AVG 2.0

Cross-linguistic Reference Grammar

Final Report

John Peterson

Table of Contents

Acknowledgements

The "Cross-Reference Grammar" project (CRG), also known under the German name AVG 2.0 (Allgemein-vergleichende Grammatik, 2. Teilprojekt), has been going on for a number of years and the present paper represents its final achievements.

It was merely by chance that I was the last in a line of researchers who worked on this project so that my name turned out to be the one given as the author of this paper. Although I had the honor of writing up a summary of the work conducted in the project, the fact that it is my name which appears as that of the "author" here must not be taken as meaning that the following pages represent merely my own work.

I was not fortunate enough to meet many of my predecessors, most of whom I only know through their publications. These include Milly Brunello, Christian Strömsdorfer and Roman Pichler, who produced much important work which has found its way in many hidden forms into the present grammar format. Others, such as Tsuyoshi Takizawa, were no longer actively involved in the project when I arrived but nevertheless gladly gave of their time to answer my questions and assist where they could.

A special word of recognition goes to the following co-workers, who helped make my stay in the CRG project especially productive:

- First, those who made sure that the computer did what we wanted it to do, and never tired of our often naïve questions as to why this or that was not possible. These include my present colleague, Irfan Bilgili, who actually got the program to work with real examples, his immediate predecessor Stefan Gering, and the creator of the basic program, Matthias Nickles. Both Stefan Gering and Matthias Nickles, while no longer involved in the project, unhesitatingly devoted much of their time (and patience!) to those of us still in the project and answered countless emails and phone calls.

- A good deal of the credit here also goes to the student researchers that I had the good fortune of working with: Felix Weigel, who worked wonders on the computer, and Eleni Kriempardis, who not only checked countless parts of the grammar format against examples taken from an array of grammars but also provided much-needed constructive criticism, pointing out where we had overlooked various possibilities. She also kindly put together the bibliography for this paper.

- Final recognition must go to the two colleagues with whom I worked most closely on the present project, Ellen Brandner and Dietmar Zaefferer. It was Ellen who originally had the brainstorm of using a tree format to describe grammatical phenomena with different kinds of nodes, and even a cursory glance at the following pages will show that the present description would not have been possible without this idea. She is also responsible for most of the work on interrogatives, the uppermost levels of the grammar format, and contributed to countless other parts of the grammar. Luckily, the tree format had already been implemented before my arrival upon the scene, so that I was able to delve directly into various aspects of the grammar. While the reader of the present paper may consider this achievement to be "obvious", my predecessors were forced to try several different possibilities before chosing the present one. Had this not been successfully completed before I joined the project, I too would have had to spend many weeks or even months with this topic, which is anything but trivial. Unfortunately, Ellen's work in the project ended much too soon, so that I have had to see the practical side of the project to its end

alone. Nevertheless, while many of the actual "nodes" in the present format carry my signature, I would never have been able to see the project to its conclusion without all the work of my predecessors, which I gratefully acknowledge here.

- Finally, Dietmar Zaefferer, the director of the project, who contributed directly in many ways to the present structure of the grammar. He was always glad to discuss the various areas of the grammar in detail and suggest alternatives, and much of the structure of the present format would have looked very different had it not been for his advice. His knowledge of both the variety of human languages and at the same time his overview of the phenomenon involved have often served to simplify the present grammar format and also to include much information which would have otherwise gone unnoticed.

To all who have contributed to this end in one fashion or another I again express my deepest gratitude.

Thus, while it is my name which appears on the cover of this paper, it is in reality all of us in the project, both present and past members, who now proudly present the culmination of their work to the linguistic community.

1. Introduction

The basic philosophy behind the Cross-Reference Grammar (CRG) project is as follows: As advances are continually being made in linguistic theory, and as our knowledge of previously undocumented languages increases, it becomes all the more important to have a medium of documenting the current state of our knowledge of languages in a uniform and consistent manner. This is of use to both theoretical linguists, positing universals of human language, as well as to field-workers, who are very likely to come across forms in languages which they cannot immediately identify. Such a medium would thus on the one hand serve to either support or disprove claims of a theoretical nature by providing sufficient examples for a given phenomenon as well as provide those with more practical needs with information on a large number of languages, any one of which may have a grammatical category similar to the one in question.

There are a number of advantages to an electronic format for such a medium as opposed to the traditional book format. These include the following, among others:

- The immediate adaptability to advances in linguistic theory. Where new insights are gained which were not known at the time of the composition of the present grammar format, these can easily be incorporated into the format without the usual trouble of republishing the grammar or the data it contains.

- Large amounts of data can be scanned for various desiderata quickly and easily.

- Constructions in two or more languages (up to the entire number of languages entered into the databank) can be compared with one another from the viewpoint of either their content, their form, or both.

- When desired, the computer allows the user to determine statistical data, such as the degree of fusion, isolation or agglutination of a respective language, which kind of information tends to be coded by which type of marking, etc.

- The data for any particular language can be continually updated, as the author's knowledge of the language increases.

- Where two (or more) languages, or two (or more) areas of a grammar are treated similarly but differ slightly in certain points, it will be possible to determine exactly where these two differ and to what extent they have the same structure.

In addition to the above-mentioned points, to which many more could be added, the electronic format requires that all data be entered using a consistent system of representation for all languages. Such a system has already been developed and used with considerable success, namely Comrie & Smith (1977), which currently serves as the basis for an increasingly popular series of language grammars. Nevertheless, this book format suffers from the problems alluded to above as well as other more "practical" problems, such as the need to choose between dividing the grammar into morphology and syntax, although the two are by no means easily separated. This also requires the author to enter much information separately which is already found elsewhere in the grammar.

For example, while discussing a grammatical point, such as "possession", the author will typically include an example. This example will probably have one or two complements and

perhaps also a finite verb. In book format, this information is entered in the form of an example and given a number after a prose description. However, if the author intends to include verb and noun paradigms in the book, which is probably what any user of the grammar expects, this information must be entered separately in paradigm form, although many parts of the paradigm for several verbs and nouns can be found in the examples. In computer format, this information can merely be entered once and glossed correspondingly - the computer is then capable, on the basis of the gloss, of recognizing e.g. whether a nominal or verbal form is part of a paradigm, thus gradually building up paradigms on the basis of the examples. Also, any given example is likely to include information on a number of areas of the grammar, such as the structure of noun phrases, word order, tense marking, intonational patterns, etc.

While the present grammar owes much of its form and inspiration to Comrie & Smith (1977), it differs in a number of points from this for several reasons. The most important is perhaps the medium chosen. For example, it is possible in book format to write a question such as Comrie & Smith's point 3.2.6.3. "Are there any restrictions between syllable initial units or clusters and syllable final units or clusters? Describe these." Although such a criterium is important, it does not make sense to pose this type of question in an electronic format, where the author must be presented not only with the question but also with all possible answers, at least to the extent that we, the members of the CRG-project, were able to locate data on this. This has required that we depart quite often from the format in Comrie & Smith. We have also often taken the liberty to diverge where we felt this was necessary or that it would be advantageous. Thus, although the criteria in Comrie & Smith (1977) were taken as a useful guideline for our work on many occasions, we have not taken it as the basis for our work.

Finally, it must be stressed that the current grammar format is only the BEGINNING of a greater enterprise. Over the years, the co-workers of the project have researched virtually all aspects of grammar from as many languages as possible, both spoken and signed. Our aim was to document the extent of variation found both in human speech and in human sign languages. While what is presented here is the culmination of our work, it goes without saying that the relatively small number of co-workers of the project will not be able to anticipate all possible grammatical categories as well as all possible means of expressing these. It is here that the electronic medium has a decisive advantage: What is not found here can easily be added without the data having to be re-entered.

We have also not been able to discuss each and every aspect of the grammar with specialists in these fields. To do so would have required much more both of our own time as well as theirs. Despite the shortcomings which are undoubtedly present in the first edition of any pioneering work (this format is to our knowledge the first of its kind), it should prove useful and we hope that it will serve as the basis on which future work can be conducted. Its goal is admittedly a very high and idealistic one and will certainly never be completely perfected. However, we firmly believe that the benefits it will provide in the future, even in its present form, will show that this idealism was well founded.

2, Basic Assumptions

The present grammar format makes as few assumptions about the nature of human language as possible. We do not, for example, assume that every language has nouns, verbs, postpositions, case markers, adjectives and adverbs, although we must certainly allow for this kind of information. These are questions which must receive an empirically based answer, and the debates presently being conducted concerning the universality of "nouns" and "verbs" show that this discussion still has a long way to go before being completely resolved.

We do however assume that every human language has means for REFERRING and PREDICATING, regardless of their type. Further, we assume that every language will have some means of MODIFYING these referentially or predicatively used expressions. We also assume that every human language will differentiate between declarative and interrogative utterances, commands, interjections, etc., but we make no assumptions as to how these are expressed.

Our basic assumptions allow us to describe these expressions as follows. If the use of an expression REFERS to an entity, then this referentially used expression has, of course, a linguistic form and what it refers to can be described with respect to a number of criteria. Let us begin with the description of the semantic content being referred to by the expression.

First, this referentially used expression may refer to a physical or a non-physical entity. Let us assume for the sake of illustration that it refers to a physical entity, such as a table. Since this is a physical entity, it can also be described with respect to its animacy features (i.e., animate or non-animate), etc. Further, unless we are dealing with an entry for the lexicon or making an entry for a nominal (etc.) paradigm, the entity referred to will have a place in the discourse-pragmatic context of a discourse. Thus, the entity being referred to can be mentioned for the first time, or it is a unique entity with which both speakers (or signers) are expected to be familiar (the sun, the moon, etc.), or it can have been mentioned in the preceding discussion once or perhaps repeatedly. On the basis of criteria such as these, a sufficiently comprehensive description of the extralinguistic entity being referred to as well as its presentational status in the current context can be achieved.

Of course, in order for this expression to be part of a language, it will necessarily also have a linguistic form. This can be a spoken word or phrase, a signed word, or, e.g. for a language which is no longer spoken, a written word. If we are dealing with a spoken language, this may be a single word, a group of words – for example a semantic head with a modifier, a quantifier and a determiner - or an entire clause. Here, we can offer a number of choices on the CATEGORIAL STATUS of the expression. If it is a word, we can enquire as to its word class, i.e., "noun", "verb", etc. No assumption is made that the referential expression must be a noun (or for that matter, a noun, verb or adjective). All possibilities known to us are offered whenever we enquire as to the word class of any linguistic form, so that even the highly unlikely possibility of the referential expression being a postposition is automatically offered. Proceeding in this fashion, it is highly unlikely that we will run into a situation where an author is faced with trying to enter information which is not enterable in the current structure. However, even in the unlikely event that this will happen, the possibility that the status is "Other" (see below) is offered throughout the format.

Returning to the form of our referential expression, if it is a noun, it may be marked for any number of different categories, including gender, number, pragmatic status, possessive marking, etc. To this end we have included a section of the grammar, to be discussed below, which includes ALL grammatical categories known to us, including TAM, gender, "noun

class", number, person, pragmatic marking, possession, diminutive marking and many, many more categories. This choice is offered in ALL contexts where overt marking is found, so that any part of speech may - at least theoretically - be marked for any type of information. Finally, we have taken note of all means of marking which we were able to catalogue in a large number of languages - from prosodic marking, pre- and postpositions, affixes (prefixes, suffixes, circumfixes, transfixes, infixes), word order changes, reduplication, etc. - and include these all under the term "Means of Marking", so that again in theory any grammatical category can be marked by any known form of marking.

By proceeding in this fashion, we have been able to keep our theoretical assumptions to an absolute minimum. As the overwhelming majority of human languages have as yet not received adequate documentation, this is undoubtedly a wise choice, as it means that no assumptions based on particular theoretical expectations have been built into the grammar, neither in terms of which categories are marked on which parts of speech, nor in terms of how these respective categories are marked. Linguistic field work on hitherto undocumented languages has often produced surprising and exciting results, and a medium of description such as the present one must be flexible enough to deal with this data to the extent that this is currently possible.

3. Methodology

The methodology involved is based on the strategy described in the preceding section and is very simple. The entire grammar makes use of three logical possibilities for presenting information and uses a "tree" format to do so.

The three possibilities represent three different types of edges in the tree:

- AND-edges or "required" edges (also HAVE-edges, since they lead from nodes that represent a phenomenon to nodes that represent its aspects). These are edges which must be followed further by the author entering the data.

  As a simple example of this we can take negation. When we are describing a case of negation, we know that it MUST have a content and a form. Hence, after affirming that he is dealing with a case of negation, the author is then REQUIRED to specify both negational content and negational form.

  Required edges are represented by a solid line in the grammar, i.e. ——— .

- EXCLUSIVE OR-edges or "exclusive" edges (also BE-edges, since they lead from nodes that represent a phenomenon to nodes that represent all its disjoint kinds). Here, the author coming from the dominating node is presented with two or more possibilities only one of which may be chosen. At least one MUST be chosen, but never more than one.

  To return to our example above, our author has determined that his or her example is a case of negation, he has hence been required to describe both negational content and form and he decides to describe negational content first. A further AND-edge leads him to the notional kind of negation. It is here that the exclusive edges play a role. The author is now presented with a choice between contradictory and contrary negation. Here the author must chose one and only one of the alternatives offered.

  Exclusive edges are represented by a checked line, i.e. - - - - - - - .

- OPTIONAL edges. These are possibilities which may or may not be needed and any combination of these - ranging from all possibilities to none - may be chosen. This type of edge is found for example where we know from experience that one or more phenomena are found in language in a certain environment but are not necessarily present in all languages.

  As an example, let us again return to the case of negation, mentioned above. When the author decides to proceed to the description of negational form, he is presented with a required edge leading to a node labeled Primary Negation Marking, but also with an optional edge leading to Sedondary Negation Marking which is relevant of course only if in the case to be described there is such a secondary marking.

  Optional edges are represented by a dotted line, i.e. ..... .

Taken together, these three possibilities allow us to deal with all aspects of grammar - both form and content. They allow us to present all possibilities for form and content in a logical fashion and ensure that data are entered into the grammar in a uniform manner. The fact that

the author proceeds in this fashion through a kind of "tree" also means that the data may be directly compared with each other, facilitating both inter- and intra-language comparison.[1]

[1] It should be noted here that it would also have been possible to use only two nodes, namely "required" and "exclusive". While this would have been possible, it would have not only increased the size of the grammar considerably, but would have greatly decreased the user-friendliness of the system. This can be illustrated by means of a simple example:
For a particular question, there are three known possibilities, a, b and c. Furthermore, we know that a, b, c and any combination of these may be appropriate in a certain situation. By not making use of the optional node we would be forced to enter the following possibilities:

1. a only
2. b only
3. c only
4. a and b
5. b and c
6. a and c
7. a, b and c

If instead of only a, b and c we also had d, e, f, g and h, it is easy to see that such a system quickly becomes unpractical. Hence, the current system with the three node types was chosen. Using optional nodes, a, b and c are all presented and any combination thereof may be chosen.

4. Language Description

In the following, the major sections of the grammar will be discussed in some detail. It will of course not be possible to discuss each and every node in the grammar, although all major principles will be discussed as well as why we have chosen certain criteria for the grammar.

The uppermost node of the grammar is immediately followed by three sections: "Language Identification", "Description Identification" and "Described Phenomenon", the first two of which may be discussed together. The three sections are "required" nodes, i.e., the author is required to answer questions pertaining to all three areas. This can be represented as follows, where the sign (+) indicates that the final node has not yet been reached:

CRGD (Cross-Reference Grammar Data)

- **Language Identification (+)**
- **Description Identification (+)**
- **Described Phenomenon (+)**

The following section takes a look at the two uppermost nodes here, while the remainder of the present paper deals with the "Described Phenomenon".

4.1. Language Identification, Description Identification

The exact structure of the two uppermost nodes is as follows, where all nodes are of the required type:

Language Identification

- **Names**
- **Genetic Relations**
- **Language Code**
- **Main Geographical Area of Use**
- **Number of Speakers**
- **Documentation Urgency**

Description Identification

- **Author**
- **Date**

Here we are interested only in a very general outline of the language involved and in establishing its unique identity for the databank, as well as the identity of the author and the date of compilation of the grammar.

"Names" refers to ALL names which are in current use to denote the language. In order to uniquely identify the language involved, a three-letter code, taken from the Ethnologue: Languages of the World (2001) will be entered under "Language Code" (where possible).

Under "Documentation Urgency", all information concerning the status of the language and its speakers which the author deems relevant may be included, such as "official language", "no longer learned by childeren", etc.

4.2. Decribed Phenomenon

Let us now return to the uppermost node, the CRGD node, under which all data from all languages is ultimately subsumed:

CRGD (Cross-Reference Grammar Data)

- **Language Identification**
- **Author Identification**
- **Decribed Phenomenon (+)**

Under the lowest node, "Decribed Phenomenon" we find the following:

Decribed Phenomenon

- **Morphosyntax and Semantics (+)**
- **Lexical Inventories (+)**
- **Phonological Inventories (+)**

In the following chapters, each of these lower nodes will be dealt with in detail. Let us begin here with the uppermost of these, "Morphosyntax and Semantics", which forms the bulk of our study.

5. Morphosyntax and Semantics

When we click on this node, we find the following:

Morphosyntax and Semantics

- **Example Entry (+)**
- **Description of Example (+)**

If we are to discuss morphosyntax and semantics, we of course need to know what it is that we are discussing the morphosyntax and semantics OF – i.e., we need a concrete example. Hence, both of the nodes here are of the required type.

"Example Entry" represents – as the name suggests – the entry of a linguistic expression into the data bank. The linguistic description of this example is to be found under "Type of Example", which begins with the question of whether the present example belongs to the lexicon or whether it is a "full" example (generally a complete sentence) and thus belongs to the grammar.

In the following discussion, we will be discussing in detail the entry of linguistic data into the databank and how this is structured. A discussion of the complete grammatical description of the data entered will be presented in the remainder of the paper.

5.1. Structure of the Data Entry

Altogether, a total of up to eleven levels will be visible to the user of the data base. There will also be two hidden levels which are essential for an automatic comparison of various data. These levels will be generated automatically on the basis of the primary data.
The levels are the following:

Visible:

Level	Description
+6	indigenous, non-Roman script
+5	standardized transcription, transliteration or Roman-based orthography
+4	sentence-level suprasegmentals
+3	phonetic transcription
+2	phonological transcription
+1	morphophonemic structure
-1	gloss
-2	complex morphology
-3	sentence constituents
-4	literal translation
-5	free English translation

Hidden (automatically generated, visible upon request):

-1'	'Templates' of Level -1
-2'	'Templates' of Level -2

In many cases, some or even most of the visible levels will not be filled in. Some examples:

- Languages which do not possess a written tradition will of course not have an indigenous script (Level +6) or a standardized, Roman-based orthography (Level +5)

- A detailed phonetic transcription (Level +3) or the representation of sentence-level suprasegmentals (Level +4) will not generally be possible for languages which are no longer spoken.

- An literal translation (Level -4) will be unnecessary if the author believes that the free English translation (Level -5) is easily obtainable from the gloss (Level -1) without extra comment.

However, at least three levels will *always* be mandatory for any entry of primary data:

- The morphophonemic structure (Level +1)

- The gloss (Level -1)

- The free English translation (Level -5)

Level +6 Indigenous, non-Roman script

This line is intended to allow users the opportunity to see what kind of non-Roman writing system has been developed for the language. Examples here include the Chinese characters, the Arabic alphabet or any of the various South Asian writing systems (Devanagari, etc.).

The information provided in this level will not be further processed by the computer and is merely intended to be used visually. It will also not be searchable.

Level +5 Standardized transcription, transliteration or orthography

Here the author has the opportunity to give one (and only one) standardized, Romanized system of transcription or transliteration, if there is one, or a standardized Roman orthography, such as for English, German, Finnish, etc.
For example: Chinese. Here the author could chose to give e.g. the Pinyin equivalent to the Chinese characters in Level +6. That is, although there are quite a number of systems of transcription for Chinese (e.g. Pinyin, Yale or Wade-Giles), there is only room for one system of transcription. Hence, it is likely that the author would chose the most commonly used system, in this case, Pinyin.
Another example is the standardized system of transliteration among Indologists in current use for South Asian languages. That is, as many South Asian languages have their own alphabetical (and largely phonemic) writing systems, a system has been developed for these writing systems which replaces each character with a combination of Roman letters and diacritics.
As with level +6, the information contained at this level will not be further processed by the computer but simply displayed. It will also not be searchable.

Example:

Hindi:	Level +6	मेरा नाम राम है।
	Level +5	*merā nām rām hai.*

Finally, this level may also be used for languages which do possess a standard, Romanized orthography, such as English, German, Irish, etc. This is especially important for those languages, such as English, French or Irish, whose standardized orthography deviates to a large extent from the actual pronunciation.

Level +4 Sentence-level suprasegmentals

At this level, sentence-level suprasegmentals, such as intonation in spoken languages, will be represented following the conventions laid out in Hirst & Di Cristo (1998).

Each intonation unit will be marked off at the beginning by the sign '[' and the end of the intonational unit will be marked by the sign ']'. This closing sign (']') will reserve a space in Level -1, the gloss (as well as in the automatically generated Level -1'), so that grammatical information represented by suprasegmentals can also be glossed.

Example: The intonational pattern of the German sentence "Mein Name ist Stefan", uttered upon entering a room.

Level +5	Mein Name ist Stefan.				
Level +4	[→	→	→	⇑ >	]
Level +2	m͡aɪn	ˈnaː.mə	ɪst	ˈʃtɛ.fan	
Level -1	1.S.POSS.NOM.S.M	name.NOM.S.M	COP.NPT.3.S	"Stefan"	[SENTFOC.FIN]
Level -5	'My name is Stefan.'				

Level +3 Phonetic transcription

This level is intended solely as the opportunity for the author to provide any information s/he feels should be provided on the exact pronunciation. That is, this level is intended as a "narrow" IPA-transcription. Here, the author can be as detailed as s/he would like. What should not be presented here is the phonemic representation of an utterance, which is represented at Level +2.

This information will not be further processed but merely displayed. Nevertheless, it will be searchable.

Level +2 Phonological transcription

This level represents the first concrete analysis away from what is actually heard (or seen) towards a representation of the distinctive units of the utterance.

Here, the author should provide a purely phonemic, or "broad" IPA representation of the sentence, avoiding all audible/visual distinctions which do not convey meaning. The information is illustrated using IPA conventions.

Example:

Level +5	Mein Name ist Stefan.			
Level +2	m͡aɪn	ˈnaː.mə	ɪst	ˈʃtɛ.fan
Level -5	'My name is Stefan.'			

Level +1 Morphophonemic structure

This level represents a further abstraction away from the spoken word/visual sign towards the meaning it conveys. It is at this level that the author visibly separates the morphs that are represented in the Level +2. All information will be displayed here in a linear fashion. This is achieved in the following way:

Continuous Units

Where the morph boundaries are segmentally separable from the surrounding morphs, this will be indicated by the sign '-' at this level (as well as in the gloss (Level -1)), following standard practice.

Example: German:

Level +5	manche
Level +1	ˈman.ç-ə
Level -1	some-P
Level -5	'some'

Where this is not possible, the morph is not further divided at this level and, following standard practice, its morphemes will be separated by the symbol '.' in the gloss (Level -1). For example, the English morph which is written as 'am' and which denotes COPULA, 1ST PERSON, SINGULAR, NONPAST, is represented at this level merely as /æm/:

Example:

Level +5	am
Level +1	æm
Level -1	COP.NPT.1.S

Discontinuous Units

Discontinous affixes present something of a challenge for a system of representation such as that of the CRG project which relies completely on a linear representation of the information. However, by using a few simple conventions, this type of information can be encoded in such a way that the computer is capable of identifying the information as belonging to one single unit and also of recognizing what kind of an affix it is.

For example, circumfixes - and the information they convey - are consistently indicated by the sign < , > in Level +1 (the reconstructed morphemic structure) and Level -1 (the gloss) while transfixes are represented along these same principles by the sign | |. The affix marked in this fashion is then attached to the preceding and following morphs in the usual way by the symbol "-", if the morph is clearly isolatable or ".", if it fuses to some extent with a neighboring morph or morphs.

A placeholder, represented by the sign '_', will be inserted into the slot where an element of the transfix or circumfix is located in actual speech. The sign '_' is only intended as a visual aid for the user and will not be read by the computer, so that a direct comparison of the example with the lexicon is possible.

In this fashion, the computer is capable of processing the information contained within the two brackets or vertical lines as belonging to one single unit and can automatically recognize what type of unit it is.

This is illustrated in the following two examples. The first is an example of a circumfix from German and the second a transfix from Arabic.

The representation of circumfixes: German

Level +5	gemachte	Standardized orthography
Level +2	gə.ˈmax.tə	Phonological representation
Level +1	_max_-<gə, t>-ə	Morphemic reconstruction
Level -1	do-<PTCP>-P	Gloss
Level -5	ʼdone'	Colloquial translation

As Level -1' is automatically created after Level -1 has been entered, the computer can then indicate in Level -1' (the so-called "Templates", see below) what kind of morph is encoded in the form of a circumfix, transfix, etc. The German example above would then have the following form in Level -1':

Level -1' LEX-<WCD>-NUM (WCD - **W**ord-**C**lass **D**erivational (or inflectional) marking)

The representation of transfixes: Arabic

Level +2	ʼka.ta.ba	Phonological representation
Level +1	k_t_b_-|a, a, a|	Morphemic reconstruction
Level -1	write-|INF|	Gloss
Level -5	ʼto write'	Colloquial translation

The corresponding template to the Arabic example would then have the form given below:

Level -1' LEX-|WCD|

The same principles can also be used to indicate infixes. In the CRG, infixes will be consistently marked by the sign > < . Here an example from Chrau (Thomas, 1971:154), a language of Vietnam:

Level +5	vanŏh	Standardized orthography
Level +1	v_ŏh->an<	Morphemic reconstruction
Level -1	know->ATTR<	Gloss
Level -5	ʼwise'	Colloquial translation

The corresponding template for the Chrau example is:
Level -1'LEX->WCD<

Level –1 Gloss

This is the first level which is completely abstract. At this level, all information associated with the morph forms at Level +1 is given. The representation is as follows:

Where the sign '-' appears at level +1 it appears here as well. Following standard practice, if the content associated with a morphophonemic form consists of several components, their respective representation will be separated from its neighbors by the sign '.'.

