TENSE, ASPECT AND THE TEMPORAL STRUCTURE OF DISCOURSE: TOWARDS AN LFG ACCOUNT

Sheila Glasbey

Department of Applied Computing
University of Dundee
Dundee DD1 4HN
Scotland, UK

srg@computing.dundee.ac.uk

Proceedings of the LFG01 Conference
University of Hong Kong,
Hong Kong

Miriam Butt and Tracey Holloway King (Editors)
2001

CSLI Publications
http://csli-publication.stanford.edu/

Abstract

This paper is based on my contribution to the workshop on Tense and Aspect at LFG01. In it, I consider the question of how information relating to tense, aspect and the temporal structure of discourse should best be encoded in LFG. I will focus in particular on what temporal information needs to be represented if we are to be able to map from the LFG formalism to some semantic representation which captures the temporal structure of continuous discourse. In particular, I will derive a preliminary set of features which are intended to be used cross-linguistically to represent tense- and aspect- related information.

Acknowledgements

I would like to thank Miriam Butt for inviting me to give a presentation at the Tense and Aspect Workshop at LFG01. I am also very grateful to all the conference participants who gave me helpful advice and comments, and contributed to a very lively and stimulating discussion.

1. Introduction

This paper addresses the question of what "temporal" information (including tense, aspect and related properties) would need to be encoded in an LFG representation, if we wished to use such a representation in order to derive the temporal structure of a piece of continuous discourse.

This very general question immediately introduces a set of more specific questions: for example:

(i) What "features" (in the LFG sense) are needed to encode the relevant information?

(ii) At what "level" should these features be represented?

Most of what follows will address question (i) – but let us first briefly consider (ii). By "level" I mean level of structure, so that one possibility might be to encode all temporal information at f-structure. Indeed, if f-structure is taken to be the sole input to semantic representation (as is stated in, for example, Butt et al 1996 p.3 and Frank and Zaenen (draft) p.1), then this appears to be a sensible way to proceed. As this stance is taken by a number of LFG researchers, I will adopt it as a working assumption here, although none of what follows is dependent upon it, as far as I can see. I will show temporal and related information represented at f-structure, but there is no reason, to my knowledge, why some of this information could not be "shared out" (e.g. some of it could appear at c-structure, provided that c-structure is visible to whatever procedure is responsible for mapping to a semantic formalism). Discussion of this matter at the workshop revealed that a number of participant researchers would wish to argue for this latter stance. I have

no objection to it, but since it appears to be less straightforward than the "f-structure" assumption, I will adopt the latter in what follows. Further research is needed to establish which approach is the more useful.

A second working assumption I will make is that temporal (and related) information should be encoded in as language-independent a way as possible. This accords with the general spirit of LFG, and would make the representation suitable as input to, for example, structural transfer for machine translation, or to mapping to a semantic formalism such as Underspecified Discourse Representation Theory (UDRT). (See, for example, van Genabith and Crouch 1997 for details of such a mapping from f-structure to UDRT.)

My current aim, therefore, is to propose a method of encoding at f-structure, in as language-independent a way as possible, all the information that is pertinent to the temporal interpretation of discourse.

I will now turn to question (i), to consider how best this might be done.

2. The contribution of temporal information to the temporal interpretation of discourse

By 'the temporal interpretation of discourse' I mean such things as determining the temporal relations (succession, inclusion, overlap, etc) between events and states as presented by the text. Of course, a vast amount of work has been done on this subject (see Kamp and Reyle 1993 for a comprehensive summary), and I cannot hope to do justice to it here. Rather than try to do so, I will simply highlight a few significant findings, and draw out some implications for possible f-structure representation.

Since what I am calling 'temporal information' is often conveyed by tense and aspect features, I will from now on refer to such information as 'tense/aspect related' (or 'TA-related') information. It should be borne in mind, however, that across the languages of the world, TA-related information can be conveyed in a wide variety of ways (e.g. tense markers may be placed on NPs in some languages[1]), so I am not intending to set any limits on what can count as TA-related information.

