbs like to want, to say or to think may. Consider the following example:

(13) Ast te let komst eh dan \{[kriigest in útbrâner] [moatst dy melde] [moatst
The school square clean-sweep

'If you come too late then you get punishment, you must report yourself and you must sweep the school playing ground.'

This phenomenon was noticed in Lerner (1991), who referred to it as a "list-in-progress". Thus it seems an idealisation to say that if and then introduce a single clause.

The similarity between the verbs mentioned above and the conjuncts suggests that both function as predicates taking an argument. In terms of Theta Theory, both assign a thematic role to a list. This is not normally assumed in the case of connectives like if and then, though nothing seems to forbid it either.

4. Conclusions

It appears then that a thematic role like theme can be assigned not just to a clause but to a list of clauses. Thus verbs like want, say, think characteristically provide structure not just to an individual clause or sentence but to a fragment of discourse. In the same vein, sentence connectives like if-then may provide structure not just to a sentence but to a discourse consisting of two lists of clauses. While discussing these issues, I noted in passing properties of spoken Frisian which are absent from written Frisian:

- complementiser repetition around a phrase
- introduction of direct speech by the preposition fan (of)

The conclusion of this article is that the scope of the predicate-argument relation is not just the sentence but the discourse. This would imply, in the case of Theta Theory, that the Theta Criterion applies not just to a NP or to S but it may apply to a list of phrases and clauses. Thus thematic roles provide a vital contribution to the structuring of discourse, be it written or spoken. The bulleted phenomena above, however, tend to be restricted to spoken language.

The problem of grouping phrases and clauses together into larger structures is not specific to spoken language. In written language, sentences must also be organised into larger structures. However, spoken language is less explicitly structured than written...
language. In this respect, the difference between written and spoken language is more a matter of degree than of kind.

References


Eric Hoekstra
Frisian Academy (Fryske Akademy)
Royal Netherlands Academy of Arts and Sciences
Doelestrjitte 8
8911 DX Leeuwarden (Ljouwert)
The Netherlands
ehoekstra@fa.knaw.nl