Lexemes will be written in lower case symbols, proper names in scare quotes, while grammatical information is written in upper case symbols.

		Level +1	**Level -1**
Examples:	Arabic:	k_t_b-|i, a:|-i: 'my book'	write-|NML|-POSS.1.S
	German:	_max_-<gɅ, t>-Ʌ 'done'	do-<PTCP>-P
	English:	æm 'am'	COP.1.S.NPT

As the gloss, following the morphophonemic reconstruction of level +1, clearly shows whether the various morphemes correspond on a one-to-one basis to the individual morphs or not (through the use of either '-' or '.'), the computer can easily compute on the basis of this level to what extent grammatical marking in a particular language is fusional, agglutinating or isolating and also what kind of content tends to be expressed by what means of marking.
Also, on the basis of the information provided at this level, the computer will automatically generate the 'Template' Level –1' (see below).

Level -2: Complex Morphology

Level -2 is based on Level -1 (the Gloss) and is intended to allow the encoding of complex morphology into the more general categories. It is entered manually by the author. An example:

The German "Perfekt", actually a verbal category which is underspecified for the opposition between PAST and PERFECT, consists of three parts: an auxiliary, marked for tense, number and person, which preceeds the participle of the lexical verb. This participle consists of the stem of the lexical verb plus the participial marking, which for most verbs is the circumfix -<gə-, -t>. This is represented as follows:

Level +5	Es hat geklappt.				
Level +2	ɛs	hat		gə.ˈklapt	
Level +1	ɛs	ha	-t	_klap_	-<gə-, -t>
Level -1	PRN.3.S.M	AUX	-NPT.3.S	_work_	-<PTCP>
Level -2		PAST / PERFECT-work-NPT.3.S			
Level -5	'It worked.'				

Note that without Level -2, the user of this grammar would not be able to recognize that the combination "AUX-NPT.3.S LEXICAL.VERB-<PTCP>" is actually the representation of a single verbal category in German.
Another example is the English progressive, formed by the auxiliary plus the participial form of the lexical verb. Below is an abbreviated example of the English sentence 'I am reading a book.' Again, only through the use of Level -2 can the user of the grammar recognize the combination "AUX.NPT.1.S LEXICAL.VERB-PTCP" as the representation of the English verbal category "PROGRESSIVE".

Level +1	a͡ɪ	æm	ˈɹiː.d	-ɪŋ	ʌ	bʊk
Level -1	PRN.1.S	AUX.NPT.1.S	read:V	-PTCP	IDEF.ART	book:N
Level -2		PROGRESSIVE-read-NPT.1.S				

Level -3 Sentence Constituents

Information at this level gives the status of the sentence constituents, such as NP(SUBJECT), NP(OBJECT), V-FIN, and V-INF, etc., and is entered manually by the author.

It is designed to allow the computer to process information directly on the sentence constituents and provides the user with important information on the structure of the sentence at a glimpse. An example from German:

Level +5	Es hat geklappt.				
Level +2	ɛs	hat		gəˈklapt	
Level +1	ɛs	ha	-t	_klap_	-<gə-, -t>
Level -1	PRN.3.S.M	AUX-	NPT.3.S	_work_:V	-<PTCP>
Level -2		PAST / PERFECT-work-NPT.3.S			
Level -3	NP(SUBJ)	V-FIN		V-INF	
Level -5	'It worked.'				

Another example, from English:

Level +5	I am reading a book.					
Level +1	a͡ɪ	æm	ˈɹi.d	-ɪŋ	ʌ	bʊk
Level -1	PRN.1.S	AUX.NPT.1.S	read:V	-PTCP	IDEF.ART	book:N
Level -2		PROGRESSIVE-read-NPT.1.S				
Level -3	NP(SUBJ)	V-FIN	V-INF		NP(OBJ)	

Level –4 Literal translation & Level -5 Free English translation

Level -4 is intended as an aid for the user to interpret the gloss. It is entirely optional and will not be searchable. Here, the author is simply given the chance to include a rough translation of all or part of the gloss which s/he feels may otherwise not be recognizable in a free English translation. The free English translation, the obligatory **Level -5**, follows and will be searchable.

Example: Pali (Middle Indo-Aryan):

Level +5	attano	atthāya
Level -1	self.GEN.S.N	sake.DAT.S.M
Level -4	'for the sake of the self'	
Level -5	'for oneself'	

Level -1' 'Templates' of Level -1

The information provided in Level -1 (the gloss) is automatically processed and on this basis, Level -1' is generated. This level specifies which category the morpheme provided in the gloss belongs to.

Examples:
'ACC' and 'DAT' both belong to the category 'case', abbreviated here as "CAS".
Also, 'PT' (past) and 'NPT' (nonpast) both belong to the category 'tense', abbreviated here as "TNS".

Confer again the following Pali example:

Level +5	attano	atthāya
Level -1	self.GEN.S.N	sake.DAT.S.M
Level -4	'for the sake of the self'	
Level -5	'for oneself'	

On the basis of Level -1 (the gloss), the Template for this example would have the following form:

Level -1	self.GEN.S.N	sake.DAT.S.M
Level -1'	LEX.CAS.NUM.GEN	LEX.CAS.NUM.GEN

LEX - lexeme
CAS - case
NUM - number
GEN - gender

With this information, the computer is capable of preparing statistical data for a number of different types of information, such as e.g. which types of categories tend to be expressed by which means of marking and, for predominantly agglutinating or isolating languages, what is the order in which the categories are presented, etc.

Level -2' 'Templates' of Level -2

This automatically generated level abstracts away from the information given in Level -2 (Complex Morphology), similar to the process in Level -1'. That is, if say a verbal category, such as the English progressive, is expressed as a periphrastic construction, Level -2 is used to convey this information. Hence, in Level -2, we would have the information PROGRESSIVE-read-NPT.1.S for the periphrastic construction 'am reading' in the following example. Level -2' would then, on the basis of this information, 'translate' this information into the more abstract information ASPECT-LEX-TNS.PERS.NUM, to denote that PROGRESSIVE belongs to the abstract category of ASPECT/ACTIONALITY, etc.

Level +5	I am reading a book.					
Level +1	a͡ɪ	æm	ˈɹi.d	-ɪŋ	ʌ	bʊk
Level -1	PRN:1.S	AUX.NPT.1.S	read:V	-PTCP	IDEF.ART	book:N
Level -2		PROGRESSIVE-read-NPT.1.S				
Level -3	NP(SUBJ)	V-FIN	V-INF		NP(OBJ)	

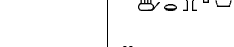
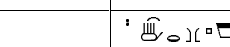
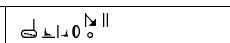
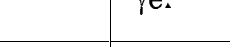
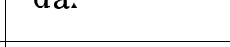
The two automatically generated levels would then have the following form, where Level -1' is based on Level -1 and Level -2' is based on Level -2:

Level -1'	PRN.PERS.NUM	AUX.TNS.PERS.NUM	LEX:V	-WCD	ART	LEX:N
Level -2'		ASP-LEX-TNS.PERS.NUM				

Using these same principles, the CRG format can eventually be programmed to create Templates for all other searchable levels where this is deemed desirable. For example, for Level +1 or +2, where the computer can eventually automatically create an inventory of phones or phonemes, etc.

Finally, to demonstrate how this approach works in a sign language, consider the following example from German Sign Language (DGS - "Deutsche Gebärdensprache"). For the purpose of demonstration, we have used the Hamburger Notational System (HamNoSys), slightly adapted to be able to include the suprasegmentals. We do not know whether HamNoSys or another system will eventually become the accepted standard system of transcription for sign languages in general, hence we will at present not comment further on this issue. Once an internationally accepted standard has been reached, this can easily be incorporated into the system.

DGS – German Sign Language

Level + 6 Standard Orthography			
Level + 5 Standard Transliteration	Frage ? REGEN DA		
Level + 4 Suprasegmentals			
Level + 3 Phonetics	 "ɣ eː d aː"		
Level + 2 Phonology	 "ɣeː"	 "daː"	
Level + 1 Morphemic Structure	 "ɣeː"	 "daː"	
Level – 1 One-to-One Gloss	rain:V	there:ADV	Q
Level – 2 Complex Morphology			
Level – 3 Sentence Constituents	V-FIN	ADV	
Level – 4 Literal Translation	'Rains there?'		
Level – 5 Free Translation	'Is it raining outside?'		

The automatically generated levels:

Level – 1'	LEX:V	LEX:ADV	MOD
Level – 2'			

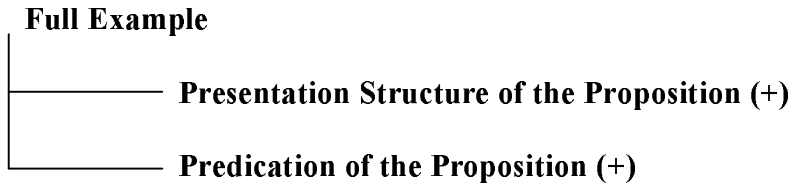
Note that, if one is able to read HamNoSys, the present example provides a much more detailed account of this question than does the standard transliteration (Level +5), although the standard transliteration is admittedly easier to read.

Due to the nature of sign languages in general, we have had to adapt our strategy here somewhat, although only minimally. If there is more than one simultaneous component, these are simply presented one over the other, which the computer will however read as two different levels, both of which correspond to one level in spoken languages.
As an example take the "mouthings" which accompany the hand movement but must be considered a part of the signs. These "mouthings" represent the use of the mouth to produce movements which have the same forms as does the production of sounds (hence the scare quotes, as the sounds themselves are not meant). As this information occurs simultaneously with the movement of the hands, this information is simply placed below that of the hand movement. The same principle can of course also be used for the suprasegmentals. For example, a polar question expecting a positive answer (such as *Isn't it raining?*) would involve both the raising of the eyebrows and simultaneously shaking the head.

This information would then simply be presented as two levels within the suprasegmental level (+4).

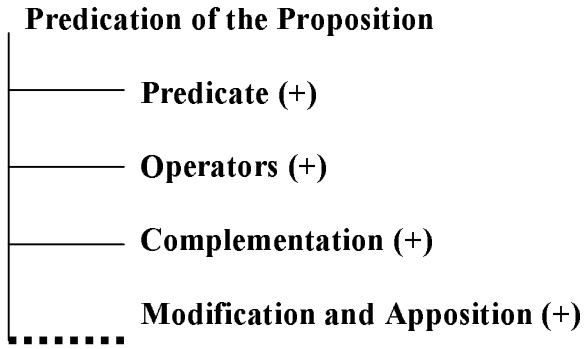
5.2. Grammatical Description

Once the author has determined that the example is not solely intended for the lexicon or a paradigm (i.e., it has a propositional content and is a "Full Example")[2] and has entered the data according to the criteria given in the section 4.2.2 above, s/he is first presented with the following:



"Presentation Structure", which deals with information packaging or pragmatic aspects of the utterance, will be dealt with separately in section 5.2.6. In the sections leading up to this, we will be concerned with what is contained under "Predication of the Proposition".

If we click on "Predication of the proposition", we find the following:



The "Operators" include tense, aspect/actionality and propositional negation, all of which will be discussed in detail in section 5.2.4.

The following three sections deal with the "Predicate" (5.2.1.), "Complementation" (5.2.2.) and "Modification and Apposition" (5.2.3.). "Modification and Apposition" here refers to sentence-level modification. However, the principles involved are the same as those for modification and apposition in general and as such they will be handled in a uniform manner.

[2] The Lexicon and Paradigms will be dealt with in section 5.

5.2.1. Predication

To define what is meant in this grammar under the term "predication", or more precisely, under the term "predicate", which is the topic of the present chapter, we must first turn to the topic of presentational structure, which will be dealt with in more detail in section 5.2.6. below.

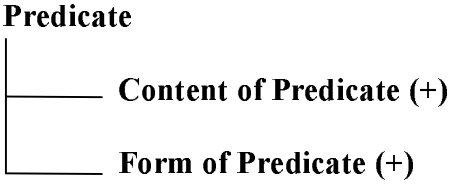
The predicate of a proposition is that part of the sentence which comprises the semantic head of the focus in a "topic-comment" or "predicate-focus" construction and all grammatical marking which this requires (cf. for example the discussion in Lambrecht, 1994). That is, a sentence in which an established topic is commented upon and whose pragmatic structure may be termed "unmarked". In marked pragmatic contexts (i.e., argument focus, sentence focus) it is that unit which most closely corresponds to this morphological unit in the unmarked construction.

It thus corresponds to the "predicate" in both constructions referred to by Bloomfield (1984:173) as "narrative predication" and "equative predication", among the many other types. In languages in which a verbal class is found, it will be this class, in languages which make use of a copula in "equative" predication, it will be the "predicate noun/adjective, etc." which appears in conjunction with this copula, as well as the copula itself.

It will in addition be that unit of the sentence which takes or can take complements and, in the case that there is at least one complement in the sentence, the predicate will be that unit whose valency determines the number of complements in the sentence as well as (at least in many languages) their respective forms.

What the "predicate" is NOT is a "verb phrase" or more generally, a verb and its object or objects, nor does it include adverbial or sentential modification.

When discussing the predicate of the proposition, we need to discuss both its CONTENT and its FORM, a dichotomy which will re-occur in virtually all of the following areas of the CRG format to be discussed below. Thus, the author will be presented with the following two REQUIRED nodes:



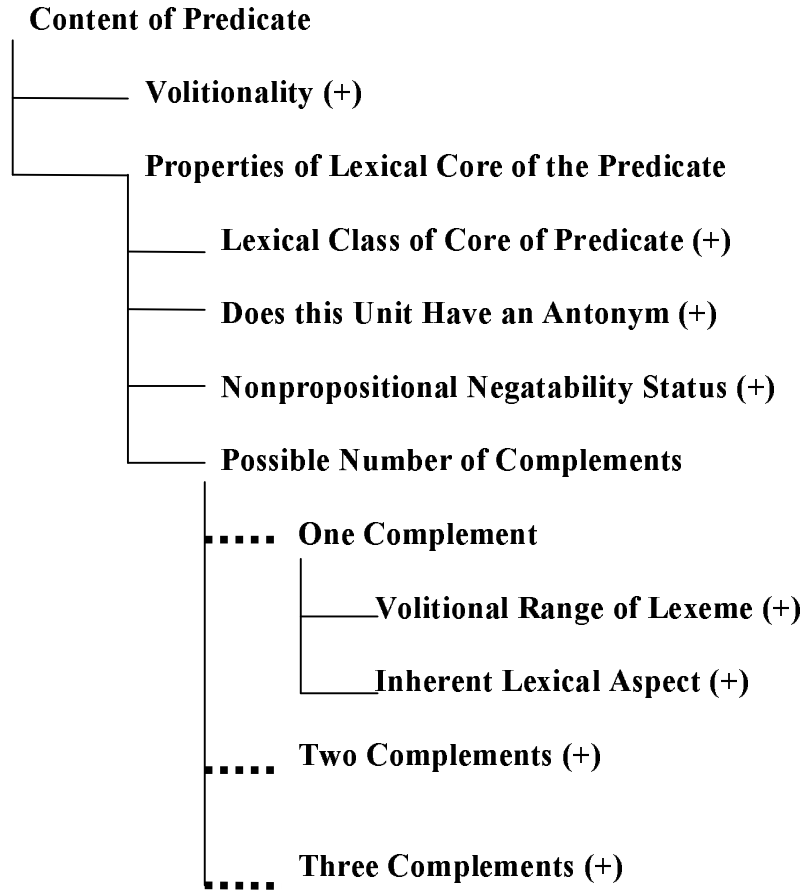
Let us begin our discussion with the CONTENT of the predicate of the proposition. This includes much information about the use of this lexeme in the actual utterance, e.g., the VOLITIONALITY of the predicate in its current use. Aspectual and actional information on the actual use of this predicate is discussed under "Operators", section 5.2.4.

We would also like to know - for the lexicon - the LEXICAL CLASS of the predicate (or rather, the lexical class of its SEMANTIC CORE), whether this predicate (core) has a (near or exact) antonym and whether it can also be, or perhaps already is, negated through lexical negation (cf. English *un-*, *dis-*, etc.).

We would also like to know for the lexicon information such as the POTENTIAL VOLITIONAL RANGE of this lexeme, the number of arguments it can take, and the inherent aspectual status (i.e., actionality) of this lexeme. As this will often depend on the number of arguments the predicate takes in an actual utterance (compare *He ran for/*in one hour* (atelic) with *He ran home *for/in one hour* (telic)), this aspectual information is attached as a required node to the respective alternatives for the number of possible complements.

Finally, the predicate may also be modified (i.e., not propositional modification but adverbial modification). For a full treatment of modification and apposition in general, see section 5.2.3.

The following diagram shows the ordering of this information:



Lexical Class of the Core of the Predicate

The determination of which lexical classes should be included here was a long and difficult process, as there is an extensive amount of literature on this topic. It was finally decided to use the criteria in Fellbaum (1993) as the basis for this classificatory system, albeit with a number of changes.

These changes primarily affect Fellbaum's classes "Verbs of Change" and "Stative Verbs", as these two criteria belong, properly speaking, to an aspectual and actional description of the lexeme and not its semantic class. "Stative" has been replaced by a number of different classes, such as "Abstracta", "Comparison, Equative or Superlative", "Identification", "Location" and "Time". The class "Verbs of Change" can be captured according to the

remaining classes listed in Fellbaum (1993) as well as by the other additional classes, to be described below.

First, to Fellbaum's class of "Creation Verbs" we have added the class of "Destruction". Also, as a number of "Verbs of Possession" listed by Fellbaum involve the transfer of an entity from one individual to another (e.g. *grant*), we have chosen to include the class "Giving or Transfer" (which includes predicates of RECEPTION as well) and to restrict "Possessional" to those predicates which denote only POSSESSION as we define this (*have, is mine,* etc.).

The class of "Movement" predicates has been greatly expanded, to include the criteria given in Talmy (1985), such as path, manner, etc.

Finally, if the author believes that s/he does not find the appropriate classification offered in this list, an additional "Other" node is offered. Also, as all nodes listed here are optional or inclusive, any combination of these may be chosen as is necessary, for example, the combination "Movement" and "Weather" for *rain,* etc. Here the complete list:

Lexical Class of Core of Predicate

- **Abstracta**
- **Body Functions and Care**
- **Cognition**
- **Communication**
- **Comparison, Equative or Superlative (+)**
- **Competition**
- **Consumption**
- **Creation**
- **Destruction**
- **Emotion or Psyche**
- **Giving or Transfer**
- **Identification**
- **Location**
- **Movement (+) (see below)**
- **Perception**
- **Physical Contact**
- **Possessional**
- **Social Interaction**
- **Time**
- **Weather**
- **Other Class - Specify**

"Movement" is rather detailed in the CRG, taking the following criteria into consideration:

- PATH OF MOVEMENT: thus *rain*, whether used transitively or intransitively, refers to downward motion, while *rise* of course refers to upward motion. The lexeme may either be specified for path of movement or not. Hence, the first node here is "required" while

the two following nodes (LEXEME IS NOT SPECIFIED FOR MANNER and LEXEME IS SPECIFIED FOR MANNER) are "exclusive" nodes, i.e., one of the two must be chosen.

If the lexeme is specified for path, the following criteria are offered:

- GENERAL DIRECTIONAL MOVEMENT, divided into,
- VERTICAL MOTION (with the nodes UPWARDS, DOWNWARDS, UP AND DOWN and OTHER) and
- HORIZONTAL MOTION (with the alternatives SIDEWAYS, FRONT TO BACK, BACK TO FRONT, SIDE TO SIDE OR BACK AND FORTH, OTHER), TWISTING MOTION, CIRCULAR MOTION, OTHER.

This is followed by the alternatives

- FROM SOURCE e.g. *leave*, with the following daughter nodes: SOURCE IS DEICTIC CENTER, SOURCE IS NOT DEICTIC CENTER, UNDERSPECIFIED
- TO GOAL, e.g. *go to*, with the following daughter nodes: GOAL IS DEICTIC CENTER, GOAL IS NOT DEICTIC CENTER, UNDERSPECIFIED
- MOVEMENT INTO
- MOVEMENT OUT OF
- MOTION PAST ANOTHER OBJECT
- MOTION WITHIN ANOTHER OBJECT
- MOTION OUTSIDE ANOTHER OBJECT
- MOTION THROUGH ANOTHER OBJECT
- MOTION ALONGSIDE ANOTHER OBJECT
- MOTION OF OBJECTS TOWARDS EACH OTHER
- MOTION OF OBJECTS AWAY FROM EACH OTHER
- RETURN MOTION
- OTHER

With the exception of the daughter nodes of "From Source" and "To Goal", all other nodes are optional or "inclusive" and can be combined with other nodes as required, e.g. "Downwards" and "From Source", etc.

The next criterion is the required node

- MANNER OF MOVEMENT. Again, the lexeme may either be specified for this criterion or not. As one of these two alternatives must be chosen, they are presented as "exclusive" alternative nodes.

If the lexeme is specified for manner of movement, the following optional alternatives are presented, of which any combination may be chosen:

- SPECIFIED FOR SPEED (specified further by the exclusive alternatives VERY SLOW, GENERAL SLOW, MEDIUM SPEED, GENERAL FAST, VERY FAST)

- SPECIFIED FOR MEANS OF MOTION. This is divided into two major, mutually exclusive groups:

- SELF PROPELLED MOTION. This is divided into the following optional alternatives, any combination of which may be chosen: ON LAND (with the daughter nodes ON FOOT (*run, walk*) USE OF HANDS AND KNEES OR FEET (*crawl*), FLAT ON GROUND (*slither*) and OTHER. After ON LAND we have the alternatives IN WATER (with the daughter nodes FLAOTING, SWIMMING, DIVING, UNDERWATER MOVEMENT and OTHER), and finally IN AIR (*fly, soar*).

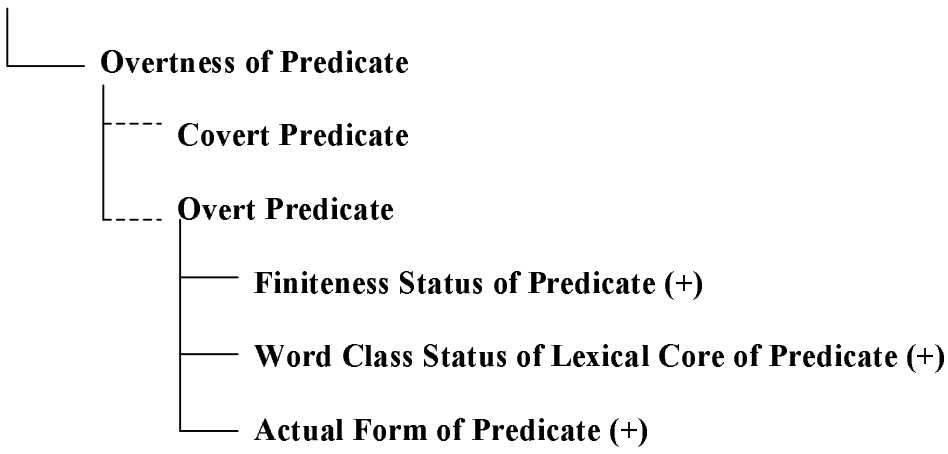
- ASSISTED MOTION. Here, there are two mutually exclusive alternatives: GENERAL ASSISTED MOTION (*ride*) and MEANS OF NON GENERAL ASSISTED MOTION, with the optional alternatives CAR, CART OR WAGON, AIRPLANE, ANIMAL (with additional possibilities), BOAT, CANOE OR SHIP (with all possible combinations included as daughter nodes), BICYCLE and OTHER.

- SPECIFIED FOR WHAT MOVES. The optional alternatives here are ANIMAL and HUMAN (for example, the difference between *gallop* and *run*), NON VEHICULAR INANIMATE OBJECTS (including METEOROLGICAL PHENOMENA such as RAIN, SNOW, etc.) and VEHICLES OF MOVEMENT, with the same alternatives as given above under MEANS OF NON GENERAL ASSISTED MOTION.

This list of criteria should be sufficient to describe virtually all movement verbs, and the reader may check for him-/herself that it will accurately describe at least the most notable differences between the following verbs of motion: *fly, leave, go to, get out of, run past, walk through, penetrate, gallop, crawl, slither, tread water, float, swim, dive, take off, land, enter,* and many others.