2.1 Narrative movement

It is well known that in multi-sentence sequences, a series of "events" is generally interpreted differently from a series containing "states". For example, in the following sequence:

(1) Mary got up. She brushed her teeth.

[1] See, for example, Nordlinger and Sadler 2000.

we have two events, and the most natural interpretation is one where Mary's teeth-brushing follows her getting up. If, however, we replace the second sentence by a stative:

(2) Mary got up. She felt very ill.

then a second interpretation is available, where Mary's feeling ill begins before she gets up and continues afterwards. Another possible interpretation is that she begins to feel ill just as, or shortly after, she gets up.[2]

If the second sentence is a progressive, then only the "overlap" reading is possible:

(3) Mary got up. She was brushing her teeth.

can only mean that Mary is already brushing her teeth when she gets up.

Thus progressives are similar but not identical to lexical statives in terms of the temporal relations that they allow. Because of the similarity, the English progressive has been taken by some to be a stativiser (Vlach 1981, Moens and Steedman 1988). There are good arguments against taking this approach, however (see Smith 1991 and Glasbey 1994) and we will not adopt it in this paper (see Section 2.2 for further discussion).

Roughly speaking, we can say that two successive sentences that are non-stative and have simple (non-progressive) aspect give rise to a "forward movement" or "temporal succession" interpretation, whereby the event described second follows the one described first.[3] If the second sentence is a progressive, then only the temporal overlap interpretation is available. If the second sentence is stative (with simple aspect), then either of these interpretations is possible.

Clearly, then, in our TA representation we will need to include stative/non-stative and progressive/simple aspect information. In fact, since other work[4] has shown that the determination of temporal structure depends not only upon these distinctions but on Vendler class (the distinction between states, activities, accomplishments and achievements, as proposed in Vendler 1967), then it appears that we may need to encode Vendler class (otherwise known as aspectual class) in our representations.

The presence of such aspectual information at f-structure will be necessary in order to enable mapping to a UDRS, or to some other semantic formalism which similarly places restrictions on possible temporal relations between events and states.

---

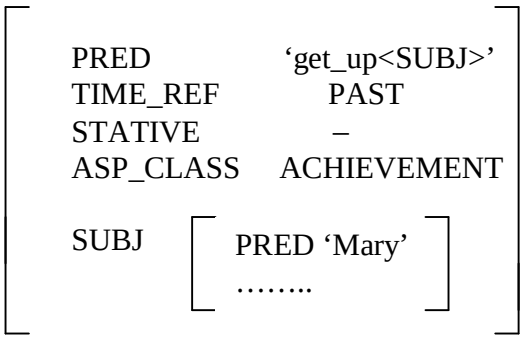
[2] Of course, world knowledge or contextual knowledge often helps to select whichever interpretation is most salient in a given context.

[3] Lascarides and Asher 1993 argue that the forward movement or "narrative progression" interpretation should be seen as a default, to be inferred in the absence of other information. Alternative interpretations, such as unordered lists and elaborations, are possible if certain other conditions are met.

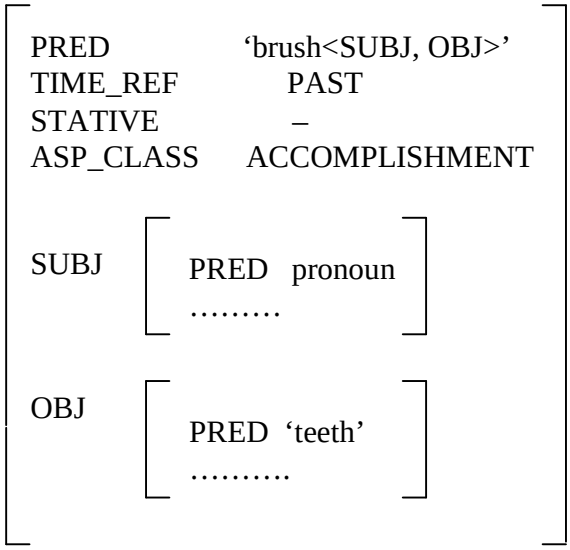
[4] See, Kamp and Reyle 1993 and Glasbey 1994, for example, for summaries.