Form of Predicate

The uppermost daughter nodes of this node are as follows:

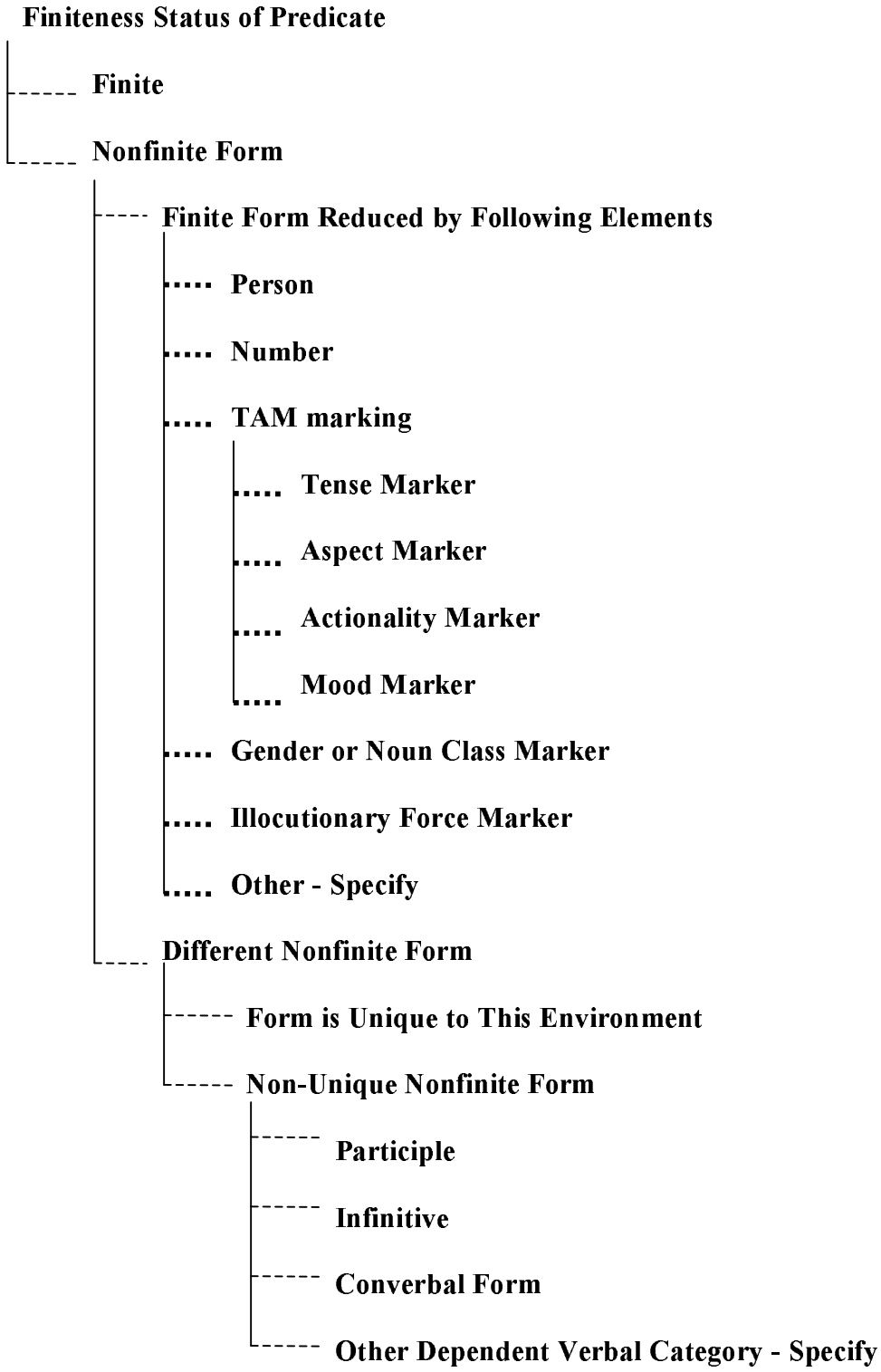


The first node is required since not every full example will have an overt predicative expression, such as the answer *Me!* to the question *Who wants to go?* or interjections, etc.

If a predicative form IS used, we would like to know a number of things, including its FINITENESS STATUS, the WORD CLASS of its lexical core, as well as a detailed acount of its ACTUAL FORM (i.e., is there overt marking of pertinent properties, is it a simple verb, or perhaps a light verb and non-referential noun, is a copula used, is it a periphrastic form, etc.).

Finiteness Status

Under the term "finite" here we mean a predicate which carries ALL marking required to be able to appear as the main predicate of a proposition. ALL OTHER FORMS are considered here "Nonfinite Forms". The "nonfinite" node is then further divided, as "non-finite" in many languages is a matter of degree, with some "nonfinite" forms having a large amount of the marking required to function as a "finite" form while lacking one or two categories (cf. for example the data in Ebert, 1993a). This is described as follows:



The question as to the word class of the lexical core of the predicate is easily dealt with as follows. It is the same basic means of inquiring as to the word class used throughout the grammar as a whole.

Word Class Status of Lexical Core of Predicate

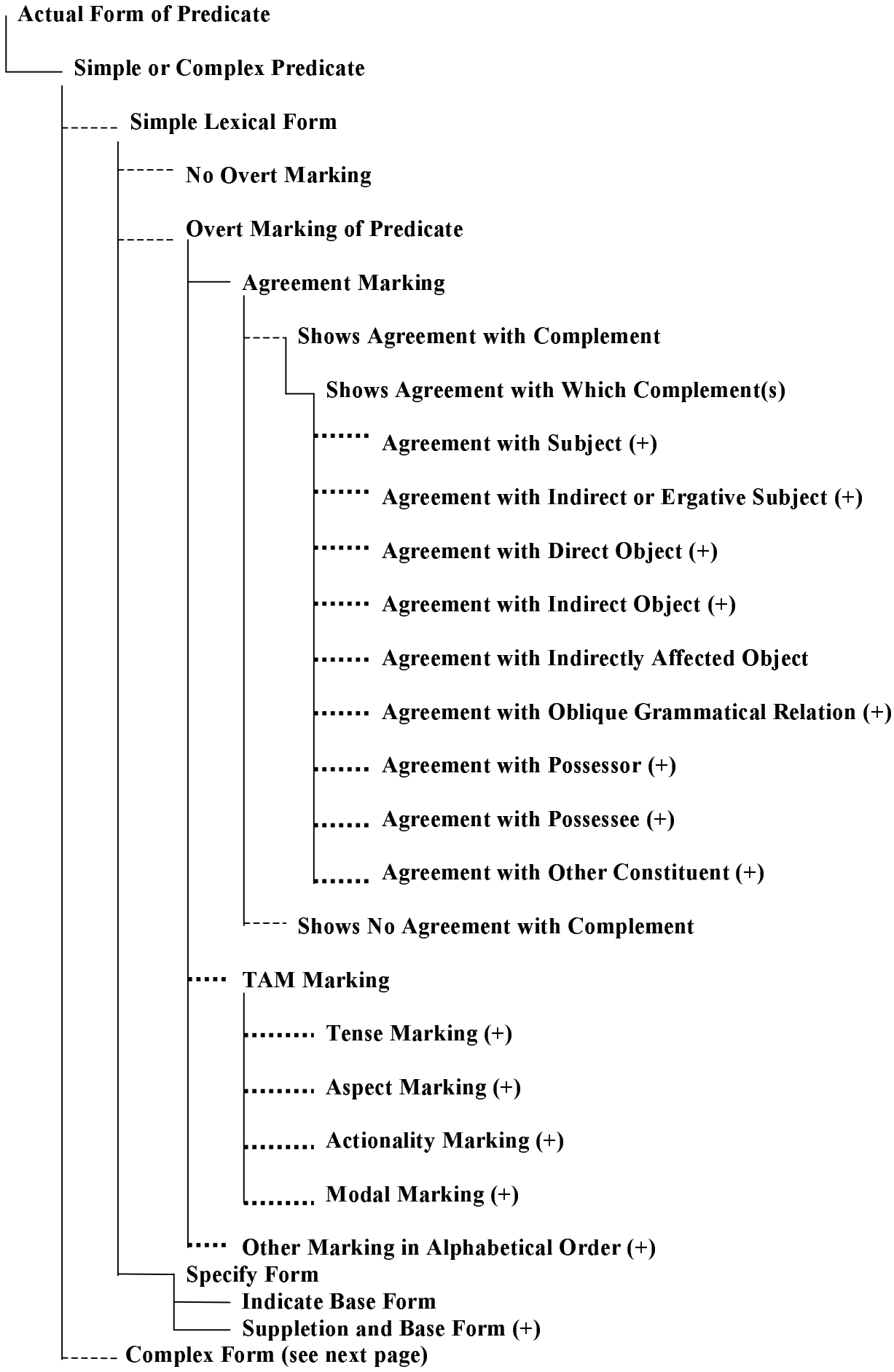
- **Verb**
  - **Pro-Verb**
  - **Lexical Verb**
  - **Numeral Verb**
  - **Negating Verb**
- **Nominal Phrase (+)**
  (for details, see section 5.2.2. on complementation, where the NP is discussed in detail)
- **Adjective**
  - **Proadjective**
  - **Lexical Adjective**
  - **Numeral Adjective**
- **Adverb**
  - **Proadverb**
  - **Lexical Adverb**
  - **Numeral Adverb**
- **Adposition**
  - **Preposition**
  - **Postposition**
  - **Can Function Both as Preposition and Postposition**
    - **Use in Current Example**
      - **Preposition**
      - **Postposition**
- **Numeral**
- **Particle**
- **Indeterminate or Underspecified**
- **Other (+)**

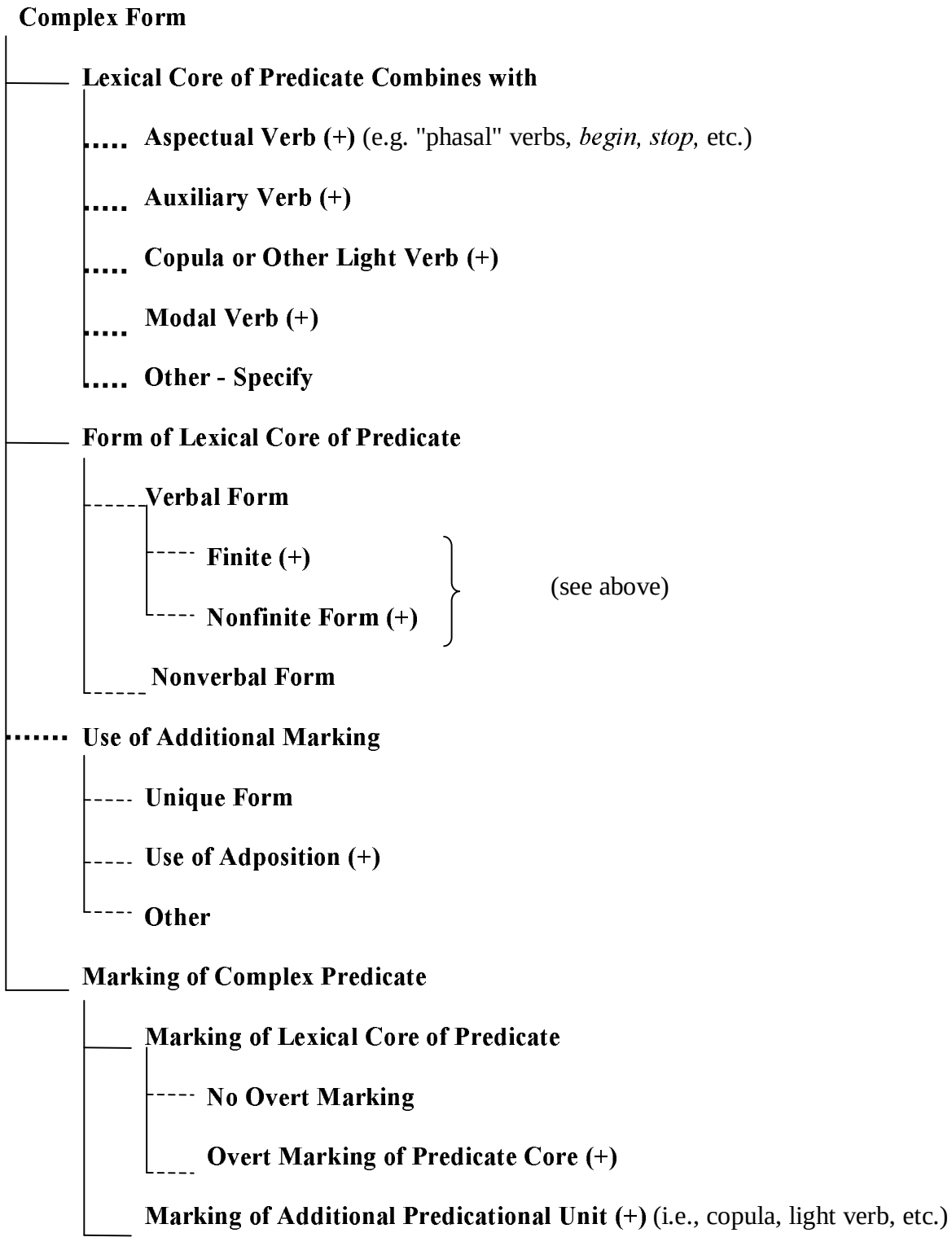
In order to be able to discuss the ACTUAL FORM OF THE PREDICATE more easily, we first divide the forms into SIMPLE and COMPLEX FORMS. This distinction divides the predicates primarily into periphrastic and non-periphrastic forms, so that we can enquire as to the marking of the different components.

A "simple" predicate consists of a single word - a verb, a "predicate noun" in a language which does not make use of a copula in such predicates, etc. All other forms, i.e, all periphrastic forms, light verbs with a non-referential noun or an adjective, etc., will be considered "Complex Forms".

Once we have determined whether the predicate consists of a single word or not, we can then proceed to inquire as to the marking of each component of the predicate - NP, lexical verb, lexical verb plus some other element (e.g. *eat up*), etc.

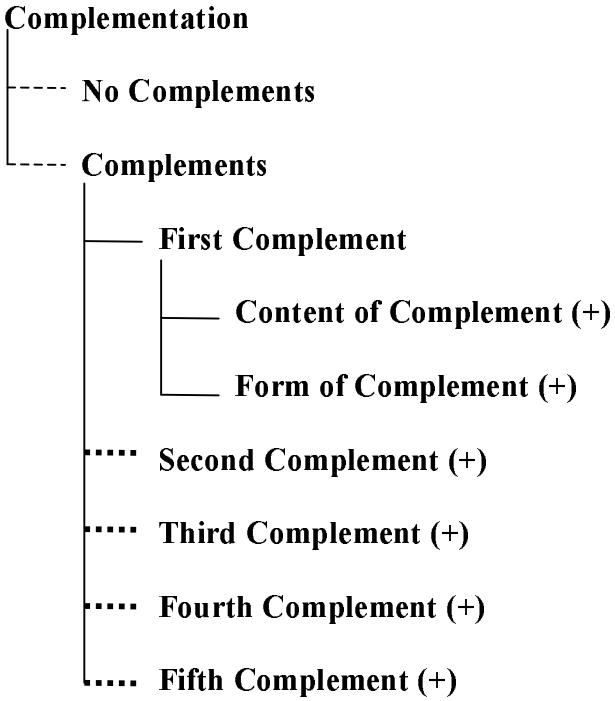
Below is a schematic representation of this structure:


Actual Form of Predicate
Simple or Complex Predicate
Simple Lexical Form
No Overt Marking
Overt Marking of Predicate
Agreement Marking
Shows Agreement with Complement
Shows Agreement with Which Complement(s)
Agreement with Subject (+)
Agreement with Indirect or Ergative Subject (+)
Agreement with Direct Object (+)
Agreement with Indirect Object (+)
Agreement with Indirectly Affected Object
Agreement with Oblique Grammatical Relation (+)
Agreement with Possessor (+)
Agreement with Possessee (+)
Agreement with Other Constituent (+)
Shows No Agreement with Complement
TAM Marking
Tense Marking (+)
Aspect Marking (+)
Actionality Marking (+)
Modal Marking (+)
Other Marking in Alphabetical Order (+)
Specify Form
Indicate Base Form
Suppletion and Base Form (+)
Complex Form (see next page)


Complex Form
Lexical Core of Predicate Combines with
Aspectual Verb (+) (e.g. "phasal" verbs, *begin, stop,* etc.)
Auxiliary Verb (+)
Copula or Other Light Verb (+)
Modal Verb (+)
Other - Specify
Form of Lexical Core of Predicate
Verbal Form
Finite (+)
Nonfinite Form (+)
(see above)
Nonverbal Form
Use of Additional Marking
Unique Form
Use of Adposition (+)
Other
Marking of Complex Predicate
Marking of Lexical Core of Predicate
No Overt Marking
Overt Marking of Predicate Core (+)
Marking of Additional Predicational Unit (+) (i.e., copula, light verb, etc.)

5.2.2. Complementation

Complementation refers to those sentential constituents whose presence is required by the valency of the predicate. As with all areas of the grammar, complementation consists of two major, logically independent, areas - CONTENT and FORM. Let us begin with a diagram of the uppermost nodes found directly under the node "Complementation":



Of course, the predicate may not require any complements at all. Thus, there are either NO COMPLEMENTS or there are COMPLEMENTS. The first of these two alternatives refers only to predicates with zero valency. Thus, if there are complements, even if these are not overtly realized, then the second alternative must be chosen. A typical example of a proposition with no complements is found in the Italian sentence *Piove* 'It is raining' which, at least in one analysis, has no complements. If required by the author's analysis, however, this sentence may of course also be interpreted as having a zero or non-overt expletive subject, in which case "Complements" should of course be chosen.

If there are complements, then there will of course be at least one complement. Hence, the first complement is a required node, while the remaining nodes are all optional. For every complement we will then need to describe both the CONTENT of this complement and the FORM it takes in a linguistic utterance. Note that the two are independent of each other here. That is, we do not anticipate that a certain semantic description, for example, "solid physical entity", will necessarily be represented by a "noun", nor that a propositional complement will necessarily be represented by a clause. Although certain tendencies are found here, we cannot preclude the possibility that a "solid physical entity" will be expressed by an entire clause.

Let us now take a closer look at the "Content" node. The uppermost nodes here are presented in the following diagram:

Content of Complement

- **Expletive Complement**
- **Non Expletive Complement**
  - **Standard or Comparee**
  - **Possessor or Possessee**

That is, semantically speaking, a complement is either an EXPLETIVE or NON EXPLETIVE complement, while it may also either be a STANDARD OR COMPAREE (in equative, comparative and superlative constructions) or it may be a POSSESSOR or POSSESSEE - or perhaps a combination of one of these last two pairs. Thus, while the complement MUST be either expletive or non-expletive, it is not necessarily a standard/comparee or possessor/possessee, although it may be one of these. The nodes "Standard or Comparee" and "Possessor or Possessee" will be dealt with further below, after dealing with the expletive and nonexpletive complements in detail.

Let us assume for the moment that we have an expletive complement, such as *it* in the proposition *It's raining*. The semantic side of such a complement is easily dealt with, since the complement has no referential properties. We would like to know, however, what grammatical role it has. The following alternatives are presented, which are the same as those presented everywhere else in the grammar with respect to grammatical roles:

- GRAMMATICAL SUBJECT. The determination of the "subject" is of course a language-specific question. This will be that argument which - if any - controls verb agreement, occupies the "subject" position of the clause, appears in the nominative, etc. It may also be that argument which is what Foley & Van Valin (1985) refer to as the "privileged NP", as it controls a number of syntactic and/or pragmatic pivots. We do not assume that every language will have the category "subject", nor do we assume that every language which possesses this category will require every sentence to have a "subject".

- INDIRECT OR ERGATIVE SUBJECT. This is that NP which appears in the ergative and which has a grammatical role comparable to that of "subject" and "object".

- DIRECT OBJECT

- INDIRECT OBJECT

- INDIRECTLY AFFECTED OBJECT. The term is borrowed from Neukom's (in press) grammar of Santali (Munda, Austro-Asiatic) and has been slightly adapted here. This refers to the NP which is indirectly affected by an action/situation without being either the direct or indirect object, but which **must be considered a grammatical role for the language concerned**. This indirectly affected object may be POSITIVELY AFFECTED, NEGATIVELY AFFECTED or simply AFFECTED IN GENERAL. Perhaps the closest English equivalent to this would be something like *on me* in the sentence *Don't you walk out on me!* It also resembles the dative in German utterances of the type *Er hat es mir weggenommen* 'He took it away from me'.

- OBLIQUE GRAMMATICAL ROLE

- NO GRAMMATICAL ROLE

- OTHER

Let us now turn our attention to the "Content" side of the non-expletive complements. In addition to the GRAMMATICAL ROLES just mentioned, this complement will also have

- ANIMACY PROPERTIES, such as "animate" (human, non-human), "inanimate" (physical entities (and their most salient properties) and non-physical entities (with the same lexical classes offered here as with predicates, see section 5.2.1. above for more details)).

- They will also have a PRAGMATIC STATUS. The alternatives here have been taken almost without modification from Gundel et al. (1993). For more details, cf. section 5.2.6.

- We may further enquire as to the DECITIC ROLE of the complement. Is it a SPEECH-ACT PARTICIPANT, and if yes, which? Or is it NOT A SPEECH-ACT PARTICIPANT, and if so, is the entity being referred to present at the speech-act situation or not?

- Further, we must clarify the QUANTIFICATIONAL STATUS of the entity referred to. Is it a COUNT ENTITY, whose quantity is exactly specifiable or not? Or is it a MASS ENTITY, and if so, can this be measured in countable entities (*cup, kilogram, cubic meter*) or not? In cases where the exact amount of a mass entity or the number of count entities cannot be specified, we would still like an approximation of whether this is "much", "little", "few", "many", etc. or perhaps "each or every individual" or "each or every set of individuals", in which case we also need a further specification of the set itself.

- We would like some information - for the lexicon - on the ANTONYM STATUS of this linguistic expression. Does it have an exact antonym? Is this an exact or near antonym? Is one element of this pair the UNMARKED member (cf. the unmarked question *How old are you?* with the marked alternative *How young are you?*)?

- Can this lexeme be NEGATED by means of lexical negation? Is the form perhaps already negated? And what type of negation does this involve, CONTRARY or CONTRADICTORY?[3]

Standard or Comparee

Of course, the complement may also be involved in an equative, comparative or superlative construction, of which it may be either the STANDARD or COMPAREE. Cf. the following examples:

He's as tall as all of us.
He's taller than all of us.
He's the tallest of all of us.

In each case, *he* is the COMPAREE while *all of us* is the STANDARD.

[3] The manner of dealing with negation will not be discussed here in any detail but will be treated in section 4.3.4.4.

The next step is of course to enquire as to the means of marking this relation, the details of which are discussed at length in section 5.2.5.

Possession

The complement may of course also be either a POSSESSOR or a POSSESSEE, and if it is one of these, it may simultaneously be the possessor of one relation and the possessee of another, as for example *secretary* in the following sentence:

Madeleine's secretary's car

Here, *secretary* is simultaneously the POSSESSEE of *Madeleine* but the POSSESSOR of *car*, information which may be important in the marking of this possessive relation.

Once we know the status of a complement in a possessive relation, we can of course also enquire as to the means of marking this status. For a detailed discussion of the MEANS OF MARKING, see section 5.2.5.

There is a good deal of further information on this possessive relation which must also be given.[4] Following the criteria in Heine (1997), we have chosen the following:

Intrinsic Relations (Required Node). This grouping consists of the following criteria:

- TEMPORAL RELATIONS (with the exclusive alternatives PERMANENT, TEMPORARY, UNDERSPECIFIED)
- PART/WHOLE RELATION. Is one of the two entities a part of the other? If so, which?
- BODY PART RELATION (listed separately as languages will often treat body parts differently from part/whole relations in general)
- KINSHIP RELATION with the required node concerning the exact relationship between the two persons (or perhaps animals) involved
- ALIENABILITY STATUS. While all of the relations given above relate to (in)alienability from a purely semantic point of view, languages differ considerably as to whether this realtion is actually treated from a morphosyntactic point of view as alienable or inalienable. Hence, this question refers to whether the relation in the current example is treated from a purely linguistic point of view as alienable, inalienable or underspecified for (in)alienability.

Extrinsic Relations (Required Node)

- CONTROL RELATION. Does one entity in the possessive relation freely dispose of the other? If so, which?
- LOCAL RELATION. Are the two entities in close proximity to one another or not, or perhaps underspecified in this respect?

[4] Pragmatic information, which of course plays a large role in possession, has already been entered above under "Content of Complement".

- OWNERSHIP RELATION. Can it be claimed of one of the two entities that it is the actual owner of the other? Although it is generally the possessor which is the owner, this is not necessarily the case, as can be seen in the NP *the dog's owner*.

Additional Language-Specific Relations

Here, there is room for criteria which may play a role in possession in some languages but which are very restricted in their occurrence. Examples include whether the relation is between an individual or a group and another entity, or whether the possessee is considered good, bad, large, small, etc. Of course, an optional "Other" node will be available here.

Finally, we have the possibility to fully describe any possible MODIFICATION of this complement. As modification is dealt with in detail in secion 5.2.3. below, we will not pursue this topic further here.

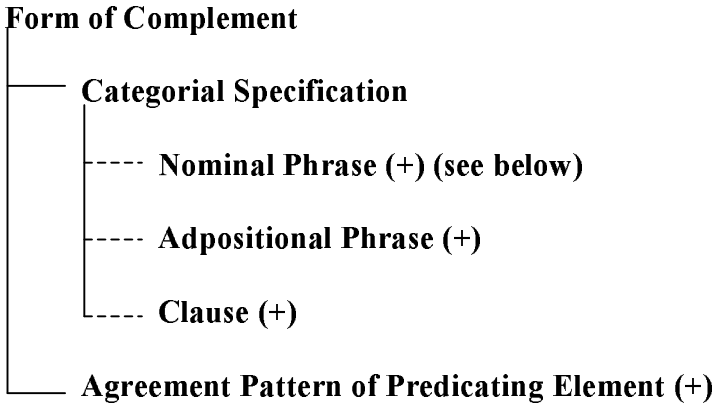
Let us now turn to the question of the FORM of the complement. In addition, we would like other formal information on this expression, for example, whether the predicate - if it shows agreement with one or more complements - agrees with this complement.

As far as form is concerned, we assume that every complement is expressed. It may or may not be OVERTLY expressed, but if it is a COMPLEMENT as we have defined this term, i.e., if its presence is required by the valency of the predicate, then it must be present in the sentence, whether overtly or not. If it is not overtly specified, then the CATEGORIAL SPECIFICATION, to which we now turn, is that of this complement if it were overtly expressed.

We group complements according to their form as follows:

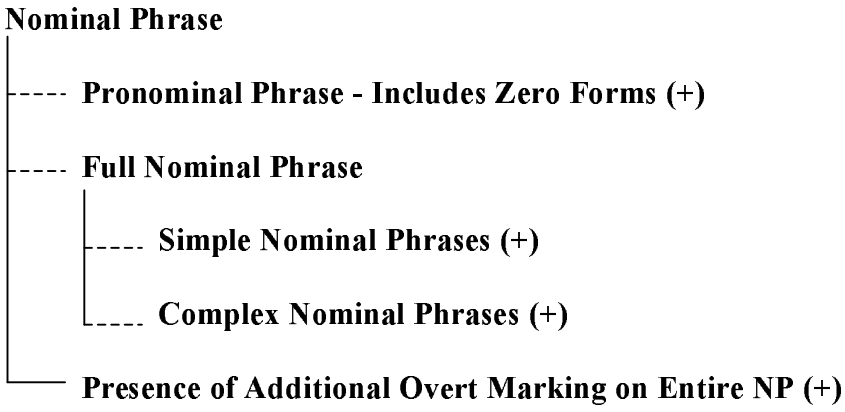
- NOMINAL PHRASE
- ADPOSITIONAL PHRASE
- CLAUSE[5]

Thus, the uppermost nodes here have the following structure:



Here, a diagram of the uppermost nodes under "Nominal Phrase":

[5] If the semantic head of the complement is a VERB, this information is to be given either under CLAUSE or NOMINAL PHRASE, whichever is most appropriate.