A preliminary attempt at a suitable f-structure for example (1) might look as follows:

Mary got up:



She brushed her teeth:



NB: I have ignored the possessive pronoun in this representation.

Assuming a system that maps from these f-structures to an appropriate semantic formalism such as UDRT, the respective values of the ASP_CLASS features of the two clauses (ACHIEVEMENT and ACCOMPLISHMENT) would contribute to determining the temporal relationship between the corresponding event referents in the combined UDRS for the discourse.

There is a problem with the above, however, in that the determination of sentencial or clausal aspect is compositional (see, for example, Verkuyl 1972, Krifka 1992, Verkuyl 1993). That is, the aspectual class of the sentence 'Mary brushed her teeth' is not determined solely by the aspectual class of the verb 'brush'. Other constituents, such as the subject and object arguments and any prepositional and adverbial phrases, may also have an effect on the aspectual class of the sentence as a whole.

In some accounts (e.g. Moens and Steedman 1988), the lexical entry for a verb is given a "basic" aspectual class, which may then be modified in certain ways by other constituents. So we might classify the verb 'eat' as an activity. If this combines with an "unbounded"[5] object argument such as 'bread', then the aspectual class of the VP 'ate bread' remains an activity. If the object NP is bounded (e.g. 'an apple'), then the VP 'ate an apple' will be an accomplishment. The aspectual class of the sentence as a whole is only determined when we additionally take into account the properties of the subject NP. If the subject NP is bounded (e.g. 'John') then the aspectual class of the sentence 'John ate an apple' is an accomplishment. If the subject NP is unbounded (e.g. 'people') then the sentence as a whole is an activity.

Prepositional phrases (etc) may also have an effect, so that while 'John climbed' is an activity, 'John climbed to the top' is an accomplishment.

Deriving aspectual class in a non-monotonic way like this would be a problem for deriving f-structures in the standard information-preserving manner employed in LFG. It is possible, however, to see the determination of sentential/clausal aspectual class as an information-preserving process. Viewed this way, sentence aspectual class is composed from information contributed by the various constituents. The verb, the subject NP, the object NP(s), and any prepositional and adverbial phrases each contribute aspectual information, which is combined according to compositional rules to give the aspectual class of the whole sentence. This type of approach has been used primarily by Verkuyl (1972, 1993) and is also employed by Krifka (1992) and others. It has been embodied in a number of implementations, including a feature-based Prolog grammar described in Glasbey 1994. Rather than assigning a basic aspectual class to the verb, which is then modified by other constituents, a feature called TERM(INATIVE) is used. This means roughly the same as 'bounded', and the feature make take the value +, – or "unspecified". Examples of verb classifications are:

'eat': UNSPECIFIED
'stroke': –TERM
'like': –TERM

A verb like 'stroke' which is –TERM will always combine with its arguments to give a –TERM VP, irrespective of the TERM properties of those arguments. A verb which is unspecified for TERM will combine to give a VP (and finally a sentence) the value of whose TERM feature is determined by the TERM values of its object and subject

[5] Various related terms are used in the literature. An closely-related alternative (used in Krifka 1992) is 'cumulative'.

arguments and any prepositional phrases (etc). Thus, for example, 'John ate an apple' would be +TERM because 'ate' has its TERM value unspecified, 'John' is +TERM and 'an apple' is +TERM.[6]

Clearly, across different languages the information contributed to overall sentence aspectual class by the various sentential constituents would vary (i.e. the mapping from c-structure to f-structure would differ). The aim, however, would be to produce a fully language-independent set of aspectual features. Much more work is needed, of course, to develop such a set of features.[7]

Examples of f-structures based on the TERM feature will be given at the end of Section 2. I will now move on to consider the progressive and related constructions.

2.2 Progressives, imperfectives and other unfinished business

We have already looked briefly at some of the properties of the English progressive and seen that it affects the temporal properties of discourse. It seems reasonable, therefore, to include a feature +/–PROG in the f-structure for the VP. This is problematic, however, given our ultimate aim of language-independence, because the English progressive has a number of properties which distinguish it from imperfectives in other languages.

One very obvious difference is that English progressives combine only reluctantly with stative verbs (? 'Mary was loving John') whereas in French, for example, the imparfait combines readily with stative verbs ('Marie aimait Jean').

There are implications here for aspectual composition and discourse structure, so we need to be careful not to use a single feature PROG cross-linguistically. An alternative suggestion is given below.

First, we should consider in more depth the proposal (mentioned earlier) made by some researchers to treat the English progressive as a stativiser. Moens and Steedman (1988) are among those who adopt this approach, but it has been shown to be unsatisfactory in a number of respects. Smith (1991) gives a range of cross-linguistic data which shows that the English progressive is not equivalent to the cross-linguistic property "imperfective".[8] Treating progressives as statives causes problems in English, too, as is shown by Glasbey

[6] It should be noted that the issue has been somewhat simplified here for the sake of exposition. Care needs to be taken, for example, with scope ambiguities in sentences like 'Everyone climbed the mountain in an hour'.

[7] It should, however, be noted that Smith (1991) bases her account of aspectual composition (which is closely related to the one just given) on data from a range of languages including English, French, Russian, Mandarin Chinese and Navajo – so at least some first steps towards language-independence have been made.

[8] There is, of course, a major assumption here that the notion "imperfective" across languages can be captured by a single feature.

(1998) in an analysis of ‘at the time’ and ‘at the same time’. The data in question is as follows:

4. Daniel climbed Ben Nevis.

5. # Jonny climbed Snowdon at the time.

6. Jonny was climbing Snowdon at the time.

7. Jonny was a young boy at the time.

8. Jonny climbed Snowdon at the same time.

9. Jonny was climbing Snowdon at the same time.

10. # Jonny was a young boy at the same time.

Sentence (4) is to be followed by one of sentences (5) – (10). It will be seen that with ‘at the same time’ the progressive in (9) does not pattern with the stative in (10). A detailed analysis is given in Glasbey 1998 – here I will simply say that in order to account for the data, it is necessary to distinguish progressives from (lexical) statives, and this runs counter to an account that treats the progressive as a stativiser.

Given that (i) it appears to be unsatisfactory to treat progressives as states, and (ii) it appears unsatisfactory to treat the English progressive in the same way as a cross-linguistic imperfective, I propose to use two features. One is PROG, which may take values of {+/–} and will be used for the English progressive (and any equivalent constructions found in other languages). The second feature, which I will call VIEWPOINT, after Smith 1991, is based on the PERFECTIVE/IMPERFECTIVE distinction, used widely in studies of the Slavic languages and others, and adopted by Smith (1991) in her cross-linguistic analysis of aspect.[9]

2.3 Perfects

Another TA feature we will need to encode at f-structure is PERF(ECT), as in:

(11) Em has climbed Ben Nevis.

The need for this feature seems uncontroversial, as the perfect (in English and a range of other languages) clearly contributes to the temporal structure of discourse. I will assume

[9] In fact Smith proposes a third value “NEUTRAL”, which she argues is needed for some of the languages in her analysis. I will retain an open mind here on whether this third value will be required. Once again, further research is needed.

here for simplicity that a single PERF feature (taking values +/–) is sufficient cross-linguistically, although this may turn out not to be the case.

2.4 Time reference: past, present and future

Up until now I have concentrated on aspectual features as opposed to tense ones. Of course, many languages have constructions which encode tense and aspect features simultaneously, and sometimes temporal reference is underspecified. It is clear, however, that we will need some way of expressing, at least, whether the temporal reference of a sentence is PAST, PRESENT or FUTURE (or some underspecified combination of these values). I propose therefore, to incorporate a feature TIME_REF which can take the above values.