NOMINAL PHRASES

NOMINAL PHRASES are grouped into PRONOMINAL PHRASES, which include zero realization, and FULL NOMINAL PHRASES.

The description of **PRONOMINAL PHRASES** is as follows:

- First, we need to know whether this is OVERTLY or NOT OVERTLY expressed. If it is overtly expressed, it requires the following description
- REALIZATION OF PROFORM, i.e., CLITIC PROFORM, STRONG or WEAK. If none of these criteria is applicable, there is also the alternative INAPPLICABLE.
- KIND OF PROFORM. This is includes the following alphabetically ordered list, which is intended to be exhaustive and which may, of course, be updated at any time as necessary: ANTILOGOPHORIC, CONTRASTIVE, DEMONSTRATIVE, FOCUSSED BUT NOT CONTRASTIVE, GENERIC PRONOUN, IMPERSONAL PRONOUN, INTERROGATIVE PROFORM, LOGOPHORIC, OBVIATIVE, PERSONAL PRONOUN, POSSESSIVE PROFORM, PROXIMATIVE OR FOURTH PERSON, RECIPROCAL, REFLEXIVE, RELATIVE PROFORMS, OTHER
- RELATION TO ANTECEDENT: ANAPHORIC, CATAPHORIC, EXOPHORIC, NOT APPLICABLE
- The most commonly marked categories, including GENDER or NOUN CLASS, PERSON, NUMBER, ANIMACY
- CATEGORY OF ANTECEDENT: (noun, noun phrase, clause, etc.)
- the OBLIGATORINESS of its use
- A full description of its actual FORM, including information on SUPPLETIVE BASES
- Finally, there will be a node where ALL ADDITIONAL CATEGORIES which are marked on the pronominal complement can be described.

We now turn to the most complex node here, the description of **FULL NOMINAL PHRASES**.

We first distinguish here between two types: SIMPLE NOMINAL PHRASES and COMPLEX NOMINAL PHRASES. By SIMPLE NOMINAL PHRASES we do not simply mean "bare nouns" but rather a nominal phrase which may or may not have an overtly expressed SEMANTIC HEAD,

DETERMINER, QUANTIFIER and/or MODIFIER. What is important here is that each of these levels may at most be occupied by ONE ELEMENT. If there are two or more elements at any level, no matter how these are joined, we have a COMPLEX NOMINAL PHRASE. Some examples may serve to illustrate this point:

SIMPLE NOMINAL PHRASES

this book
the big one
the small table
two books
my watch
your three books

COMPLEX NOMINAL PHRASES

blue, red or green books	-	Complex modifier
this or that table	-	Complex determiner
two or three books	-	Complex quantifier
your or his watch	-	Complex possessive determiner
my book and your watch	-	NP-level Complexity
me and her	-	NP-level Complexity

SIMPLE NOMINAL PHRASES

First, we ask for a schematic representation of the form here, so that consultants can see at a glance how nominal phrases are structured in the respective language. This information will also automatically be fed into the **Inventory of Simple Nominal Phrases** (see section 6.). Two examples:

the two red books	⇒	DET QUANT ADJ N
my table	⇒	POSS.DET N

Next, we enquire about each of the individual components of the nominal phrase, beginning with the SEMANTIC HEAD, generally a noun, although we of course also enquire as to the word-class status of this element.

This semantic head may be overtly expressed or not, and if it is overtly expressed, it may be a proform, lexical noun, etc. In addition, it can also be classified with respect to its most salient semantic properties (ANIMACY, etc.). Finally, we would like to know of any OVERT MARKING which the semantic head carries. This does not refer to the overt marking of the NOMINAL PHRASE AS A WHOLE, but rather just to that marking which the semantic head may carry independently of the marking of the entire nominal phrase. If there is overt marking on the entire phrase, this is encoded elsewhere (see "Presence of Additional Overt Marking on Entire NP" in the diagram above).

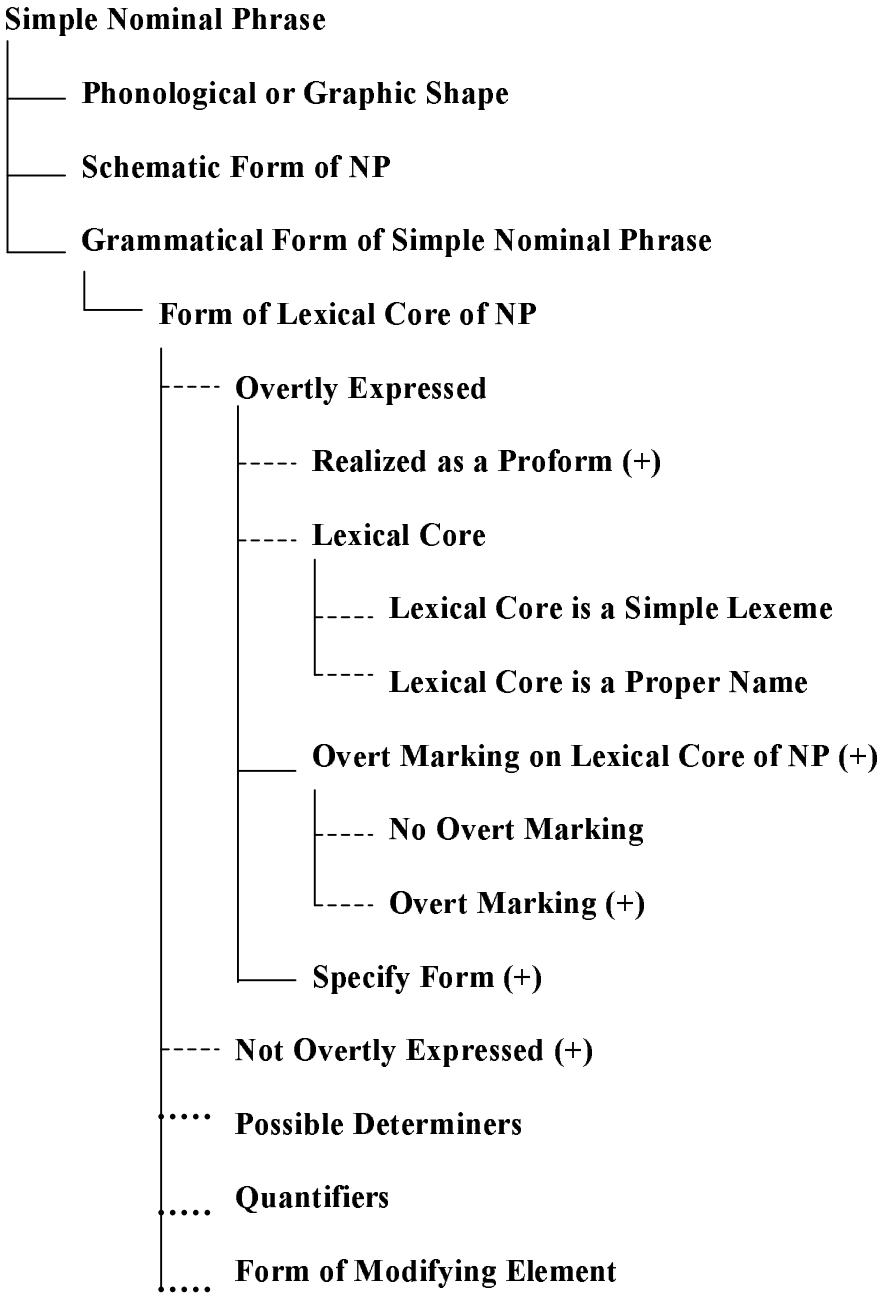
We then turn to the remaining individual components, all of which are offered as optional nodes, as the nominal phrase may or may not have a determiner, quantifier, or modifier.

OTHER COMPONENTS OF NP

We would first like to know if there is a determiner and, if so, if it is a **demonstrative or an article**, or perhaps a unit whose status is indeterminate with respect to these two categories. If it is an article, we need to know what type of article, i.e., definite or indefinite. Also, the form may be further specified for DISTANCE (PROXIMAL, MEDIAL, DISTAL with further distinctions), and a number of languages also further specify this information with respect to ALTITUDE and TYPE OF PERCEPTION, which are hence included here as optional nodes. Further, we would like to know if this is an INTERROGATIVE FORM and also of course, its actual FORM, any OVERT MARKING which it carries (including the possible use of a classifier and a description of its form and general functions!), and the OBLIGATORINESS of its use. As these criteria - with the exception of the classifier - have all been dealt with elsewhere, we need not go into detail on these points here.[6] Similar comments then pertain to the form of any explicit **quantifier** (again, including the possible use of a classifier!) and **possessive determiners**. Finally, the form of any modifying unit is fed into the data along the lines of the criteria discussed in section 5.2.3. below.

To sum up, below is a diagram of the uppermost nodes under "Simple Nominal Phrase".

[6] As the description of the classifier follows the same basic principles as all other parts of the grammar, the author felt that this need not be dealt with in a separate section. The basic issues here are of course its form and which element it attaches to (and where on this element), if it attaches to an element at all. On the content side, we are primarily interested in whether it is a general classifier or whether it is restricted semantically to certain types of entities, i.e., flat, round, long, etc.



COMPLEX NOMINAL PHRASES

As is to be expected, the description of complex nominal phrases is, as its name suggests, much more complex than that of simple nominal phrases. However, the basic principles involved are the same.

We first distinguish between COMPLEXITY AT ONE OR MORE SUB-NP LEVELS, i.e., where one or more of the levels WITHIN the single nominal phrase (determiner, quantifier, modifier, semantic head) is occupied by two or more components, such as *this or that book, green or blue eyes*. The other alternative here is that of COMPLEXITY AT THE NP LEVEL, i.e., a complex nominal phrase which consists of two (or more) components, both (or all) of which have the necessary minimal form required to be able to stand alone as a full nominal phrase. For example, *my book and your watch, blue eyes and red hair, me and you*, since *my book, your watch, blue eyes, red hair, me* and *you* could all also function as complete nominal phrases in

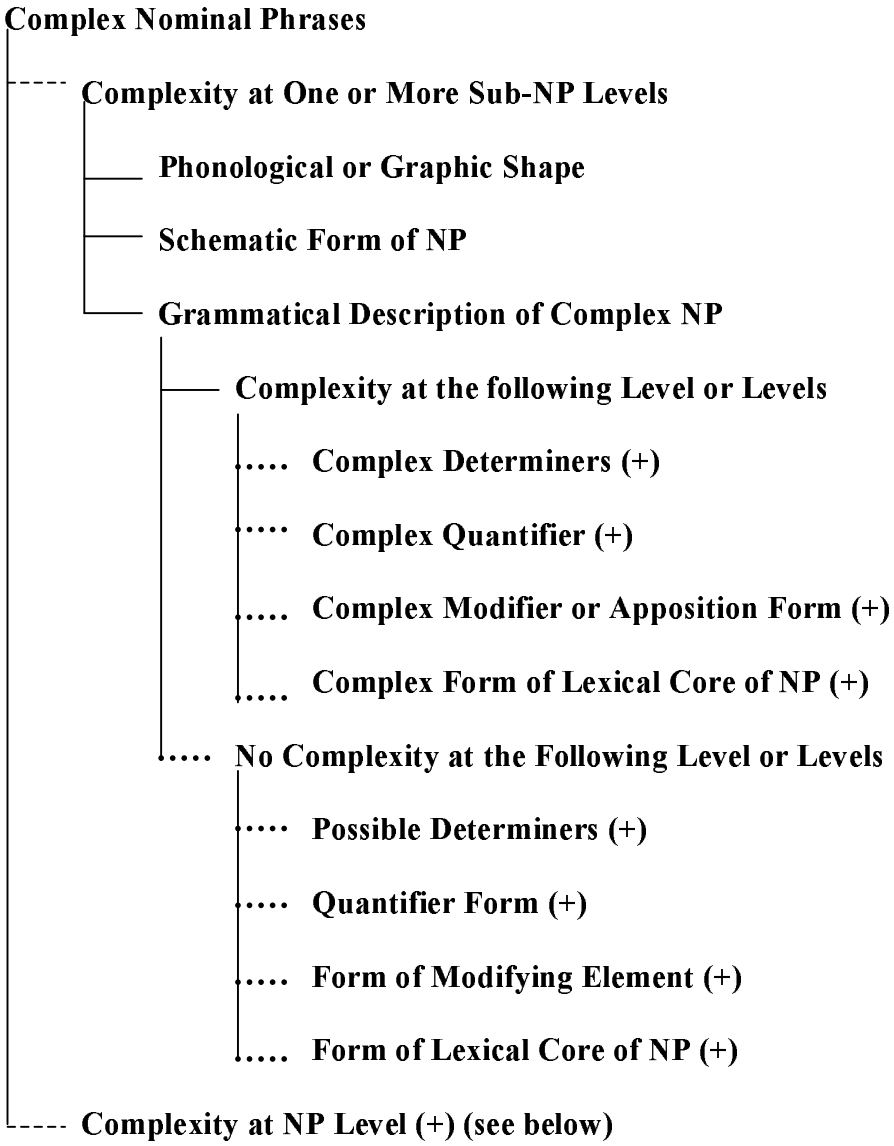
their own right. This distinction will help us considerably in describing the form of complex NPs.

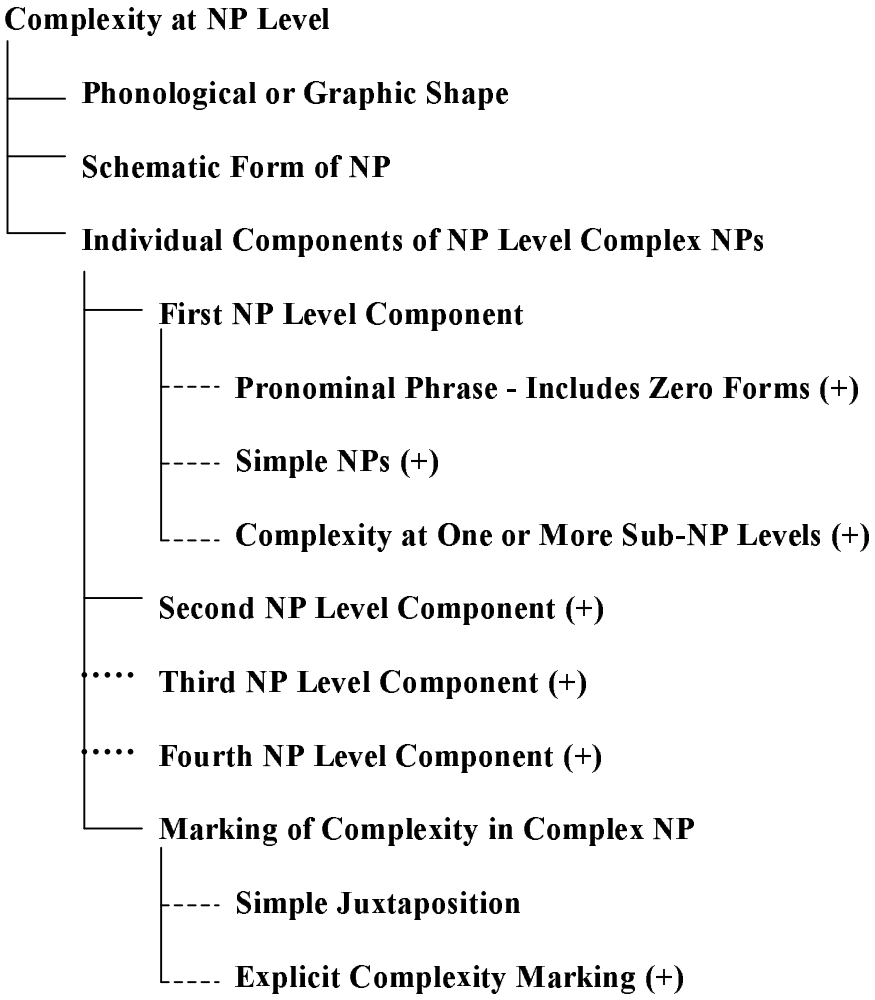
For example, if we have a complex NP with complexity at the uppermost NP level, we can simply treat this for the moment as two simple NPs and describe these separately. For example, the first "NP" may be a pronominal phrase and the second a simple full NP, or vice versa, etc. We then simply enquire as to how these two units are joined, for example through a conjunction (*and*, *or*), a clitic (Latin *-que*, Sanskrit *-ca*, and of course which word(s) it attaches to and where on this word/these words), simple juxtaposition, etc.

If there is complexity at a sub-NP level, we may have complexity at any combination of sub-NP-level units, while other sub-NP-level units are simple. Thus, we may have a simple determiner and a complex quantifier (*those two or three books*), vice versa (*these or those five books*), etc.

Thus, for each sub-NP level, we can have either a simple or a complex unit. Simple units are described in the manner described above under "Simple Nominal Phrases". Complex determiners, quantifiers, etc., are dealt with in a similar fashion to complex NPs with complexity at the NP level just described - each unit of a complex component is described separately, and then the manner in which these are joined is described.

Here, the uppermost nodes of the description of complex NPs:


Complex Nominal Phrases
Complexity at One or More Sub-NP Levels
Phonological or Graphic Shape
Schematic Form of NP
Grammatical Description of Complex NP
Complexity at the following Level or Levels
Complex Determiners (+)
Complex Quantifier (+)
Complex Modifier or Apposition Form (+)
Complex Form of Lexical Core of NP (+)
No Complexity at the Following Level or Levels
Possible Determiners (+)
Quantifier Form (+)
Form of Modifying Element (+)
Form of Lexical Core of NP (+)
Complexity at NP Level (+) (see below)



Adpositional Phrases

There is not much that need be said here, as we merely need to know whether the adposition is a preposition or postposition, or whether it can function as both. If it can be used either as a pre- or postposition, we of course need to know it status in the current example.

We then need to know the categorial status of the dependent of the adposition. This is generally a nominal phrase, however there are also many languages where adpositions can also take an entire clause, which must not necessarily be nominalized, such as in many of the Tibeto-Burman languages of Nepal. Thus, both of these possibilities will be offered here.

If the adposition takes a nominal phrase as its dependent, the information on this nominal phrase is exactly the same as in the section above devoted to the description of nominal phrases and need not be repeated here. If the dependent is a clause, this information is entered along the criteria discussed in sections 5.2.1. - 5.2.4.

Clause

The means of encoding this information is essentially the same as that given in sections 5.2.1. - 5.2.4. and need not be given here. These sections deal with the predicate, the complements of the clause, possible clause-level modification, and the operators tense, aspect/actionality, modality and negation.

The only additional information we need to discuss here is the manner in which this clause is marked as the complement of the predicate, for example through a conjunction (*that, if, whether*), etc., and the schematic pattern, which is then automatically fed into the **Inventory of Dependent Clause Patterns** (see section 6.2.4.).

5.2.3. Modification

We now turn our attention to the topic of MODIFICATION, or more precisely, MODIFICATION AND APPOSITION, as we can deal with the two phenomena using the same basic principles.

"Modifiers" in the CRG are all types of NONOBLIGATORY additional specifications which serve to further identify a referent or to qualify a proposition, predicate or some other constituent. Thus, this category subsumes, in addition to ADJECTIVES, ADJECTIVE PHRASES and ADVERBS, non-obligatory ADPOSITIONAL PHRASES (*in the evening, on Saturdays*) and "adverbial" clausal modifiers (*while running down the street; after I talked to her*), as well as both restrictive and non-restrictive RELATIVE CLAUSES.

Bhat (1994) and Dixon (1982) were of great assistance to us in determining above all the semantic criteria here, and many of the following criteria have been taken directly from these sources. Of course, as these two works are primarily dedicated to the adjectival class, we have also added further criteria, as we are not interested here only in adjectives but in all forms of modification.

Instead of beginning here as usual with the semantic content of the modifier, we must first differentiate between the MODIFIER itself and the MODIFIED ELEMENT. The status of the modified element is of course necessary to determine the scope of modification, and in many languages this will be closely related to the form of the modifier itself, although this is not necessarily the case. Does the modifier modify an entire PROPOSITION, a PREDICATE, the REFERENCE to an entity, another MODIFIER, etc.? We would also like to know here the categorial status of the modified element, such as NOUN, VERB, CLAUSE, ADJECTIVE, etc., as this will again often be closely related to the form of the modifier itself. At this level we are only interested in a very general description of the modified element, as a description of this element will be dealt with in detail in another part of the grammar, depending on the status of the modified element. For example, a modified "noun" will likely be dealt with under "Complementation", a modified verb will likely be dealt with in detail under "Predication", etc. Thus, the uppermost nodes of the tree are as follows:

Modification and Apposition

- **Modification or Apposition**
  - **Modification**
  - **Apposition**
- **Content of Modified Element**
  - **Modified Illocutionary Force**
  - **Modified Proposition**
  - **Modified Predicate**
  - **Modified Expanded Predicate ("verb phrase")**
  - **Modified Argument**
  - **Modified Modifier**
  - **Other**
- **Form of Modified Element**
  - **Modified Sentence**
  - **Modified Clause**
  - **Modified Verb Plus complement**
  - **Modified Verb**
  - **Modified Noun**
  - **Modified Proform**
  - **Modified Adjective**
  - **Modified Adpositional Phrase**
  - **Modified Adverb**
  - **Other**
- **Modifying Element (+)**
- **Order of Modifying and Modified Element (+)**

The remainder of the "tree" here deals with the content and form of the modifier itself, as well as any restrictions on its use. Here we are interested in three main areas: CONTENT, FORM and RESTRICTIONS ON THE USE AND FORM OF MODIFYING ELEMENT, all of which are required nodes:

Modifying Element

- **Content of Modifying Element (+)**
- **Form of Modifying Element (+)**
- **Restrictions on Use and Form of Modifying Element (+)**

As usual, let us begin with the semantics of the modifier, i.e. its "Content". The first piece of information we need to know about the modifier before being able to describe it in any detail is whether it is a PROPOSITIONAL or NONPROPOSITIONAL modifier, as the information we would like to know about the modifier depends above all on this question.

If it is a NONPROPOSITIONAL modifier, we would like to know the following criteria, among others:

- First, we would like to know the PERMANENCY STATUS of this modifier. It has often been noted that while adverbs tend to be restricted temporally, adjectives are compatible with both high and low time-stability (cf. *momentary* and *permanent*). Here, consultants of the grammar may be interested in any number of criteria, for example a comparasion of time-stability with the lexical class of the modifier, or with the status of the modified element, etc.

- Does the correct interpretation of this modifier depend on the presence (whether overtly stated or not) of another modifier? E.g. 'good', as in 'good ball' (i.e., 'good because it is round, light, etc.'), 'good boy' (i.e., 'good because he is well-behaved'), etc.

- We would then like to know something about the semantic properties of this modifier. This involves discussing both its UNMARKED (or "default") SEMANTIC TYPE as well as its ACTUAL SEMANTIC USAGE. For example, an English adjective such as "sick" has an unmarked semantic type of referring to an animate being in poor health. However, this same adjective can be found in environments where its usage must be considered marked, such as the exclamation *You're a sick person!* when referring to bad behaviour. Here, we divide the various lexical classes as follows, primarily for convenience:

- **Primary Modifiers**. These include, among others, the LEXICAL CLASSES given in Dixon (1982) and Bhat (1994). To these ten classes we have added a further number of classes (see below), such as LOCATION, DEGREE, etc., as we are not interested only in adjectives but in all types of modification and apposition. Here is a full listing of these classes. Except where these are mutually exclusive (e.g. "equative", "comparative", "superlative", etc.), these are all of the optional or inclusive type, so that a single modifier can belong to more than one class simultaneously.

Semantic Types of Primary Modifiers

- Degree
  - Type of Degree Relation
    - Simple Degree Modifier
    - Comparative
    - Equative
    - Superlative
- Distributive
- Human Propensity
- Locationals
  - Directionality
  - Position
  - Source
  - Destination
  - Path
- Nonlocal Goal Orientation
  - Preparedness
  - Usefulness
  - Purposive
- Place of Origin, Nationality, etc.
- Physical Property
  - Dimension
  - Health
  - Physical Property Pertaining to Substance (e.g. *hard, soft*)
  - Shape
  - Color
  - Age
  - Gender
  - Amount or Number
  - Animateness (+)
  - Value
  - Other Condition or Physical Property - Specify
- Possessive Relation
- Properties of Actions or Situations
  - Difficulty
  - Speed
  - Cause
  - Volitionality Properties (+)
  - Manner
  - Amount or Number
- Similarity
- TAM
  - Temporal Modifier
    - Anterior Only (+) (further specifications for remoteness)
    - Partially Anterior and Partially Simultaneous
    - Simultaneous Only
    - Partially Simultaneous and Partially Posterior
    - Posterior Only (+) (further specifications for remoteness)
    - Specified Period of Time
  - Aspect Modifier
  - Actionality Modifer
  - Modality
    - Source of Judgment is an Argument in the Sentence (+) (further specifications follow)
    - Source of Judgment is Another Person or Persons (+) (further specifications follow)
- Other

- **Secondary Modifiers**. These are linguistic expressions which normally predicate or refer to an entity but which may also function as modifiers, such as *water* in *water-jug* or *take-away* as in *take-away food.* This can further be divided into PHYSICAL ENTITIES and NON-PHYSICAL ENTITIES, with further divisions according to substance, predicate class etc., the details of which can be found elsewhere in the grammar and need not be listed again here (see sections 5.2.1. and 5.2.2.).

- A further criterion is that of SCALAR vs. NON-SCALAR modifiers, a topic which is closely interwoven with the following criteria: does the modifier have an ANTONYM, and if it does, is this a NEAR or EXACT antonym? Is one element of this pair the UNMARKED member (cf. the unmarked question *How old are you?* with the marked alternative *How young are you?*)?

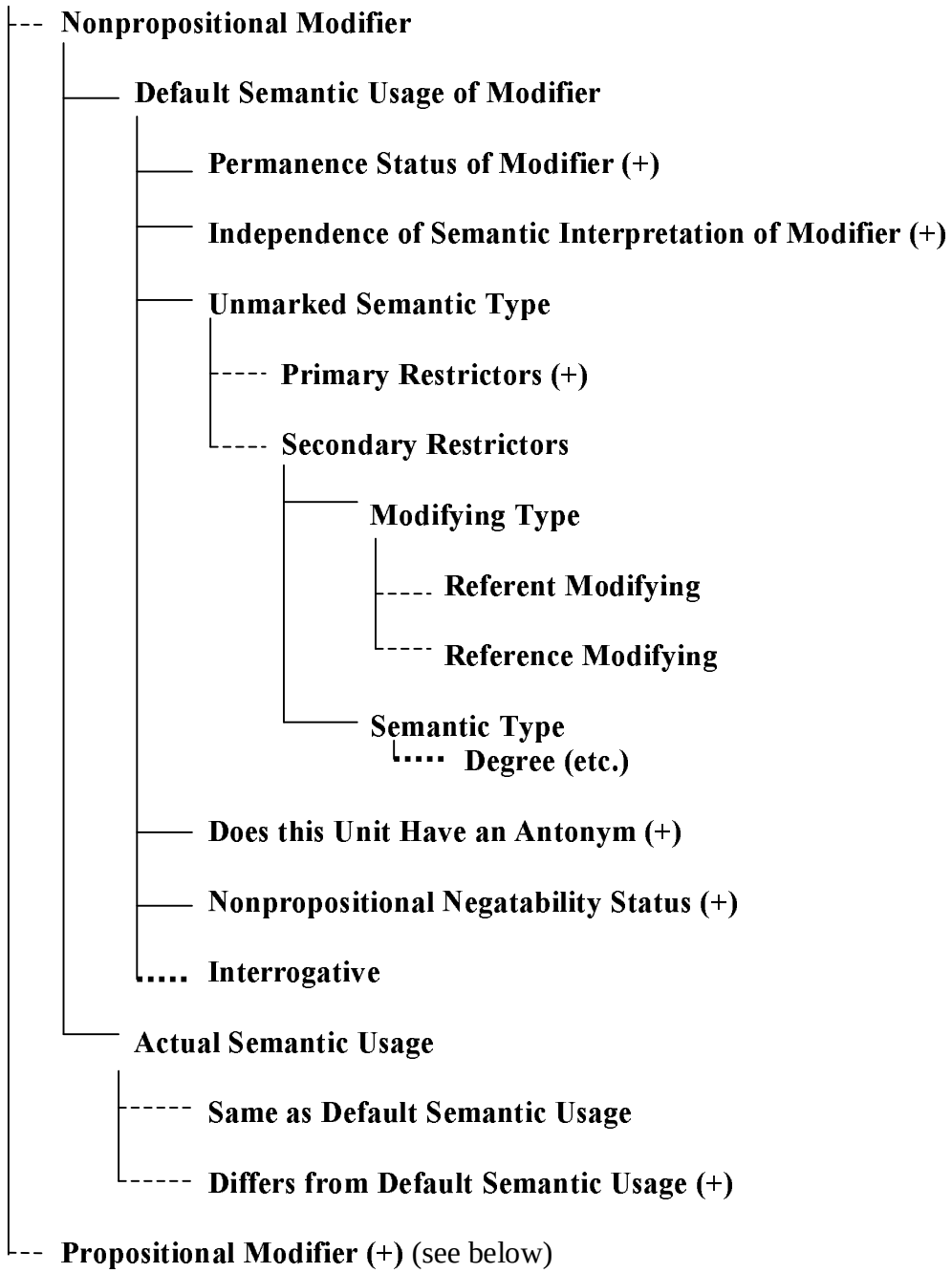
- Can this modifier be NEGATED by means of lexical negation? Is the form perhaps already negated? And what type of negation does this involve, CONTRARY or CONTRADICTORY?[7]

- Is this an INTERROGATIVE modifier?

- Finally, what is the MODIFICATION TYPE of this modifier, REFERENT MODIFYING or REFERENCE MODIFYING?

Below is a representation of the uppermost nodes just described:

[7] The manner of dealing with negation will not be discussed here in any detail but will be treated in section 5.2.4.4..



On the other hand, the modifier may be a PROPOSITIONAL modifier, in which case we would like to know the following details:

- Is it a RELATIVE CONSTRUCTION (in the sense of Keenan, 1985)? If so, is it a RESTRICTIVE or NON-RESTRICTIVE modifier / appositive?

- Does it (perhaps in addition to the last criterion) involve an INTERPROPOSITIONAL SEMANTIC RELATION to the main proposition? If so, what is the nature of this relation? This involves not only the use of subordinate clauses introduced by a conjunction, but also clauses containing a converb, of the type *(After) Having finished cleaning up*, among others. The interpropositional semantic relations which have been included (in addition to an "Other" node!) are the following:

- **Vague or Indeterminate**. While this may seem to be a strange starting point, much recent work on the interpropositional semantic relations found in sentences with an "adverbial" modifying clause whose predicate is a converbal form (cf. the various contributions in Haspelmath & König, 1995, Peterson, in preparation) repeatedly stresses that the interpropositional semantic relation is often vague. This is similar to the use of the conjunction *and* in English when joining two independent clauses, although with a converb, the clause whose predicate is a converbal form is necessarily dependent or subordinate.

- **Relation is Not Vague or Indeterminate**. These are: Relations involving EXPECTATION, such as CONDITIONALS, CONCESSIVES, CONDITIONAL CONCESSIVES and COUNTERFACTUALS.

- Tense / Aspect / Actionality / Mood (see "Semantic Types" above for more details)

- CAUSE

- PURPOSIVE

- LOCATIONALS, which may be divided up as follows: DIRECTIONALITY, POSITION, SOURCE, DESTINATION and PATH

- ATTENDANT CIRCUMSTANCES

- RESULT

- Finally, we must include here information on the clause itself, such as the PRESENTATION STRUCTURE of this propositional information as well as information on the possible use of a PREDICATE and COMPLEMENTS, as well as PROPOSITIONAL MODIFICATION of the propositional modifier. The principles involved here are the same as those involved with predication (5.2.1.) and complementation (5.2.2.) in general, and need not be dealt with again here in detail. Note that this node will also contain information on the OPERATORS (TENSE, ASPECT, MODALITY) of the propositional modifier itself.

Let us now turn to the FORM of the modifying element. In order not to assume that a certain type of modifying element, such as "color" or "relative construction", presupposes a certain form, all possible forms will be available for all possible types of modifying elements.

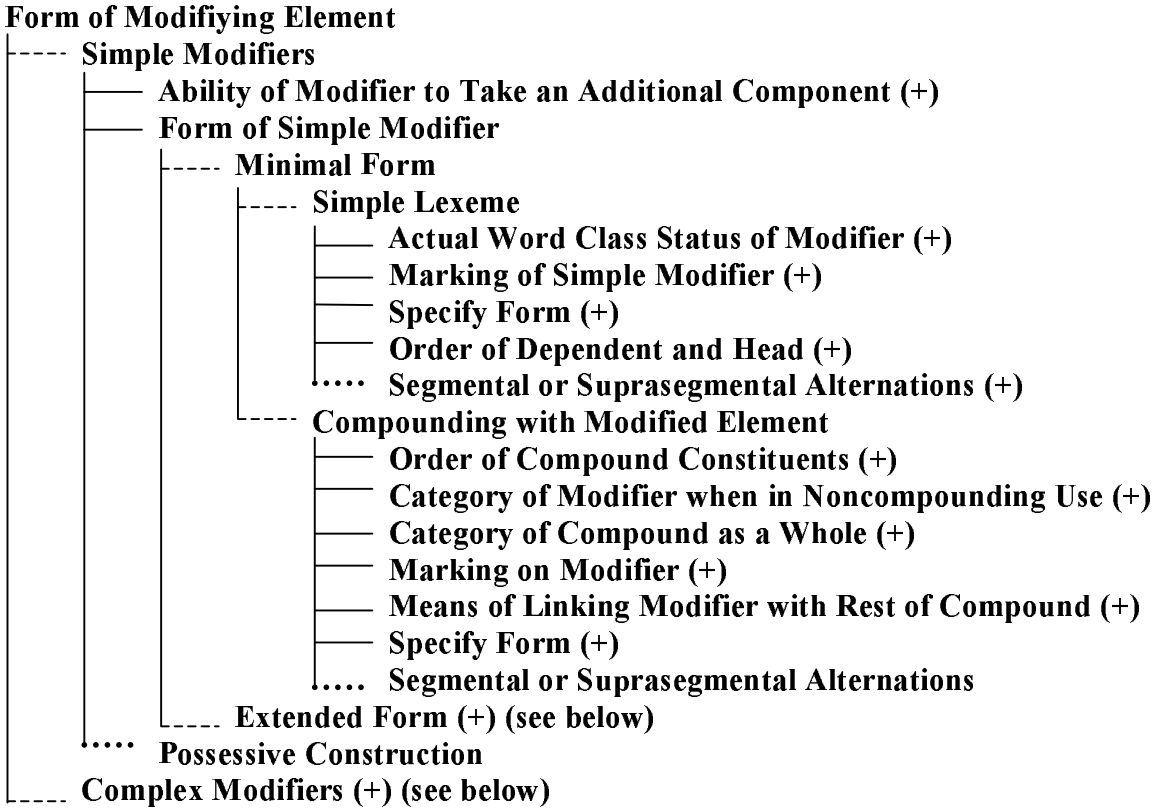
We would first like to know something about the actual form of the modifying element itself. For convenience, we divide the forms of modifiers into "Simple" and "Complex" forms. "Simple" forms involve those where the modifier is expressed by a single word, or forms a compound with the modified element, appears as inflectional or derivational marking on the head or, perhaps even through segmental or suprasegmental alternations. All other forms are "Complex". Simple forms are further divided into "miminal" and "extended" forms. Minimal forms have the minimal amount of marking which a modifier in this language must have to function as a modifier, while extended forms show further marking, such as derivational marking (e.g. adjectives derived from nouns, adverbs derived from verbs, etc.).

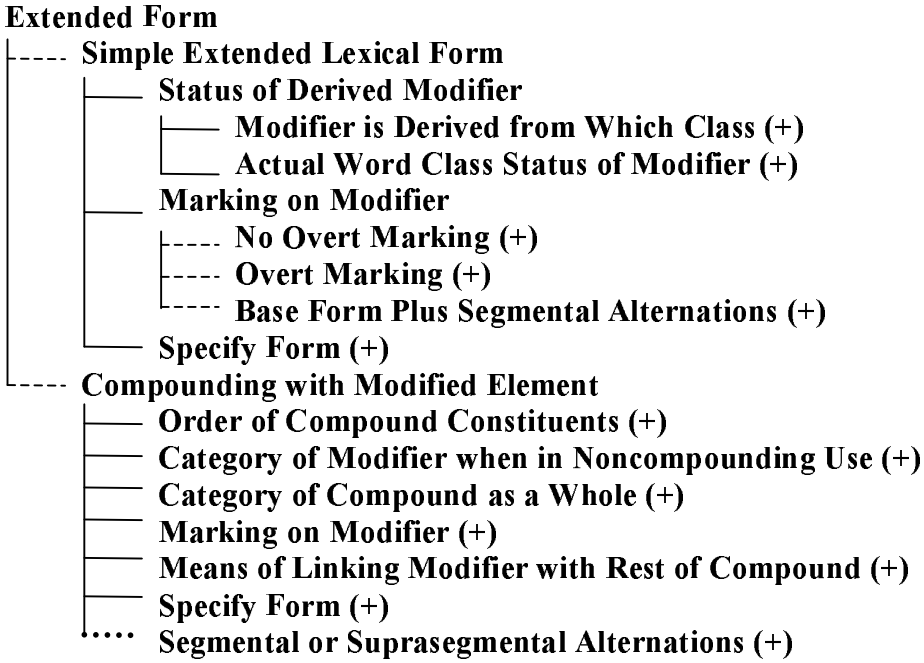
If we are dealing with a simple modifier, we would first like to know here - partly for the lexicon and partly for a complete description of the present example - whether the modifying element can take an ADDITIONAL COMPLEMENT or not, such as *proud* in English, which cannot

take a complement when in attributive function (**pround of his new job man* as compared with the German phrase *ein auf seinen neuen Job stolzer Mann*). If this modifier can take an additional complement, we would of course like to know whether the current example contains an additional complement and if so, we need information on both the content and form of this complement. As this information is identical with that for complementation in general (section 5.2.2.), we need not discuss this further at this point.

If the modifier does not form a compound with the modified unit, we need to know what marking the modifier takes (case, gender, etc.), and any "secondary" modifications, such as secondary (supra-)segmental alternations, and of course its form and the respective order of modifier and modified element. In the case of extended modifiers, we will also want to know what class the modifier is derived from, what overt marking - if any - is present to denote this derivation, etc.

If the modifier forms a compound with the modified unit, we need to know the respective order of the two units, the category of the modifier when not used in a compound, the category of the compound as a whole, whether a linking unit (or "Fugenelement") is used, as well as any marking which the modifier may take. Finally, we need to know the form of the modifier and, of course, any secondary segmental or suprasegmental alternations.
This has the following form:





If we have a complex modifier, it will have one of the following forms:

- a complex COMPOUND (e.g. *all-you-can-eat buffet*). Here, in addition to asking for information on the order of head and dependent, we need to know the status of each component of the compound, whether it is a modifier, determiner, noun, verb, adposition, etc. This information must be requested for each of the components individually. Finally, we need to know the category of the compound as a whole.

- POSSESSIVE CONSTRUCTION. As this information is dealt with elsewhere in the grammar, we need not go into detail. For the consultant, a search function will help locate all information desired.

- a NONPOSSESSIVE ADPOSITIONAL PHRASE

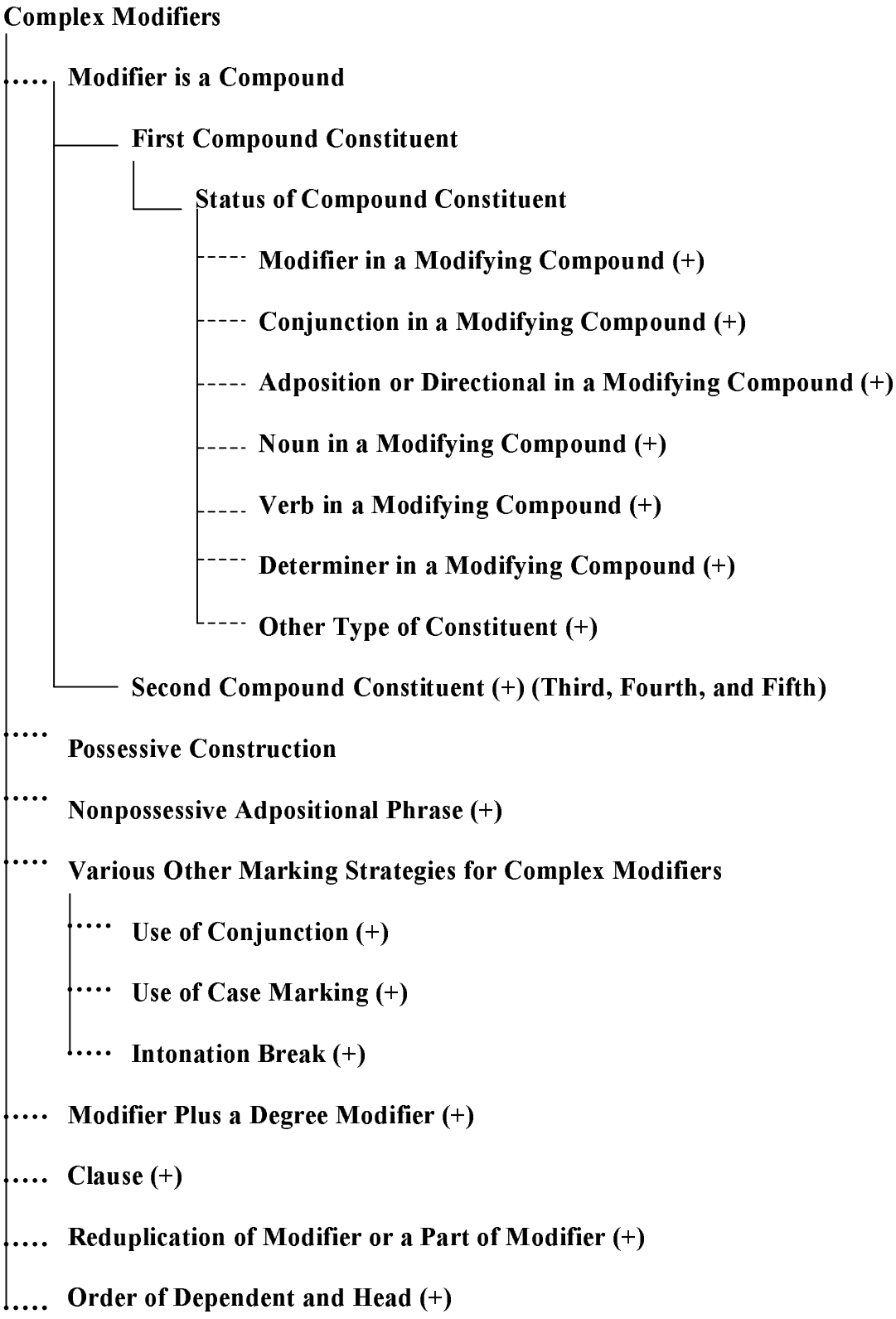
- MODIFIER PLUS DEGREE MODIFIER. Note that the degree modifier must also be entered separately, since it modifies the modifier. However, the two together of course also modify the head, e.g. *very big car*.

- CLAUSE. Of course, the modifier may have a clausal form, such as relative clauses in a number of languages. The information contained here is virtually identical with the information presented in sections 5.2.1. - 5.2.2. - i.e., we need to know information about the predicate, the complements, etc.

- REDUPLICATION of a modifier or a part of the modifier.

- Finally, there are a number of "miscellaneous" types of marking which can occur with any number of these marking strategies, such as the use of conjunctions, case marking, or an intonation break.

Without going into detail here, the node "Complex Modifiers" has the following form:



Under "Order of Dependent and Head", we must first enquire as to the continuous status of both constituents, as either or both may be discontinuous. Then, the possibilities "Juxtaposition" and "No Juxtaposition" are offered, and all corresponding combinations are offered, the details of which we need not go into here. We also enquire into the obligatoriness of this order.

We now come to the last aspect of modification and apposition, that of the RESTRICTIONS on either form, content or both in certain situations.

This is a required node, that is, the author must click on this node and chose at least one of the two following (mutually exclusive) possibilities: NO RESTRICTIONS, RESTRICTIONS.

If there are restrictions, then we would of course like to know what type of restrictions these are. Below is a list of the restrictions which have come to the attention of this author and which have thus been included in the grammar. It is perhaps needless to say that this list cannot be expected to be exhaustive. Hence, an "Other" node has been included as well. All restrictions are of the optional type, so that more than one may be chosen, as necessary:

- ATTRIBUTIVE, PREDICATIVE OR NOMINAL USE RESTRICTIONS. That is, a modifier may only appear in one or two of these uses.

- INFLECTIONAL RESTRICTIONS. This may be restricted to certain TAM categories, case, number, or person restrictions, certain valencies, or gender or noun classes.

- RESTRICTIONS ON THE USE WITH DETERMINER AND OR PRONOUNS. While some modifiers may only occur with an article, others may not be compatible at all with a determiner. Also, it is quite common that proforms in general are not compatible with modifiers.

- RESTRICTIONS ON RELATIVIZATION

- RESTRICTIONS ON THE USE WITH CLITICS

- RESTRICTIONS ON THE GRAMMATICAL RELATION WHICH THE MODIFIED EXPRESSION CAN HAVE IN AN UTTERANCE

- RESTRICTIONS ON THE TYPES OF MODIFIERS THAT A MODIFIER MAY MODIFY

- RESTRICTED TO EQUATIVE, COMPARATIVE OR SUPERLATIVE CONSTRUCTIONS (or not allowed with these constructions)

- Restrictions on ECHO-WORD FORMATIONS

- Finally, the modifier may be restricted to the current context - i.e., it is only found in a fixed expression.

5.2.4. Operators

The following section actually consists of four sections, each of which is properly a chapter in its own right. These are TENSE, ASPECT AND ACTIONALITY, MODALITY and NEGATION. Each will be dealt with individually. Common to all four areas is the usual division between the CONTENT of the respective "operator" and the FORM through which this content is expressed, if it is expressed at all. While negation is, to our knowledge, always overtly expressed, this may not be the case with the other three "operators", as a large number of languages have an "unmarked" form or are underspecified with respect to one or more of these three operators, especially tense and aspect.

5.2.4.1. Tense

Much of the following derives either directly or indirectly from Comrie (1985), while often deviating somewhat from this as necessary. Let us begin with the CONTENT side of this operator.

First, there is the question as to whether we are dealing with temporality at all, that is, is an utterance necessarily connected to time or, as in gnomic propositions, is the utterance considered "atemporal". Comrie (1985:40) notes with respect to this category, "Just as we claim there can be no separate habitual tense, distinct from the present, likewise we are led to claim there can be no universal tense, i.e. a tense that is used for truths that hold at all time. Thus, a sentence like *cows eat grass* is claimed to refer only to the present moment, the interpretation of this as a universal truth being on the basis of structural and extralinguistic factors beyond the meaning of the present tense."

While there is good reason for assuming that there is no such thing as a "universal tense", this is an empirical question and from the point of view of a cross-referential grammar such as the present one, it would be wise not to exclude such a possibility. Hence, the first question in our "tree" will deal precisely with this topic, i.e., is there a relation to external time or is there atemporality?

Another interesting point mentioned by Comrie is the possibility that the relation to external time is CYCLIC, i.e., it does not refer to one particular time or to a particular duration of time, but to a situation which is "located relative to some cyclically recurring event, of which the ones known to me to be relevant are different parts of the 24-hour cycle, i.e. morning, afternoon, evening, day, night." (Comrie, 1985:17).

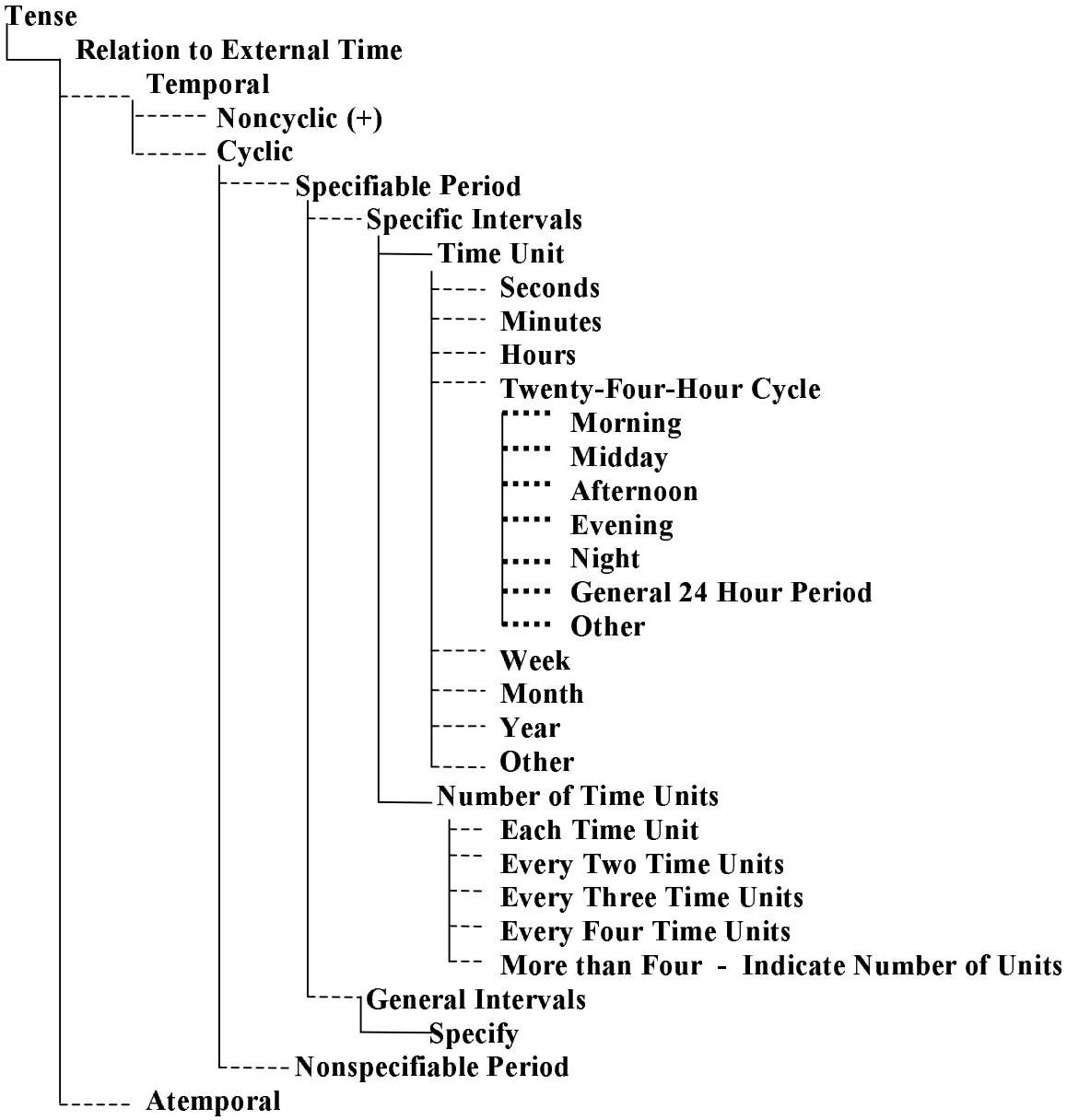
Again, a grammar which is intended to be used for all human languages and which makes as few *a priori* assumptions as possible should not exclude such a possibility. Also, although Comrie only explicitly mentions cycles on a 24-hour basis, there is no reason to assume, if such a category does occur, that it could not be related to an annual cycle (e.g., harvest time, planting time, the rainy season, etc.). Again, this is an empirical question. Hence, after determining that an utterance has a temporal relation to external time, the next necessary criterion is whether this temporal relation is CYCLIC or NONCYCLIC.