2.5 Summary: Proposed TA-related features for f-structure

In sections 2.1 – 2.4, I have proposed a number of TA-related features that will be needed in our f-structures. These are STATIVE {+/–}, TERM {+/–}, PROG {+/–}, VIEWPOINT {PERFECTIVE/IMPERFECTIVE}, PERFECT {+/–} and TIME_REF {PAST/PRESENT/FUTURE}.

In order to illustrate the use of the above features, I will now give a number of lexical entries for verbs, and an f-structure for an associated sentence, to indicate how the required feature composition is achieved.

Lexical entries for verbs:[10]

ate :	V(↑ PRED)	=	‘eat < (↑ SUBJ)(↑ OBJ) >’
	($\uparrow_{\mu}$ AUX)	=	–
	($\uparrow_{\mu}$ FIN)	=	+
	(↑ TERM)	=	
	(↑ PERFECT)	=	–
	(↑ TIME_REF)	=	PAST
	(↑ VIEWPOINT)	=	PERFECTIVE
	(↑ PROG)	=	–
	(↑ STATIVE)	=	–

Note that the lack of a value for particular feature means that the feature value is unspecified.

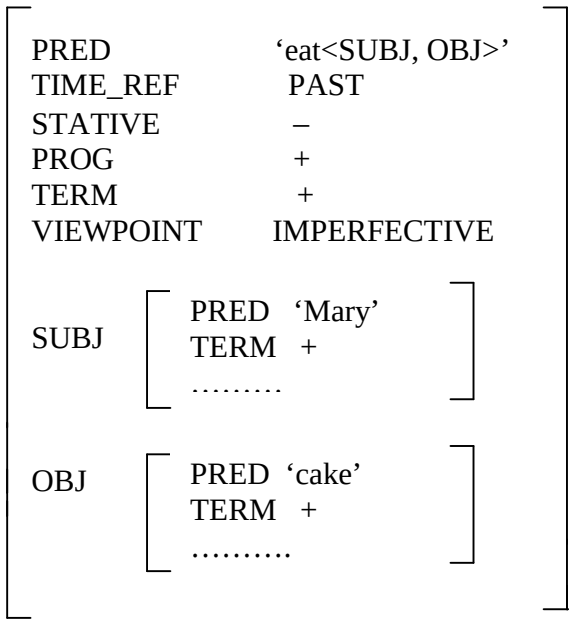
[10] I use LFG terminology employing a level of morphological representation (m-structure) here, following Butt et al 1996.

stroked: V(↑ PRED) = ‘stroke < (↑ SUBJ)(↑ OBJ) >’
($\uparrow_{\mu}$ AUX) = –
($\uparrow_{\mu}$ FIN) = +
(↑ TERM) =
(↑ PERFECT) = –
(↑ TIME_REF) = PAST
(↑ VIEWPOINT) = PERFECTIVE
(↑ PROG) = –
(↑ STATIVE) = –

eating: V(↑ PRED) = ‘eat < (↑ SUBJ)(↑ OBJ) >’
($\uparrow_{\mu}$ AUX) = +
($\uparrow_{\mu}$ FIN) = +
(↑ TERM) =
(↑ PERFECT) =
(↑ TIME_REF) =
(↑ VIEWPOINT) = IMPERFECTIVE
(↑ PROG) = +
(↑ STATIVE) = –

was: V_{aux} ($\uparrow_{\mu}$ AUX) = +
($\uparrow_{\mu}$ FIN) = +
($\uparrow_{\mu}$ DEP AUX) = –
($\uparrow_{\mu}$ DEP VFORM) = PROG
(↑ PERFECT) = –
(↑ TIME_REF) = PAST
(↑ VIEWPOINT) = IMPERFECTIVE
(↑ PROG) = +

F-structure for 'Mary was eating a cake':



3. Additional TA-related features

I have sketched a set of possible TA-related features and shown how they might be used to obtain compositionally-derived f-structures. Before closing, I would like to consider whether there are any additional TA-related features that should be included.

One clear case seems to be the need to encode in some form the distinction between accomplishments and achievements (or some closely-related distinction, as explained below). This need is evident when we look at English sentences containing 'when' constructions. Sandström (1993) draws attention to the following type of example:

(12) When Mary arrived at the station she bought her ticket.

(13) ?When Mary drove to the station she bought her ticket.

Sandström points out that in (12), the relation between the arriving event and the ticket-buying event is one of temporal succession. The effect is similar to that conveyed by the two-sentence sequence:

(14) Mary arrived at the station. She bought her ticket.

In (13), however, it is very difficult, if not impossible, to get this interpretation, and (13) sounds rather strange as a result. (Another reading, perhaps best described as an "occasion" reading[11], can be obtained by some speakers, but the temporal succession reading is clearly absent.)

Sandström identifies the difference between (12) and (13) as being due to the fact that 'Mary arrived at the station' is an achievement while 'Mary drove to the station' is an accomplishment.

Glasbey (1995, 2001) argues that the crucial distinction is not one of "punctuality" (the standard difference between achievements and accomplishments) but rather involves properties of the subject NP. Glasbey's arguments are based in part on examples such as:

(15) When the soup cooled down we drank it.

And:

(16) When the snow melted away we went for a walk.

The standard tests for aspectual class (see Dowty 1979) classify 'the soup cooled down' and 'the snow melted away' as accomplishments – but the temporal succession reading is readily available. Glasbey (2001) proposes that the criterion for the temporal succession reading in this construction is a "thematic" property of verb's subject NP argument, and concerns whether the subject is described as undergoing a change of state. This property, which she calls SUBJECT STATE TRANSITION or SST, can be seen as a feature of the subject NP, which is placed there by a property of the verb which is clearly distinct from considerations of punctuality.

I therefore propose an additional feature, called SST {+/–}, to be placed on the subject NP of the sentence, the value of the feature being derived from a lexical feature of the verb. The verb 'arrive', for example, which has the value +SST, would have the following feature specification as part of its lexical entry:

(↑ SST) = +

while, in contrast, a verb like 'eat', which has the value –SST, would have the feature specification:

(↑ SST) = –

The value of the SST feature would then be carried up to clause level.

It is interesting to note that here we have a feature which intuitively belongs to the subject NP and yet affects the TA properties of the clause as a whole. This is reminiscent in some

[11] A paraphrase of this reading would be 'On the occasion that Mary drove to the station she bought her ticket'.

ways of the way that tense information in some languages is conveyed (at least in part) by tense markers placed on NPs.[12] One noteworthy difference, however, is that in the SST case, the value of the SST feature for the clause is derived from the SST value of the verb, rather than some feature marked on the subject NP. It would be interesting to see if there are any languages that mark SST (or similar) information either fully or partially by inflection of the subject NP.

Again, further work is needed, but we have shown, at least, that the list of TA-related features suggested above is incomplete. It seems highly likely that more such TA-related features will be uncovered. The challenge is, as I said earlier, to make the set of TA-related features as language-independent as possible, while, of course, encompassing all the data.

4. Conclusion

I have shown that in a feature-based linguistic framework such as LFG, in order to determine the temporal structure of discourse we need to identify a set of cross-linguistically valid TA-related features.

I have suggested that in LFG it may be appropriate for such a set of features to be encoded at f-structure. This proposal is, however, based on the assumption, stated earlier, that f-structure constitutes the sole input to semantic interpretation. This assumption appears to be somewhat controversial in the LFG community, and clearly further work is needed to clarify and resolve the matter.

Having reviewed the effects of a range of TA-related features on the temporal structure of discourse, I have proposed a set of such features that aims as far as possible to be language-independent (and unquestionably falls short of that aim). I have also outlined briefly how the required f-structures could be derived from lexical entries for verbs and other constituents, in an information-preserving manner.