Finally, if it is cyclic, we would like to know the nature of this cycle, for example if this deals with a SPECIFIABLE TIME PERIOD (24-HOUR CYCLE or parts thereof, MONTHS, YEARS, etc.) and how many of these units are involoved. This will most likely be each and every time unit involved (for example, every morning, every evening), although the possibility should be included that it will refer to every two, three or four of these units. Also, since different

languages (and cultures) divide the day up differently, these units within the 24-hour period will all be optional nodes, so that the author may combine these as necessary.

It may also be that we are not dealing with a specifiable period of time, but merely a situation which occurs at GENERAL INTERVALS other than those offered here. In this case, the author is required to specify what this interval may be.

The last alternative is that there is NO PARTICULAR SPECIFIABLE PERIOD, i.e., we have an action or situation which holds again and again on a regular basis, but not on any specific temporal basis. This is shown in the following diagram:



Before proceeding, we must introduce three terms which will be central to the following discussion:[8]

[8] The following definitions make no pretence at being complete but are simply intended to give the reader some orientation for the following discussion. As the present work is merely intended to be a general guide to the use and layout of the grammar, we will not deal with these theoretical issues any more than required, especially as

Speech-Act Time - The moment in time at which the speech act takes place.

Event Time - The moment in time at which the event or situation referred to held, holds or will hold. It may be a specific moment in time or a prolonged period.

Reference Time - Either a point in time or a prolonged period of time, which may or may not be identical to either the speech-act time or event time, or both, which serves to provide a point in time as a reference point to which the event time may be located.

Some examples: (SAT - speech-act time; RT - reference time; ET - event or situation time)

1. I'm reading a book now.
SAT = ET = RT - present moment

2. Before I got there yesterday, he had already finished cleaning the house.

SAT - present moment
RT - when I arrived yesterday, exact moment unspecified
ET - prior to RT

3. He told me yesterday that he would do it.

SAT - present moment
RT - yesterday, exact moment unspecified
ET - sometime after reference time, unspecified as to whether this is before or after SAT

By defining the relations between these three times,[9] it is possible to describe sufficiently any temporal relation holding between an event or situation and the moment when this event or situation is referred to in an utterance. For example, these times may be simultaneous, such as in the case of the present, where all three are identical. Or reference time and speech-act time may be simultaneous, while the event time precedes this moment, e.g. the "present perfect" in English. Finally, event and reference time can coincide and precede (past) or follow (future) the speech-act time.

Hence, we must define the relation between the Speech-Act Time (SAT) and the Reference Time (RT), as well as between the Reference Time and the Event (or Situation) Time (ET). As the principles at work here are the same for both relations, there is no need to discuss both relations here in detail but merely to discuss the principles at work using one of these relations as an example.

---

there already exists an extensive literature on this topic. The interested reader is referred to Comrie (1985) and Kiparsky (1998), both of whom refer to Reichenbach (1947).

[9] A second reference time has also been incorporated into the grammar, although it is not yet clear whether this will be necessary. The principles involved here are the same as those described above when using one reference time, except that these relations become much more complicated through the presence of a fourth time. As there is nothing to be gained in a discussion of a model using four points of temporal reference, we will restrict our discussion in the following to a model using three temporal points of reference, i.e., with only one reference time (RT).

Using the relation between the Reference Time and the Speech-Act Time as an example, note that the two times may at least partially overlap, or may have no temporal overlap whatsoever. If there is no overlap, than either the RT precedes the SAT or the RT follows the SAT.

If there is overlap, this may be partial or complete overlap, i.e., the two times may be partially or completely identical. If they are only partially identical, there is the question of whether RT precedes and overlaps with SAT or whether RT overlaps and follows SAT. There is also the possibility that the SAT is completely contained within the RT or the (theoretical) possibility that the RT is completely contained within the SAT.

There is also the question of REMOTENESS between these three temporal points of reference. It is well known that many languages have special tenses for actions which occurred in the immediate past as opposed to those which occurred long ago. There are also a number of other possibilities which languages exhibit in this respect. In this connection Comrie (1985: 87f) writes:

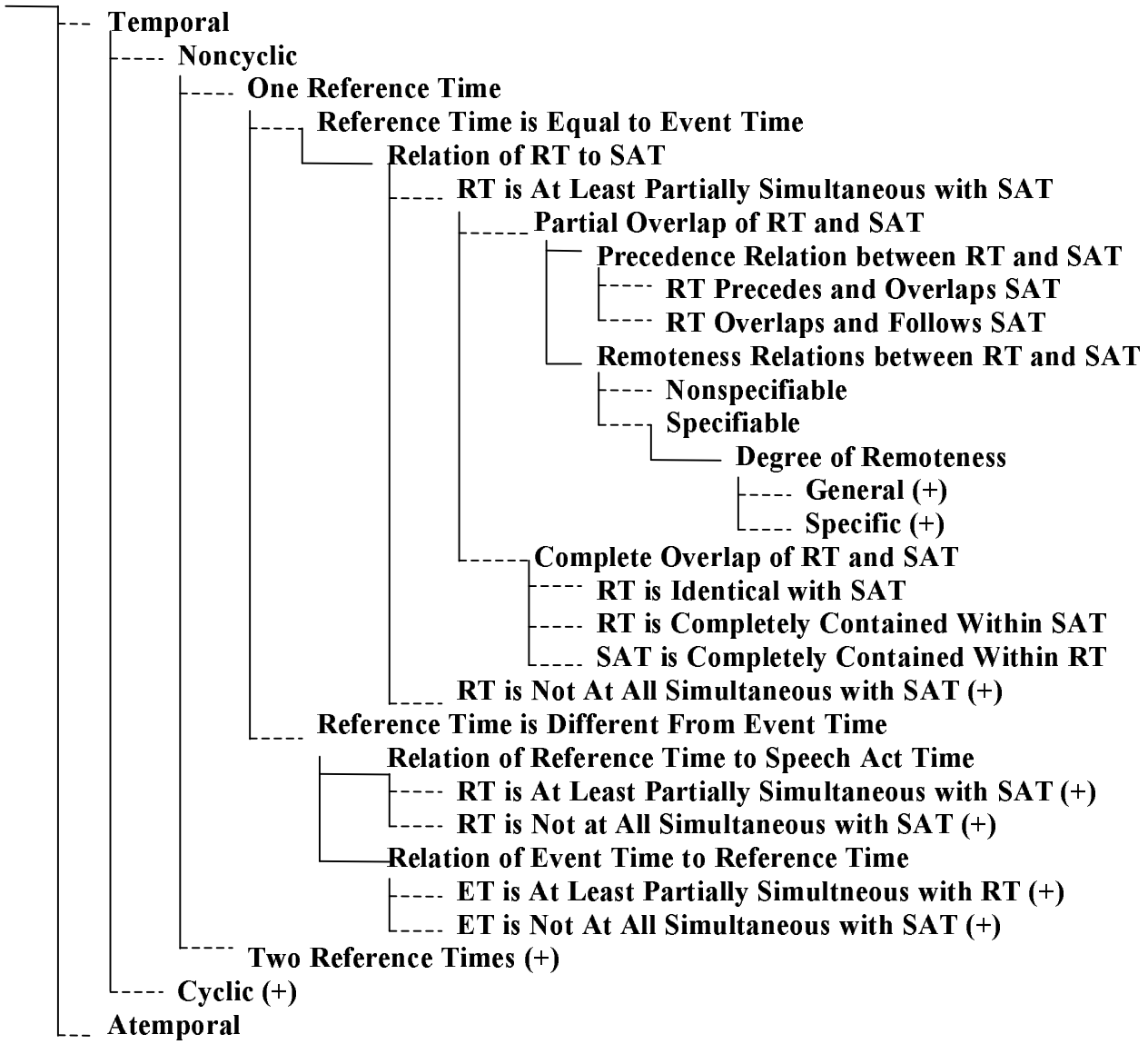
> "In relation to the number of distinctions, it is also necessary to specify precisely what the cut-off points are for the various distinctions, e.g. a distinction between 'recently' and 'longer ago', or between 'today' and 'earlier today', or a distinction between 'this year' and 'before this year', etc. [..] The commonest cut-off point seems to be that between 'today' and 'before today', to which we can give the Latinate names hodiernal and pre-hodiernal. Another common cut-off point is that between 'recently' and 'not recently', i.e. between recent and non-recent. [..] Another cut-off point found recurrently in the past is between 'a few days ago' and 'more than a few days ago', i.e. non-remote versus remote. Other cut-off points seem to be language specific. Thus the Mabuiag dialect of Kalaw Lagaw Ya has a special tense for 'last night'. Distinctions with cut-off points prior to 'a few days ago' are also found: one especially prolific system is that of Kiksht, where in addition to cut-off points based on the change of days (e.g. 'today' versus 'yesterday'), there are also cut-off points based on the change of years (e.g. 'this year' versus 'before this year'); but so far, this example stands unique.

We thus need an exact specification of the temporal distance between the Speech-Act Time and the Reference Time, as well as between the Reference Time and the Event (or Situation) Time. This is achieved in the following manner.

First, the DEGREE OF REMOTENESS can be either GENERAL or SPECIFIC in nature. If it is GENERAL, than the question remains as to whether it is REMOTE (past) or NONIMMEDIATE (future) or NONREMOTE (past) or IMMEDIATE (future). If it is a specific interval of time, it may be a matter of a single DAY, in which case it may be either EARLIER OR LATER ON THE SAME DAY or BEFORE OR AFTER THE SAME DAY, i.e., on the same day or one day earlier/later.

The same general possibilities hold for intervals of TWO DAYS, THREE DAYS, WEEKS, YEARS, although this is somewhat more complicated. For example, with TWO DAYS it may be before these two days, the same day and the evening before, the same day and the following morning, today vs. yesterday, etc. As this is merely a question of the combination of these details, we can skip over this section here. As usual, it may not have been possible for us to anticipate all possibilities, so that the node OTHER is provided here as well.

For a view of the uppermost nodes under the content side of TENSE, confer the following diagram.



After discussing the temporal content in some detail, let us now turn to the question of form, i.e., how is this temporal content expressed, if it is overtly expressed at all.

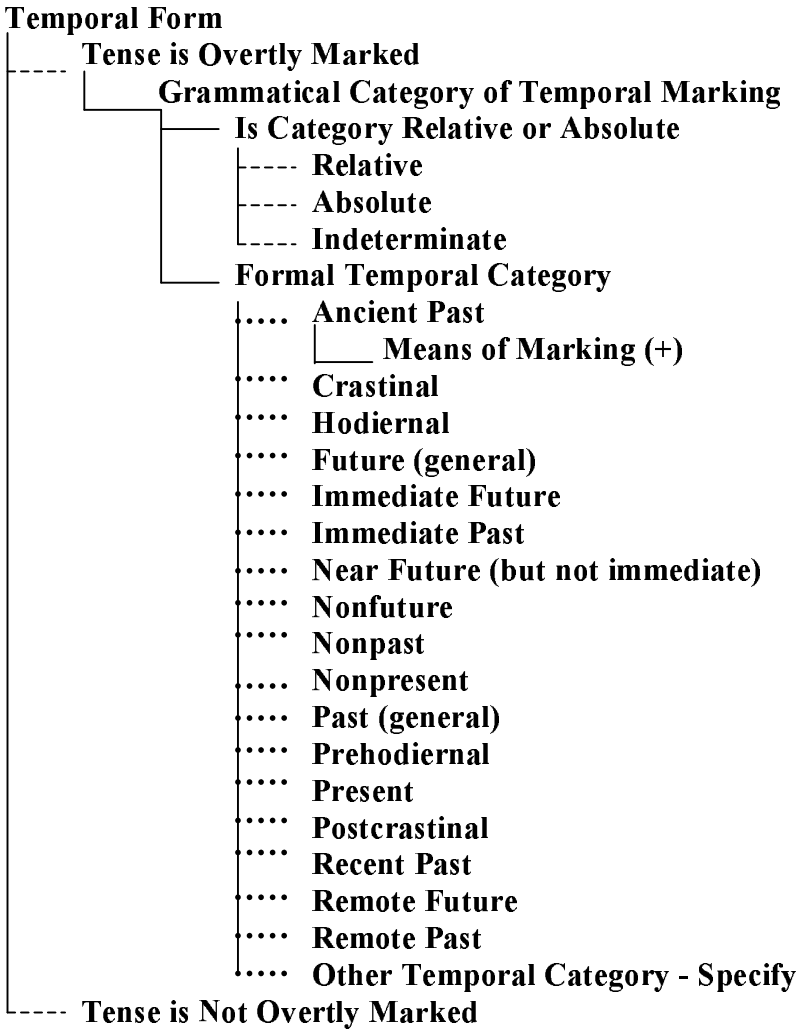
The first question is of course WHETHER the temporal relation is expressed at all, as "tenseless" languages are by no means uncommon. However, temporal MARKING here is also considered to include specification by means of a temporal modifier, so that the presence of an adverbial temporal specifier in these languages would also be considered to be the presence of temporal marking, not just tense marking on the predicate itself. Cases with no overt marking are thus restricted to either "tenseless" languages where no indication of the temporal location of the event/situation is given or languages which generally have tenses but where some temporal relations (generally an "unmarked" relation) are either not explicitly marked or are underspecified.

If tense is indicated, we would like to know which temporal category is meant. This list includes the following categories, taken primarily from Bybee et al. (1994):

ANCIENT PAST, CRASTINAL, HODIERNAL, FUTURE (GENERAL), IMMEDIATE FUTURE, IMMEDIATE PAST, NEAR FUTURE (but not IMMEDIATE), NONFUTURE, NONPAST, NONPRESENT, PAST (GENERAL), PREHODIERNAL, PRESENT, POSTCRASTINAL, RECENT PAST, REMOTE FUTURE, REMOTE PAST, OTHER TEMPORAL CATEGORY - SPECIFY

Following each of these categories is a required node named "Means of Marking". Here, we are concerned with the way in which the respective category is expressed, such as through a modifier, inflectional or derivational marking, adpositions, particles, segmental alternations, etc. As this topic is dealt with in detail elsewhere in the present paper, it will not be dealt with here. For detailed information, cf. section 5.2.5.

Finally, we would like to know if this temporal category is a RELATIVE or ABSOLUTE category. With the exception of the node "Means of Marking", the "tree" thus has the following form:



5.2.4.2. Aspect and Actionality

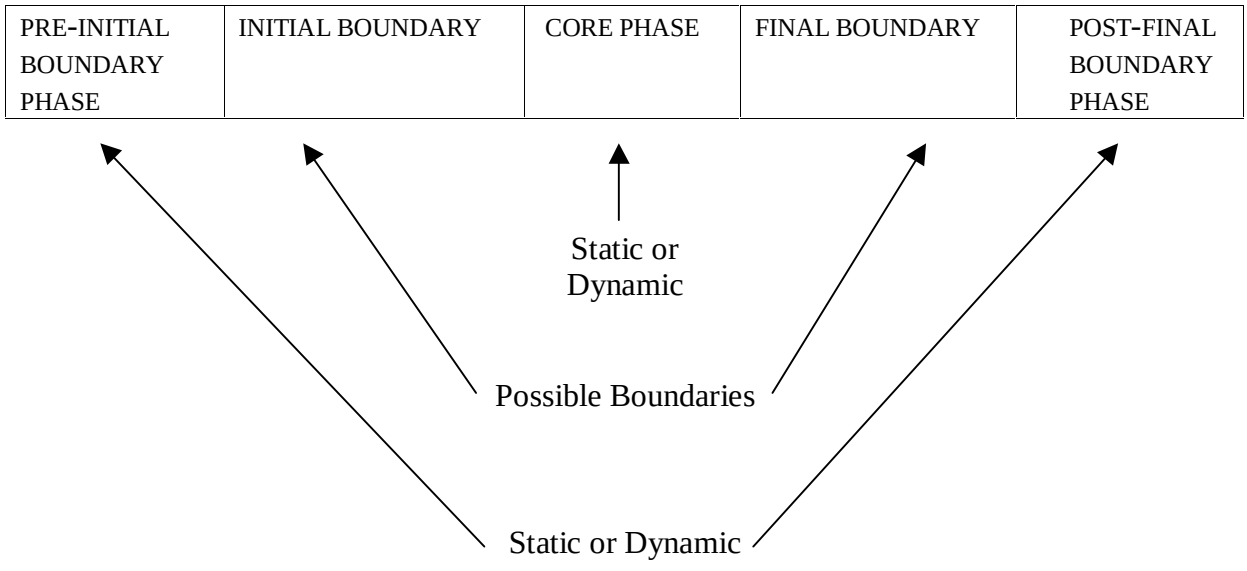
A number of sources provide the criteria for the treatment of aspect and actionality. These include Breu (2000), Comrie (1976), Sasse (1991b, 1991c), Carlota Smith (1991) and Van Valin & LaPolla (1997). The author was also able to participate in a working seminar on the subject of aspect and actionality at the University of Zurich in June, 2000 and gather information from the several talks held there.

Although there are several aspects of "aspect" on which virtually all of the authors mentioned above would probably agree, there are also a large number of differences in their approaches to the topic, and it was therefore our goal to provide as general an outline here as possible, so that as few *a priori* assumptions as possible need be made. We thus provide only a very basic "skeleton" of criteria which we believe apply to any event or situation.

The only assumptions which we were forced to make, and which seem reasonable to us, are the following, on which the above mentioned authors all would most likely agree: aspect and actionality are primarily concerned with the presence or absence of boundaries or limits. The perfective is primarily concerned with the presence or setting of boundaries, while the imperfective is concerned with their absence or their removal. This will depend to a large extent on the inherent lexical aspect or "actionality" of the predicate and its complements.

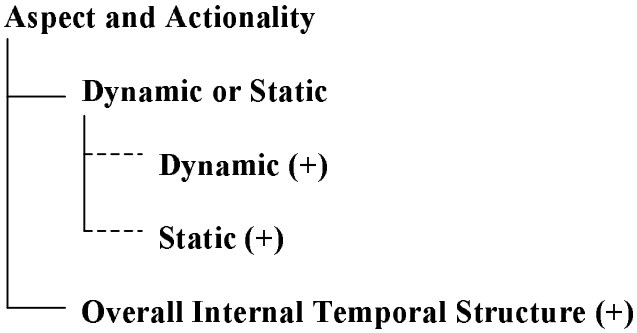
Further, both events and situations may be either SPONTANEOUS, such as the presence of stars in the sky or rain falling, while others are INDUCED, i.e., they require the presence of an external circumstance to take place or to continue (*sit, stand, smile, fear, be afraid of,* etc.). Whether or not this is volitional is not of interest here but is dealt with elsewhere (see "Predication", section 5.2.1.). It is important to stress here that we are referring to "induced" or "spontaneous" in a purely linguistic sense. That is, while we know that rain falls due to the force of gravity (among other reasons), the utterance *rain falls* in English does not presuppose the presence of any external circumstances - rain simply falls, with no cause, as far as the English language is concerned. In a language where "Gods pour water", this may of course be different.

We have also had to make a (minimally) arbitrary decision as to the number of individual phases which may be involved in events and situations. The maximal structure allowed for here is the following, represented schematically:



That is, we do not assume the possibility of more than two boundaries with intervening periods of states or dynamic situations but assume that such situations will be dealt with by other means, such as habitual or iterative categories, etc.

In the following, the uppermost nodes of Aspect and Actionality are presented:



Thus, whatever event or situation we are describing, it will be either DYNAMIC (or "non-static") or STATIC (i.e., pure-state). "Static" here refers to situations in which there is no change involved for the period in question - one temporal point within the range in question is exactly equivalant to every other point within the same range and there is no perceptible movement. The only possible boundaries here are purely arbitrary and involve the input of external energy. That is, if no external energy is applied, the situation described will continue on indefinitely. Note that this is not to say that states are never induced: One may, e.g., "be afraid" of something (e.g. heights). Thus, this state is INDUCED, as there is no fear without the knowledge that there are heights or the experience one has made in such situations. However, there is no external energy applied here to "make fear" - it is a pure state. There is thus a difference between external "energy", which keeps a situation going or ends it, and external "circumstances", which can be responsible for the presence of a situation.

"Dynamic" actions or situations, on the other hand, require some form of outside energy to take place or continue. One temporal point within the range in question will not be exactly identical to every other temporal point within the same range and there may be perceptible movement. This encompasses spontaneous change, the maintenance of a situation through external energy (e.g. *holding a door open*), situations where there is perceptible movement (e.g. *rain falling*) and activities in general (for a more detailed discussion, cf. Van Valin & LaPolla, 1997:84).

Cf. the following examples, which should help to clarify the situation:

Dynamic

(1) He held the door open.
(2) The door was held open.

Static

(3) The door was open.

Here, a detailed knowledge of the language is necessary, as one must know whether a particular predicate in a particular language implies external energy (such as *hold* in English) as opposed to predicates which require no external energy to continue (*be open*).

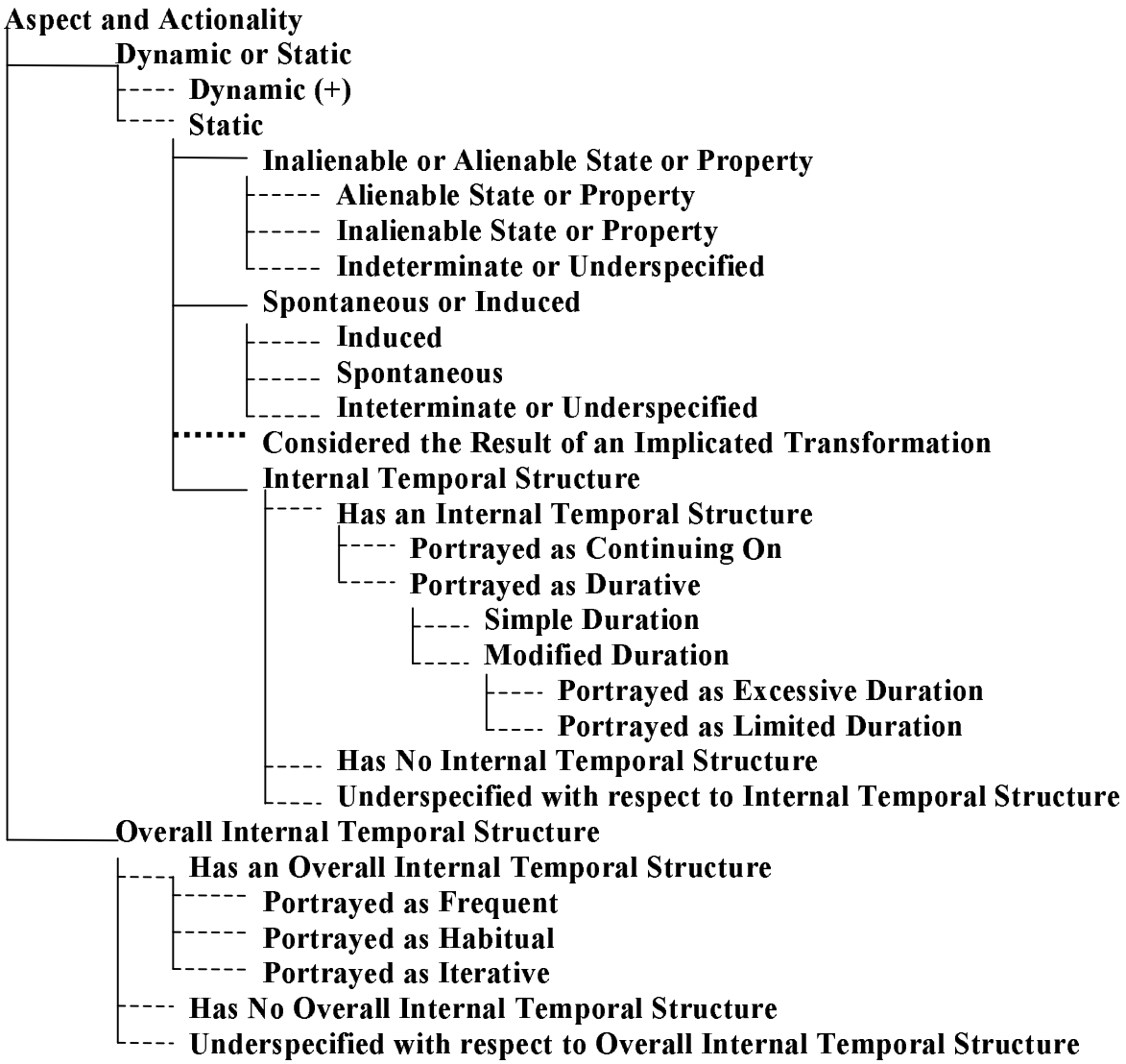
In effect, this relates to the types of possible boundaries involved. That is, although *hold*, as well as *be open,* have arbitrary initial and final boundaries, *hold* depends on the application of external energy and the end of the situation comes with the end of this external energy.

On the other hand, with *be open* the situation holds as long as there is NO external energy applied - that is, the boundaries here have diametrically opposed requirements to those of *hold*: Without external energy, nothing will change. To end the state, one must APPLY energy, instead of ending this external energy, as was the case with *hold.*

We would like to know about the OVERALL INTERNAL STRUCTURE of an event or situation. That is, the overall internal structure including possible boundaries and pre- and post-boundary phases, such as frequent, habitual or iterative, or does the event/situation perhaps have no overall internal temporal structure, or is the overall internal temporal structure underspecified, as e.g. is the case with the English "past tense" (cf. *I worked there*, where we do not know whether this was a single occurrence (*I worked there once*) or perhaps frequentive, habitual or iterative (*I worked there when I was a kid, I worked there every summer,* etc.). The overall internal structure refers to the ENTIRE situation - i.e., all phases which may be found between the initial and final boundaries or, more generally, between the pre-initial and post-final phases.

This is not to be confused with the INTERNAL TEMPORAL STRUCTURE. This refers to the the INTERNAL temporal properties of an action or situation, i.e. durativity and continuousness. There may explicitly be no internal structure, as with most perfective predicates, or the internal temporal structure may be underspecified. As an example of this last possibility, cf. again the English phrase *I worked there,* which gives no information as to whether this situation continued over a longer period of time (e.g. *for years and years*) or whether it is portrayed as a single situation with no internal temporal structure.

Let us now take a brief look at the semantic description of STATES. This is presented in the following diagram:



There are two criteria in this "tree" which we have not yet discussed, CONSIDERED THE RESULT OF AN IMPLICATED TRANSFORMATION and INALIENABLE OR ALIENABLE STATE OR PROPERTY.

The first of these two criteria is attached to the "Static Tree" as an optional node, as it may or may not hold in a specific case. What is meant here are states which are considered to have arisen as the result of a previous change of state, whose results they are. E.g. *The glass is broken*, which refers to the state of glass at the moment of speech and implicates that this present state is the result of a previous event, namely that of *breaking*. However, this is not explicitly stated in the sentence above, it is only implicated. This is the well known class of RESULTATIVE forms, the tests for determining which are well known and need not be given here (e.g., compatability with adverbs such as *still*, etc.).

The other not previously mentioned criterion here, that of ALIENABILITY, refers to the semantic differences in pairs such as *Firemen are altruistic* (inalienable) and *Firemen are available* (alienable). In the first example, a statement is made about an inherent property of firemen in general (whether or not it is true), while the second example tells us about the non-inherent state in which they are found at a particular moment. Other pairs include *She's a woman* (inalienable) and *She's here* (alienable), while some may be ambiguous without proper context, e.g. *She's sick.* This is also closely related to the individual/stage-level distinction, while not necessarily being identical to it, and there is of course room here for language-

specific criteria. For example, while *She's an adult* may reasonably be considered an inalienable property in English, other languages may treat this as an alienable property, i.e., adulthood is simply one of a number of phases through which individuals may pass. The author will have to make this decision him-/herself, based on the morphosyntactic properties of the language involved.

Turning now to the semantic description of DYNAMIC situations, there are a number of criteria we are interested in. For example, how many possible boundaries are involved - one (*find out about, learn of, get to know, become*) or two (activities in general). Are they inherent boundaries (*become, eat up*) or arbitrary (again, activities in general). Also, of these possible boundaries, are they explicitly transgressed (*I stopped reading*), explicitly not transgressed (*I am reading*) or underspecified in this respect (e.g., with the past tense in a large number of languages).

Note that it is certainly possible for a predicate to denote an inherent final boundary while at the same time denoting that this final boundary has not yet been reached. A prime example of this is found in the English sentence *He's eating up all the food.* Here, the predicate has an inherent final boundary, denoted by the marker *up* in the predicate *eat up* (as opposed to *eat*). However, the use of the "present progressive" denotes that this telic activity has EXPLICITLY not reached its natural conclusion, i.e., he is STILL IN THE PROCESS OF a goal-oriented activity.

However, there is much more involved here than merely the presence or absence of boundaries and their number. For example, a predicate with one boundary may consist only of a boundary (*reach, cross over*) or a boundary plus either an initial or final state or activity. While this may not be a common property of European languages, it is an exceptionally common pattern in many Asian languages (cf. e.g. Ebert, 1995), where e.g. a lexeme which translates as 'know' may actually mean 'find out/ know', or a lexeme which translates as 'put on (clothes)' may actually mean 'put on/wear', that is, a change of state followed by a state or activity. What is important here is that this initial or final state or activity is implied (and not merely implicated) by the predicate. For example, while *die* refers to a process with a final point, and we are all aware of the resultant state, this state (i.e., 'being dead') is not an actual part of the predicate itself. The predicate refers only up to the end of the process, but not the state which results from this process. Once again, the usual test for this differentiation are adverbs such as 'still' Cf. **He has still died, *He still died*. There is of course the possibility of saying *He is still dead,* although it must be admitted that the adjective *dead,* while being closely related to *die*, is not a part of the paradigm of *die*. Thus, *die* in English can only refer to the change of state itself, which, as the following sentence shows, can be a drawn out change of state.

(1) He's dying.

Now confer the following Nepali sentence:

(2) *pānī (ajhai) uml-eko cha.*
water still boil-PERF.PART AUX.NPT.3.S.NF
'The water has boiled up. / The water is still boiling.'

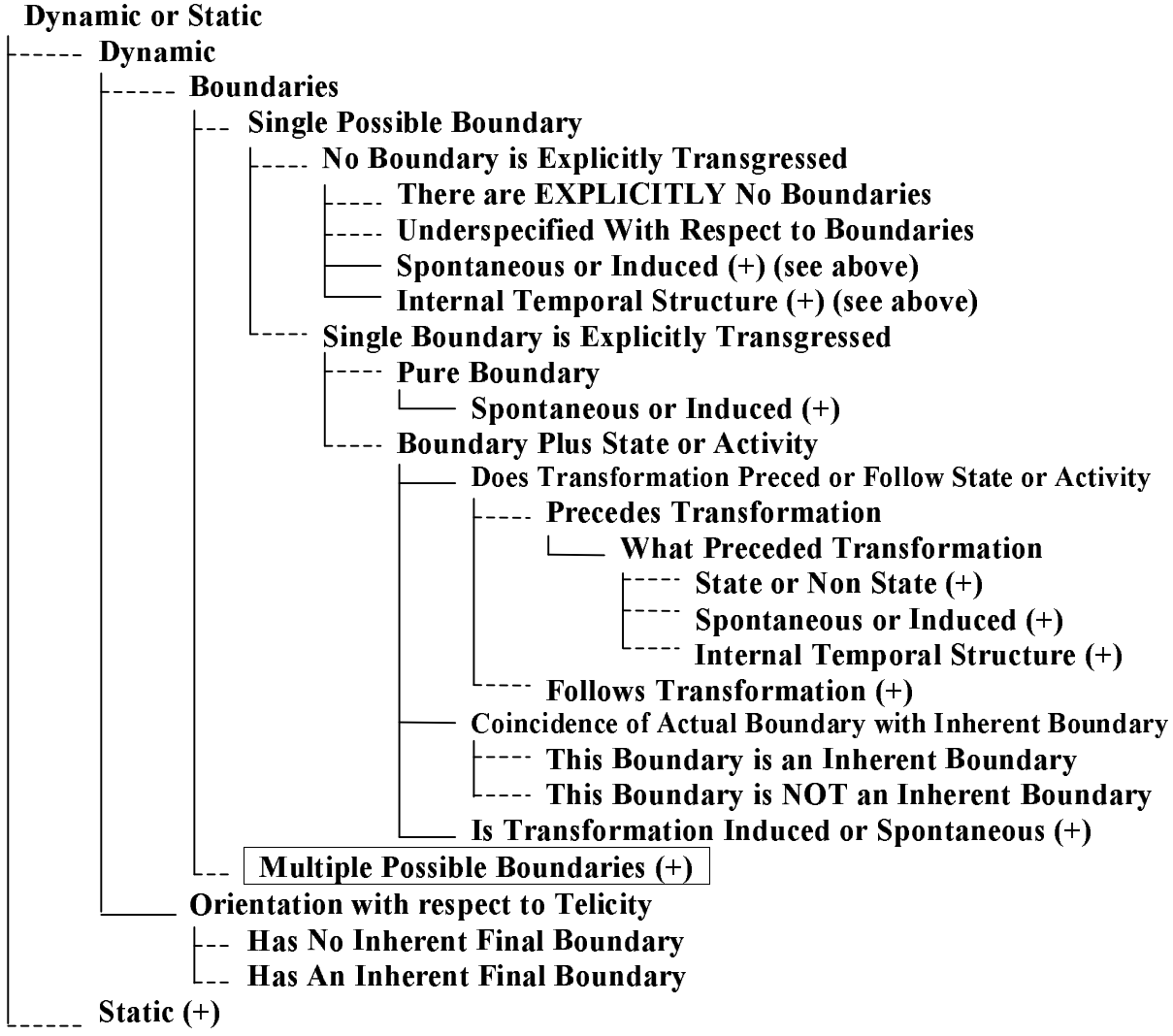
Here, the predicate refers possibly either to a change of state or to its possible resultant state. Both interpretations are not simultaneously present in either of the two interpretations, however. In the first, it is the water coming to a boil, in the second, the state which results from this. That is, the predicate *pānī uml-* in Nepali is underspecified with respect to which

interpretation is meant, and one must conclude that it involves both the change of state as well as the resultant state itself. In this fashion, the author also provides a good deal of data for the lexicon, i.e., that the predicate here has an underspecified inherent aspectual structure.

In order to describe the utterance fully, we must therefore enquire as to the number of possible boundaries, whether only a boundary is involved or a possible preceding or following state (as well as the intermediate stage, when there are two possible boundaries), whether the boundary or boundaries has or have actually been crossed, and whether these are inherent or arbitrary boundaries.

Finally, we need to know whether each phase involved is SPONTANEOUS or INDUCED, and whether a non-boundary phase is STATIC or DYNAMIC. There is finally the question of the internal temporal structure, which has already been dealt with above.

The following two diagrams present the DYNAMIC side of the "Aspect Tree" in detail. To facilitate matters for those who are only interested in the telicity of the predicate involved and who are not interested in a detailed discussion of the actual structure of the situation involved, the required node ORIENTATION WITH RESPET TO TELICITY has been added here.



Multiple Possible Boundaries

- **Pre Initial Boundary Phase**
  - **State or Non State (+)**
  - **Spontaneous or Induced (+)**
  - **Internal Temporal Structure (+)**
- **Actual Initial Boundary**
  - **Is This Boundary Explicitly Transgressed**
    - **Boundary is Explicitly Transgressed**
      - **Spontaneous or Induced (+)**
      - **Coincidence of Actual Boundary with Inherent Boundary (+)**
    - **Boundary is Not Explicitly Transgressed (+)**
- **Core Phase**
  - **State or Non State (+)**
  - **Spontaneous or Induced (+)**
  - **Internal Temporal Structure (+)**
- **Actual Final Boundary**
  - **Is This Boundary Explicitly Transgressed (+)**
- **Post Final Boundary Phase (+)** (same structure as Pre-Initial Boundary Phase)

We now turn to the FORM in which the aspectual and actional information is encoded, if it is encoded at all. This follows the same principles as does temporal marking and hence need not be discussed here in detail. The categories are taken almost without change from Bybee et al. (1994). Note that all nodes are of the optional type, so that they may be combined as necessary.

Aspect and Actionality Form

- **Aspect and/or Actionality is Overtly Marked**
  - **Grammatical Category of Aspectual or Actional Marking**
    - **Dynamic or Dynamically Neutral Categories**
      - **Atelic or Delimited**
        - **Means of Marking (+)** (also attached to all following categories)
      - **Continuous (+)**
      - **Excessive Duration (+)**
      - **Frequentive (+)**
      - **Habitual (+)**
      - **Imperfective (general) (+)**
      - **Inceptive (+)**
      - **Iterative (+)**
      - **Limited Duration (+)**
      - **Perfective (+)**
      - **Progressive (+)**
      - **Telic or Completive (+)**
    - **Stative Only Categories**
      - **State Changes (+)**
      - **State Continues (+)**
      - **State Ends (+)**
      - **State Exists (+)**
- **Aspect and/or Actionality is Not Overtly Marked**

5.2.4.3. Modality

The criteria for modality are taken almost without further modification from the criteria in Zaefferer 2001, a typology of modality in general. These criteria cover not only modality in its more traditional sense, such as deontic and epistemic modality, but also what Zaefferer refers to as "illocutionary modalities" ("illokutionäre Modalitäten"), including non-sentential illocutions (interjections, etc.) and sentential illocutions (expressive illocutions and volitive illocutions, the latter of which include prohibition, permission, etc., but also questions, assertions and declarations).

For further details on the criteria involved, the reader is referred to Zaefferer 2001. Here we will merely list the criteria and show how they are integrated into the "Modality Tree".

As an utterance may simultaneously express different types of modality, all major branches are of the optional type, so that these may be combined as needed. Only where these possibilities are mutually exclusive, such as with questions/declarations/assertions, will we make use of exclusive nodes.

Only minor changes have been made here, usually due to the difficulties involved when converting the information in any written text into the "tree" format used here.

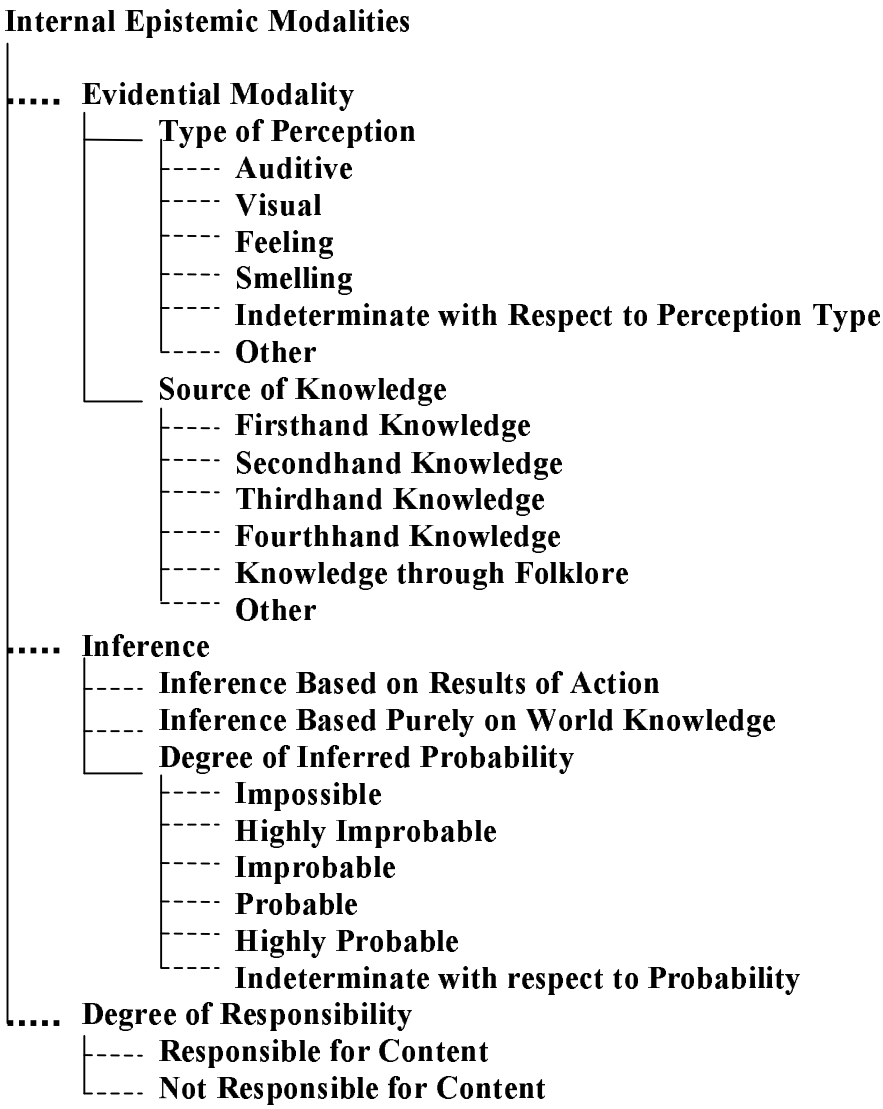
A minor change has also been made in the "Internal Epistemic Modalities", following the model laid out in Peterson (in press). Here, not only is a distinction made between "Type of Perception" and "Source of Knowledge", as in Zaefferer (2001), but a further differentiation has also been made between "Inference based on the Results of an Action" and "Inference Based Purely upon World Knowledge" (cf. also Willet 1988), as there are many languages where this difference is expressed grammatically. To achieve this, the model in Peterson (in press) has been used, where "inference" is not considered a type of "evidence" *per se* but rather is a closely related phenomenon.

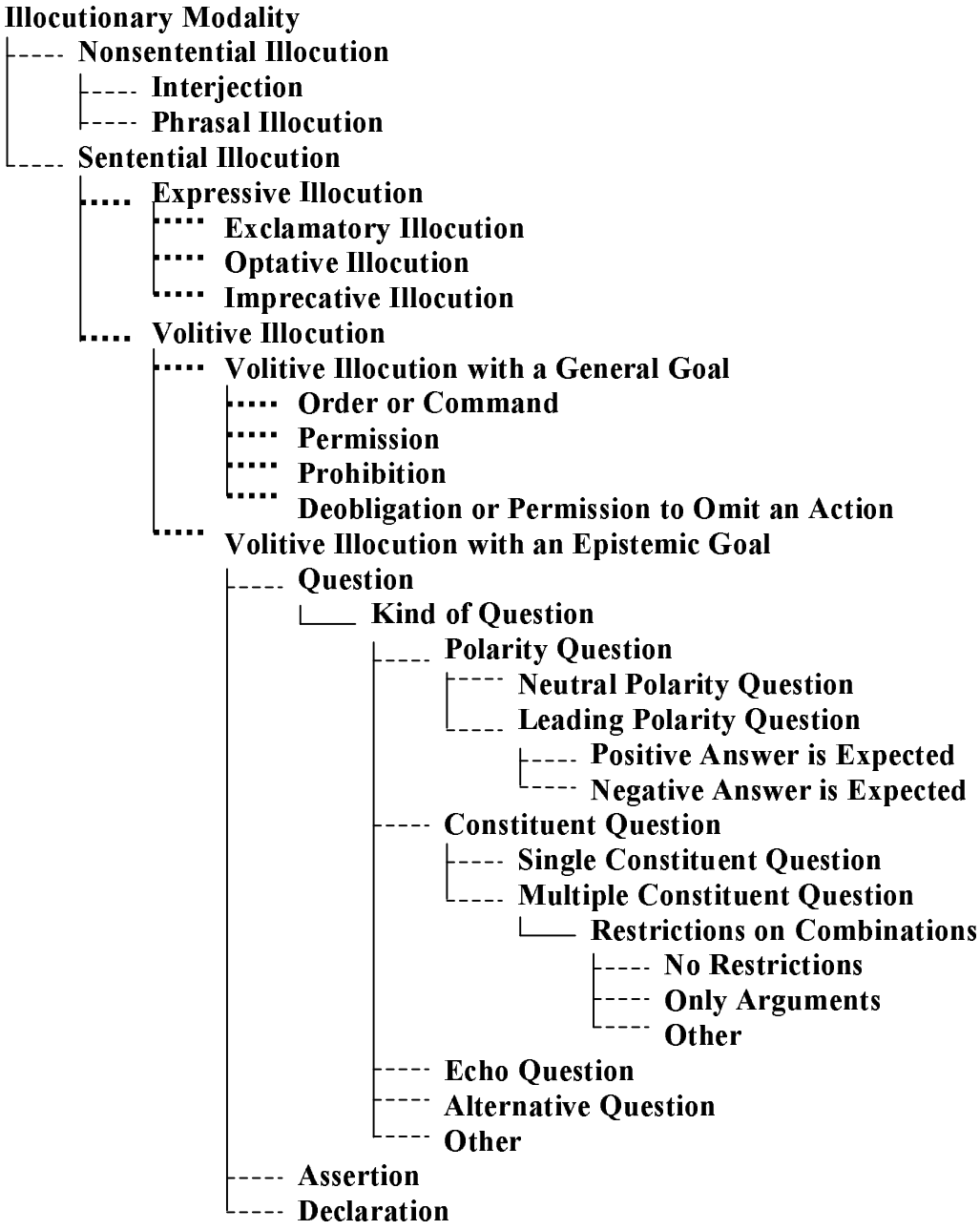
That the two are in fact separate phenomena, although they are clearly closely related, can be demonstrated as follows:

"Inference Based Purely on World Knowledge" implies no first- or second-hand knowledge of an event or situation, since the "source of knowledge" is merely the speaker's own past experience and what s/he infers from this past experience. Here, there is no direct contact to any outside situation. On the other hand, "Inference Based on the Results of an Action" implies first-hand knowledge of a situation. However, the reverse dependency does not hold: While the absence of first- or second-hand knowledge does automatically imply "Inference Based Purely on World Knowledge", first-hand knowledge by no means automatically implies "Inference Based on the Results of an Action" but may also mean that the speaker personally witnessed an action.

Hence, the two types of inference are listed here separately from the types of evidence.

- **Modality**
  - **Internal Modalities**
    - **Actional Internal Modalities**
      - **Tendency**
      - **Disposition**
        - **Acquired Ability**
        - **Nonacquired Ability**
      - **Influence**
      - **Circumstance**
      - **Andative**
      - **Venitive**
      - **Conative**
      - **Simulative**
      - **Deontic**
    - **General or Stative Internal Modalities**
      - **General Tendential Modalities**
      - **General Dispositional Modalities**
      - **General Influential Modalities**
      - **General Circumstantial Modalities**
      - **Imminence Modalities**
      - **Conditional Modalities**
      - **Similative Modalities**
      - **Alethic or Ontic Modalities**
    - **Internal Attitudinal Modalities**
      - **Internal Expressives**
        - **Internal Exclamatives**
        - **Internal Optatives**
      - **Internal Telic Modalities**
        - **Internal Volitive**
        - **Internal Intentional**
      - **Internal Epistemic Modalities (+)** (see the following pages)
      - **Internal Evaluative Modalities**
      - **Internal Expectational Modalities**
    - **Other**
  - **Illocutionary Modalities (+)** (see the following pages)


Internal Epistemic Modalities
Evidential Modality
Type of Perception
Auditive
Visual
Feeling
Smelling
Indeterminate with Respect to Perception Type
Other
Source of Knowledge
Firsthand Knowledge
Secondhand Knowledge
Thirdhand Knowledge
Fourthhand Knowledge
Knowledge through Folklore
Other
Inference
Inference Based on Results of Action
Inference Based Purely on World Knowledge
Degree of Inferred Probability
Impossible
Highly Improbable
Improbable
Probable
Highly Probable
Indeterminate with respect to Probability
Degree of Responsibility
Responsible for Content
Not Responsible for Content


Illocutionary Modality
Nonsentential Illocution
Interjection
Phrasal Illocution
Sentential Illocution
Expressive Illocution
Exclamatory Illocution
Optative Illocution
Imprecative Illocution
Volitive Illocution
Volitive Illocution with a General Goal
Order or Command
Permission
Prohibition
Deobligation or Permission to Omit an Action
Volitive Illocution with an Epistemic Goal
Question
Kind of Question
Polarity Question
Neutral Polarity Question
Leading Polarity Question
Positive Answer is Expected
Negative Answer is Expected
Constituent Question
Single Constituent Question
Multiple Constituent Question
Restrictions on Combinations
No Restrictions
Only Arguments
Other
Echo Question
Alternative Question
Other
Assertion
Declaration

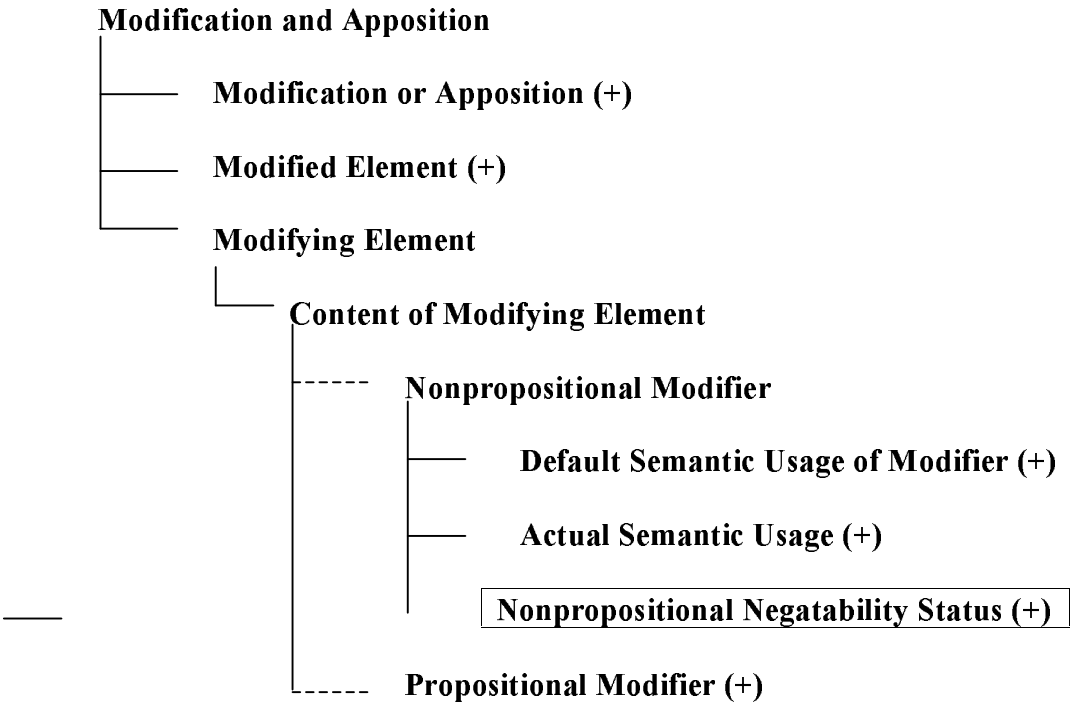
5.2.4.4. Negation

The criteria for the treatment of negation are taken primarily from Payne (1985). As with other areas of the grammar, negation involves two major sections to be dealt with: content (what is negated, type of negation) and form. Finally, we would also like to know as much as possible about the possible restrictions of certain types of negation in different languages. Let us begin with a description of negational content.

The first thing we need to know about negation is of course what type of negation is involved: CONTRADICTORY (or "mutually exclusive") negation, such as *smoker / non-smoker*, or CONTRARY negation, where the two terms "represent opposite poles along a given dimension and leave room for other possibilities between them, as for instance with *intelligent* and *unintelligent*." (Payne, 1985:241).

The next criterion on the content side is whether we are dealing with PROPOSITIONAL or NONPROPOSITIONAL negation. Propositional negation involves, as its name implies, the negation of all types of propositional information, generally negation of sentences (assertions, commands, interrogatives) and clauses. Nonpropositional negation involves the negation of nonpropositional modifiers and lexical negation in general (negative adjectives, adverbs, nouns, etc.). As the two areas deal with completely different grammatical domains, it is not possible to have a single "Negation Tree" in the present format. Instead, propositional negation will be offered as a "tree" when discussing predication, clauses in general, propositional modification, propositional complementation, etc., while nonpropositional negation will be dealt with in the appropriate contexts (e.g. the lexicon, nonpropositional modification, non-propositional complementation, etc.). A search function will help the user locate all instances of negation in the grammar in its present state. For the future, the use of hyperlinks is planned, so that it will be possible to have a virtual link of all instances of negation to a superordinate "Negation Tree".