Much further research is needed, of course, both to work out the fine details of the above sketch, and to establish, as far as possible, a fully language-independent set of TA-related features that will do the work required for discourse interpretation, translation between languages, and whatever else we may ask of it.

[12] An example is the Arawak language Chamicuro, spoken in Peru (discussed in Nordlinger and Sadler 2000), where clausal tense information is encoded on the definite article.

References

Butt, M., Niño, María-Eugenia, Segond, Frédérique (1996). Multilingual processing of auxiliaries within LFG. *In* D. Gibson (ed) *Natural Language Processing and Speech Technology: Results of the 3rd Konyens Conference,* pp 111-122. Mouton de Gruyter.

Dowty, D. (1979). *Word Meaning and Montague Grammar*. Dordrecht : D. Reidel.

Frank, A., Zaenen, A. (draft). Tense in LFG: Syntax and morphology. Xerox Research Centre Europe, 6 chemin de Maupertuis, 38240 Meylan, France.

Glasbey, S. (1990). *Tense and Aspect in Natural Language Processing – A Situation Theory Approach.* MSc thesis, Department of AI, University of Edinburgh.

Glasbey, S. (1994). *Event Structure in Natural Language Discourse.* PhD Thesis, Centre for Cognitive Science, University of Edinburgh.

Glasbey, S. (1995). 'When', discourse relations and the thematic structure of events. Workshop notes of the 5th International Workshop TSM'95, Château de Bonas 23-27 June 1995. Institut de Recherche en Informatique de Toulouse – Université Paul Sabatier, INPT and CNRS. Equipe de Recherche en Syntaxe et Sémantique – Université Toulouse-Le-Mirail and CNRS.

Glasbey, S. (2001) Event structure, punctuality and 'when'. Draft, University of Dundee.

Glasbey, S. (1998). Progressives, states and backgrounding. In S. Rothstein, ed, *Events and Grammar,* pp. 105-124. Kluwer Academic Publishers, Dordrecht/Boston/London.

Kamp, H. and Reyle, U. (1993). *From Discourse to Logic: Introduction to the Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht: Kluwer.

Krifka, M. (1992). Thematic relations as links between nominal reference and temporal constitution. In Sag, I., Szabolsci, A. (eds), *Lexical Matters.* CSLI publications, Stanford, Ca, USA.

Lascarides, A., Asher, N. (1993). Temporal interpretation, discourse structure and commonsense entailment. *Linguistics and Philosophy* 16 pp 437-493.

Moens, M. and M. J. Steedman (1988). Temporal ontology and temporal reference. *Computational Linguistics* 14(2), 15-28.

Nordlinger, R., Sadler, L. (2000). Tense as a nominal category. *In* Butt, M. and Holloway King, T. (eds), *Proceedings of the LFG00 Conference*, University of California, Berkeley: CSLI Publications.

Sandström, G. (1993). When-clauses and the temporal interpretation of narrative discourse. *Technical Report 34, Department of General Linguistics, University of Umeå.*

Smith, C. (1991). *The Parameter of Aspect.* Kluwer Academic Publishers, Dordrecht.

Van Genabith, J., Crouch, R. (1997). On interpreting F-structures as UDRSs. In: ACL-EACL-97, Madrid, Spain, Proceedings of the Conference, 1997, pp. 402--409.

Vendler, Z. (1967). Verbs and Times. In *Linguistics in Philosophy*, pp 97-121. Ithaca, N.Y: Cornell University Press.

Verkuyl, H. J. (1972). *On the Compositional Nature of the Aspects.* Dordrecht: D. Reidel.

Verkuyl, H. J. (1993). *A Theory of Aspectuality. The Interaction between Temporal and Atemporal Structure.* Cambridge Studies in Linguistics, Vol 64. Cambridge: Cambridge University Press.

Vlach, F. (1981). The semantics of the progressive. In P. J. Tedeschi and A. E. Zaenen, eds., *Syntax and Semantics*, Vol 14: *Tense and Aspect*, pp 271-292. New York: academic Press.