Regardless of whether we are dealing with propositional or non-propositional negation, we need to know the scope of negation. In the case of non-propositional negation, this question need not be stated overtly, since the domain has already been given, only AFTER WHICH is the question of polarity dealt with. An example from modification:



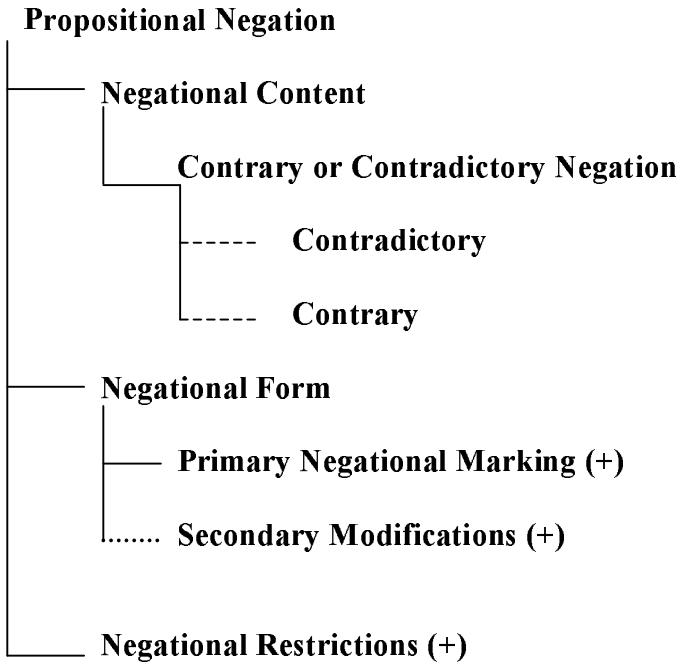
If the author has arrived at "Nonpropositional Negatability Status", then it is obvious that s/he is clearly dealing with the (non)negatability of a nonpropositional modifier, which is all that can be within the scope of negation here.

The same holds true for propositional negation, as information on propositional negation is either at the level of the description of a proposition or of a clause (complement or modifier). Thus, there is no room here for ambiguity as to the scope of negation.

Ideally, it would also be desirable to have a hyperlink connecting presentation structure (section 5.2.6.) with negation, as the two are intricately interwoven in terms of the the scope of negation. This is planned for a future version of the grammar but is not possible at present for technical reasons. Nevertheless, the user of the grammar will of course be able to click on both "Propositional Negation" and "Presentation Structure" and have both windows open at the same time, allowing for quick and easy comparison.

We turn now to the FORM of negation. This consists of a PRIMARY form of negative marking and is often accompanied by SECONDARY MODIFICATIONS, which are however not found in all languages. Hence, while primary negation is a required node, secondary modifications will be presented as an optional node.

Under primary forms we mean, following Payne (1985:207), "the *addition* of a negative morpheme to a corresponding positive sentence." This is in opposition to the secondary modifications, which never seem to indicate negation alone but require the presence of a primary form of modification. We will deviate here somewhat from Payne and include "Phonological Alternations" under the primary means of negation, as there is some evidence, although not completely conclusive, that tone changes alone can indicate negation (cf. Dahl, 1979:82). The uppermost nodes of propositional negation thus have the following form:



Primary Negational Marking

First, we must take note of the fact that while in some languages propositional negation is marked only once, in others it is marked twice or more. Also, in languages which generally mark negation only once, negative words or constructions can often be found more than once in the sentence, where each negative word (etc.) has its own negating force. Cf. the following examples:

Standard English:

Single Negational Marker:

(1) I don't know anything.

Multiple Negational Markers, each with own negating force:

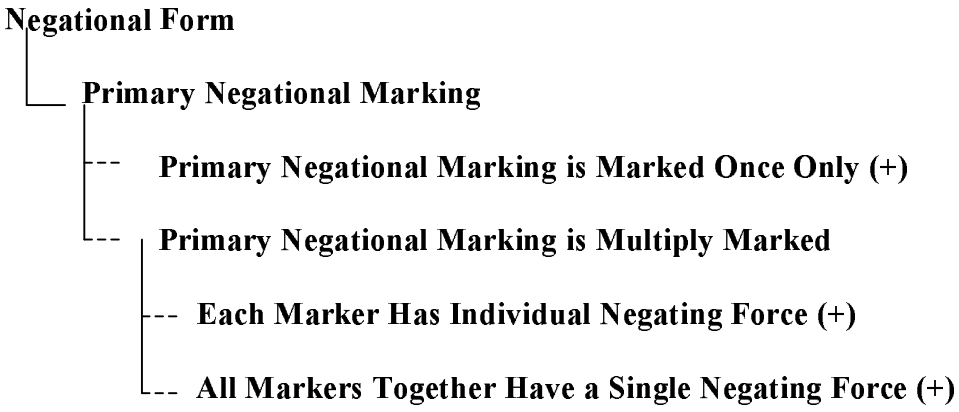
(2) I didn't take nothing (i.e., I did take something).

Spanish

Multiple Negational Markers with single negating force:

(3) No sab-e nada.
NEG know-3.S.PRS nothing
'S/he doesn't know anything.'

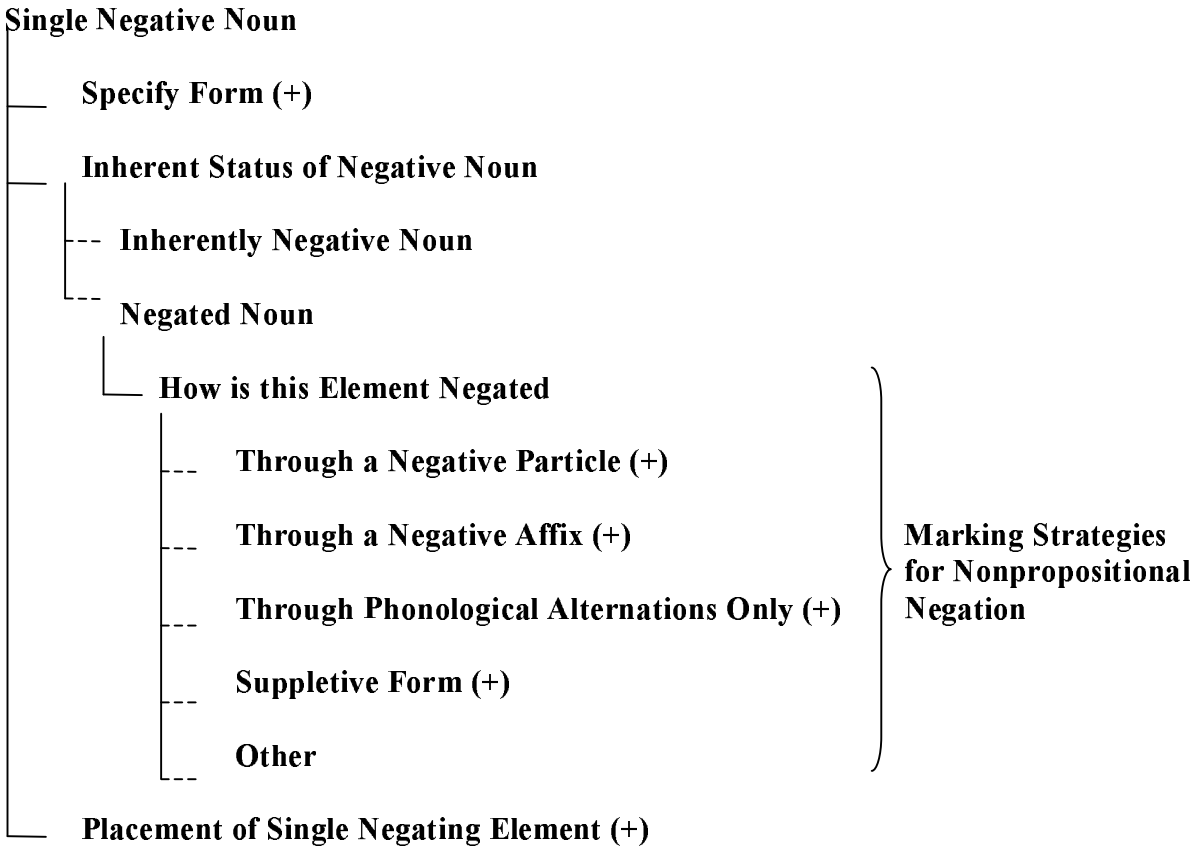
Hence, we must first ask the author to indicate whether negation is marked only once or more than once, and if it is marked more than once, whether each occurrence of negation has its own negating force. Thus, the pertinent part of the "tree" looks as follows:



Payne (1985) dstinguishes between three types of "sentential negation": "standard negation", i.e., "that type of negation that can apply to the most minimal and basic sentences" (Payne, 1985:198), NEGATIVE QUANTIFIERS (with the two classes NEGATED and INHERENTLY NEGATIVE QUANTIFIERS) and NEGATIVE ADVERBS (again with NEGATED and INHERENTLY NEGATIVE subgroups).

"Standard negation" can be divided up among the following types of marking: NEGATIVE VERBS, NEGATIVE PARTICLES, MORPHOLOGICAL NEGATIVES and the comparatively rare NEGATIVE NOUNS. We add to this list "Phonological Alternations", which are said to be the primary means of negation in at least one language (cf. Dahl, 1979:32). We will not deal with this topic here in any detail, however, as it is dealt with elsewhere (cf. section 6.).

Concerning negative quantifiers, adverbs and nouns, we first need only enquire as to whether sentential negation is achieved by a "negative" adverb, quantifier or noun. Next, we can enquire as to the inherent status of this negative form, i.e. either inherently negative (e.g. *scarcely*, *barely*) or negated (e.g. *not often*, *nowhere*). If it is negated, there is of course the question as to how this is done. Here, we can simply insert the remaining possibilities for nonpropositional negation, which we present below in abbreviated form. These same possibilities hold for nonprositioanl negation in general and this "tree" may thus be inserted in all parts of the grammar where nonpropositional negation is found. The example given here is from the alternative "Negative Noun":



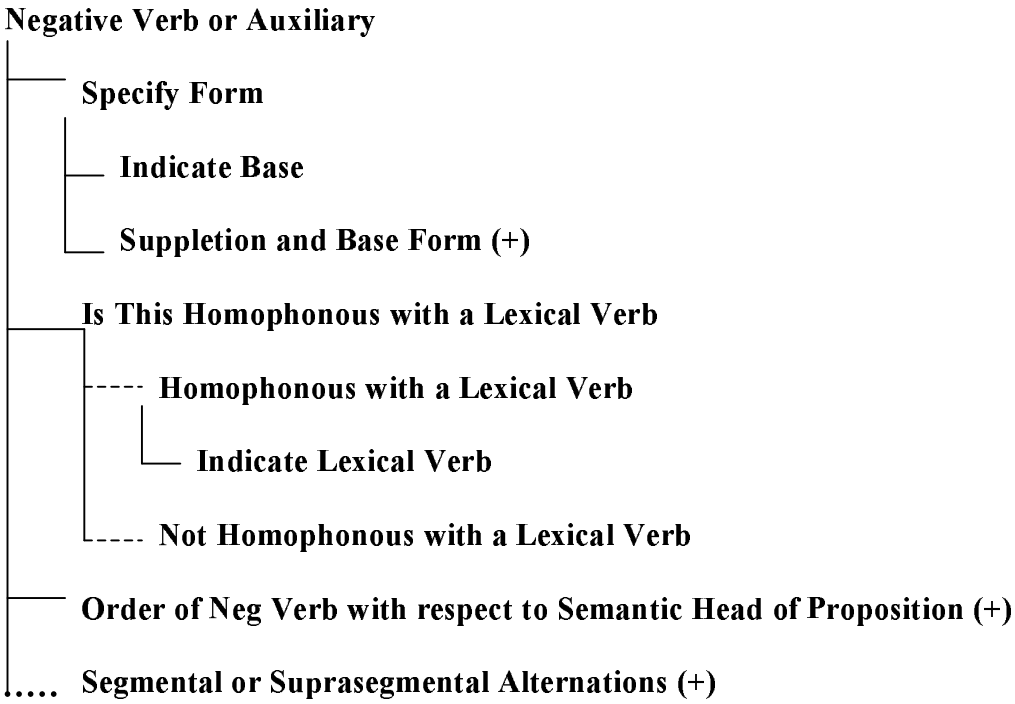
(For reasons of space, it is not possible to present the "tree" here in its entirety. However, the general stragegy should be clear by this point.)

Turning now to the remaining marking strategies for propositional negation, we can begin with the NEGATIVE VERBS, found e.g. in many Polynesian languages. Included here are also NEGATIVE AUXILIARIES, such as for example English.

Payne (1985:207ff) differentiates between "higher" or lexical negative verbs and auxiliary negative verbs. Also, among the "higher" negative verbs, there are those which have either all or only some of the characteristics of a typical main verb, such as the ability to take a sentential complement, while others are more restricted, often to the point that it is difficult to decide if we are dealing with a "higher" or auxiliary verb (cf. Payne's discussion of Tongan and Fijian, 208ff).

We can, however, skip over much of this information here, as the status of the predicate is dealt with in detail elsewhere in the tree (section 5.2.1.), where the form of the predicate is described in detail, including questions such as finiteness, form, complementation, etc. Here, we need only enquire as to whether the negative verb is homophonous with a lexical verb or not, and if it is, with which lexical verb. If we are dealing with an auxiliary verb, or even with two conjugated verbs, both of which carry to certain degrees person marking, TAM marking, etc., this information will all be entered under the node "Predication" (section 5.2.1.).

Of course, as with all forms of negation, we also need to enquire as to the placement of this negative word. However, as this topic is of a mere "mechanical" nature, we will not deal with it here in any detail. Thus, the "tree" for negative verbs is as follows:



Similar strategies are also used to describe negative particles (i.e., invariant negative forms), morphological negatives (while we do not restrict these to derivational forms only, as does Payne, but include inflectional forms as well), and negative nouns. In each case, we must also enquire as to the placement of these markers, and in the case of morphological negatives, we must also take into consideration the fact that the negative marker may occur verb internally as well as before or after other words. Cf. Nepali:

(4) u ga-e-na-thyo
3.S go-PFV-NEG-PT.3.S.NF
'He didn't go'

Cases such as these are easily dealt with by using a simple analogy to the description of placement for affixes, which can have the form of prefixes, suffixes, infixes, etc., the details of which are not of concern here, as they are of a more "mechanical" nature.

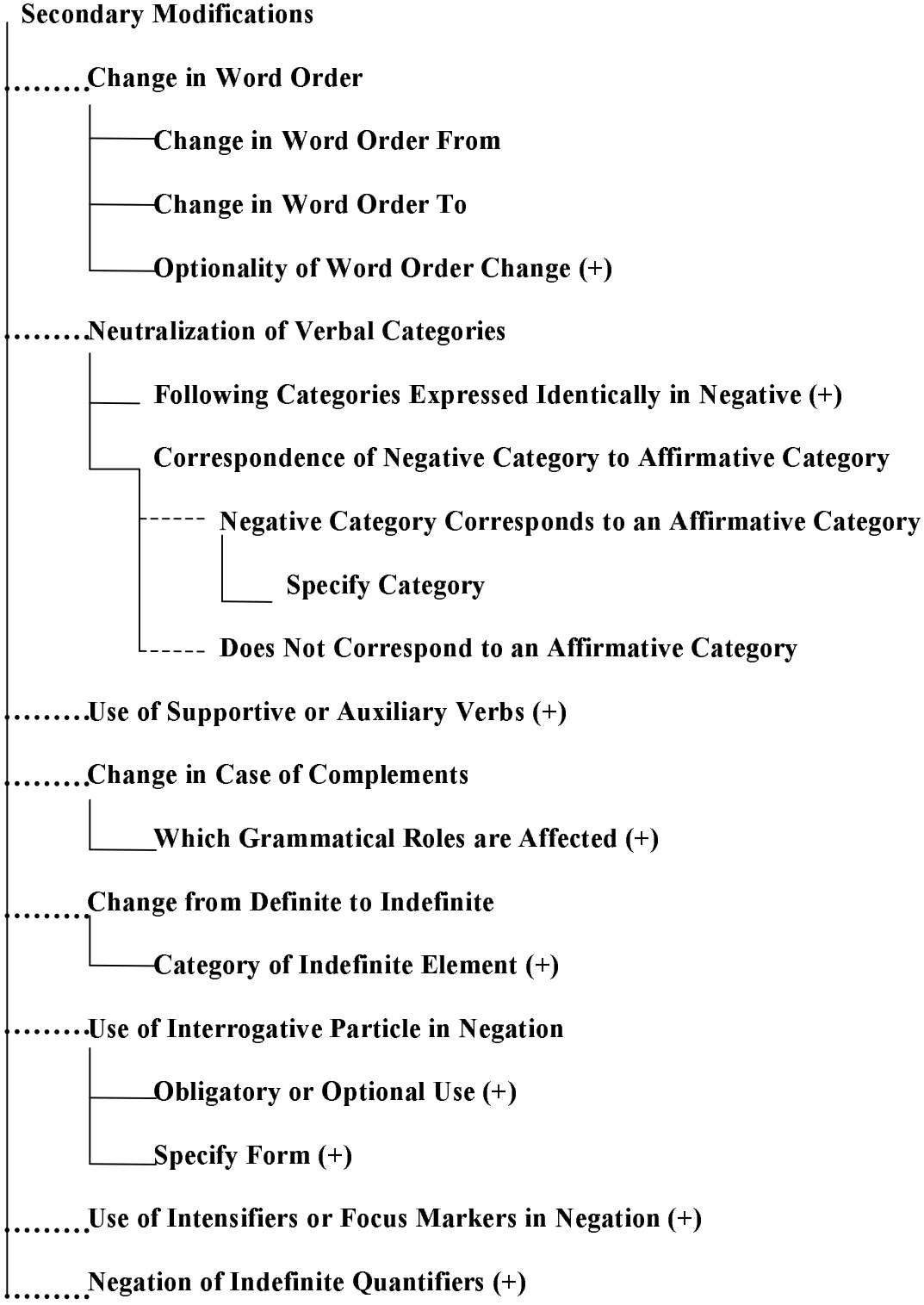
Secondary Modifications

As noted above, not all languages show secondary modifications in addition to the primary means of negational marking, hence "Secondary Modifications" are joined to "Negational Form" as an optional node - it may be clicked on if necessary, but is not required.

There are a number of secondary modifications mentioned in Payne, including CHANGE IN WORD ORDER, CHANGE IN TONE, NEUTRALIZATION OF TENSE DISTINCTIONS, USE OF SUPPORTING VERBS (other than negating auxiliaries) and CHANGE IN NOUN CASE. The category "Change in Tone" is already included in the tree under primary marking as the optional node "Segmental or Suprasegmental Alternations" and will not be dealt with here (cf. section 6.). To Payne's criteria we have further added CHANGE FROM DEFINITE TO INDEFINITE, USE OF INTERROGATIVE PARTICLE IN NEGATION, USE OF INTENSIFIERS OR FOCUS MARKERS IN NEGATION and NEGATION OF INDEFINITE QUANTIFIERS, as these were encountered in a number of the grammars consulted. We have also expanded Payne's category of neutralization of tense distinctions to

include all verbal categories, as we have also found instances, such as Nepali, where aspectual rather than temporal distinctions are neutralized in negation.

As any possible combination of these secondary modifications may be found, all have been added in the form of optional nodes. By now the reader will be familiar enough with the general strategy involved, so that we need not go into the details of this part of the grammar but simply list the uppermost nodes.



Negational Restrictions

We now come to the last topic to be dealt with under the heading "Negation" - that of possible restrictions on types of negative markers, restrictions on the negation of constituents of various pragmatic statuses, restrictions on the form of negative markers in main vs. subordinate clauses, etc. The main purpose of this is to allow the author the chance to include much of the information which is generally included in grammars in prose form. However, while we believe we have been able here to anticipate many of the restrictions found in various languages, this section of the present grammar cannot be considered exhaustive and cannot fully replace a thorough prose description. To this end, the node "Other Negational Restrictions" has been added and the author will be given ample opportunity to include any pertinent information in the form of prose comments.

The following restrictions have been included in the grammar:

- Pragmatic Restrictions on Presence of Arguments in Negation
- Restrictions on the Use of a Determiner
- Retrictions on the Negation of Adverbials
- Restrictions on the Negation of Quantifiers
- Restrictions on the Form of the Negative marker. This includes the following restrictions:
  - Restrictions according to Word Class
  - Restrictions according to TAM categories
  - Restrictions according to the status of the clause (i.e., main vs. subordinate)
- Other Negational Restrictions

For the sake of brevity, the shortest restrictional "tree" is presented here:

Restrictions according to the status of the clause

This marker is Restricted to the following clause types

Independent clauses

Dependent clauses

Dependent clauses in general

Restricted to the following types of dependent clauses

5.2.5. Marking

Marking is dealt with in the most liberal fashion possible in the grammar. It is a general "tree" which is inserted into all parts of the grammar where overt marking can be expected. This includes all referential and predicating expressions as well as individual words (e.g. determiners, quantifiers, adjectives, adverbs, nouns, verbs, auxiliaries, adpositions, etc.).

The "Marking Tree" consists of two major parts: the grammatical categories which are marked, and the strategies which have been found in languages to mark any one of these categories. The grammar thus avoids the restrictions of anticipating only certain marking on certain words or phrases, and offers each means of marking for each grammatical category which is overtly marked. Thus, while e.g. person marking is a TYPICAL feature of verbs, there is no reason why it should be excluded as a possible feature for prepositions (which IS found in a number of languages) or adverbs. Also, while we know of no language where number marking is achieved through tone, there is no reason to exclude such a possibility for language in general.

To help both the authors and the consultants of the grammar orient themselves, we have grouped the various categories as follows: "Marking Which is Typical of Nominal Phrases", "Marking Which is Typical of Predicates", "Miscellaneous Marking Categories" and "Other Types of Marking". This, however, is only to avoid having an extremely long list of categories - altogether there are 24 major categories, not counting all the differet KINDS of categories which each group contains. For example, "Gender" is given as a single category - WHICH gender (masculine, feminine, neuter, non-masculine, non-feminine, etc.) is dealt with in detail under the main tree itself. It should be noted that this grouping into "typical" categories for nominal phrases, etc., has no theoretical implications but is merely intended as a practical aid: ALL NODES are listed here as optional, so that ALL may be selected in each case, and of course any combination of these. The categories are listed in alphabetical order in each subdivision. Here the individual categories:

Overt Marking

Marking Which is Typical of Nominal Phrases

Alienability or Inalienability Marking (+) (perhaps in conjunction with general possessiional marking)

Classifier Marking (+)

Grammatical Relation Marking (+). Note: this is not identical with "case" marking, as case marking, while having "Grammatical Relation Marking" as its base, is found elsewhere in grammar, e.g. in conjunction with adpositions, often in comparative/superlative constructions, etc. In addition, "Grammatical Relation Marking" can also be realized by adpositions, which we would not necessarily wish to refer to as "Case Marking". Further down in the "tree", under "Means of Marking", which will be dealt with shortly, there will be a required node which asks whether this marking is to be considered "case" marking and if so, which case. See below.

Possessional Marking (+)

Marking Which is Typical of predicates

Diathesis and Valency Marking (+) (e.g. ACTIVE, PASSIVE, APPLICATIVE, CAUSATIVE, TRANSITIVE, INTRANSITIVE, etc.)

Inverse Marking (+)

Person Marking (+)
Predicate Nominal or Adjectival Marking (+)
Relative Form Marking (+) (e.g. the so-called "relative" verb form found in some languages)
TAM Marking (+)

Miscellaneous Marking Categories
Additive or Restrictive (+)
Altitudinal Marking (+)
Comparative, Equative or Superlative Construction Marking (+)
Coreferentiality or Noncoreferentiality (+)
Distance Marking (+)
Distributive Marking (+)
Empathy or Subjective marking (+) (e.g. ENDEARMENT, DISLIKE, DIMINUTIVE, AUGMENTIVE)
Gender or Class Marking (+)
Honorific Status (+)
Number Marking (+)
Pragmatic Marking (+)
Word Class Marking (+) (e.g., NOMINALIZER, ATTRIBUTIVIZER, etc.)

Other Types of Marking - Specify

Once an individual category has been selected as being overtly marked, there arises the question as to HOW this category is marked. Hence, each category has the required node "Means of Marking", which is identical in each case and offers all possibilities known to us for the marking of any grammatical category. These are:

Marking is Achieved through Additional Separate Word
Marking is Achieved by Clitic
Inflectional and Derivational Marking
Marking is Achieved by Word Order
Marking through Sentence Level Suprasegmentals
Zero Marking or Lack of Other marking
Segmental or Suprasegmental Alternations

In the following, a brief overview of each category is given:

- **Marking is Achieved through Additional Separate Word**
  - **Modifier**
  - **Conjunction**
  - **Adposition**
    - **Preposition**
    - **Postposition**
    - **Can Function Both as Preposition and Postposition (+)**
  - **Verbal Form or Form Derived from a Lexical Verb**
    - **Finite Verb Forms (+)**
    - **Nonfinite Form (+)**
  - **Use of Determiner**
    - **Status of Determiner (+)**
  - **Use of Quantifier**
  - **Specify Form (+)**
- **Marking is Achieved by Clitic (+)**
- **Inflectional and Derivational Marking**
  - **Affix**
    - **Suffix**
    - **Prefix**
    - **Circumfix**
    - **Infix**
    - **Transfix**
  - **Reduplication**
    - **Reduplication of Entire Word**
    - **Reduplication of Part of Word (+)**
- **Marking is Achieved by Word Order (+)**

Segmental and suprasegmental alternations must also take into account the different nature of phonological inventories of spoken and signed languages. This is achieved by first requesting the author to decide (required node) whether the language is signed or spoken, and then proceding accordingly. However, as we had no data at hand for signed languages dealing with this area of the grammar, we have as yet not been able to provide criteria for signed languages here. As research progresses in this field of sign-language studies, this node can easily be added.

Segmental or Suprasegmental Alternations

- **Segmental Alternations and Processes**
  - **Spoken Language Segmental Alternations and Processes (+)**
  - **Signed Language Segmental Alternations and Processes (under construction)**
- **Suprasegmental Alternations and Processes**
  - **Spoken Language Suprasegmental Alternations and Processes (+)**
  - **Signed Language Suprasegmental Alternations and Processes (under construction)**

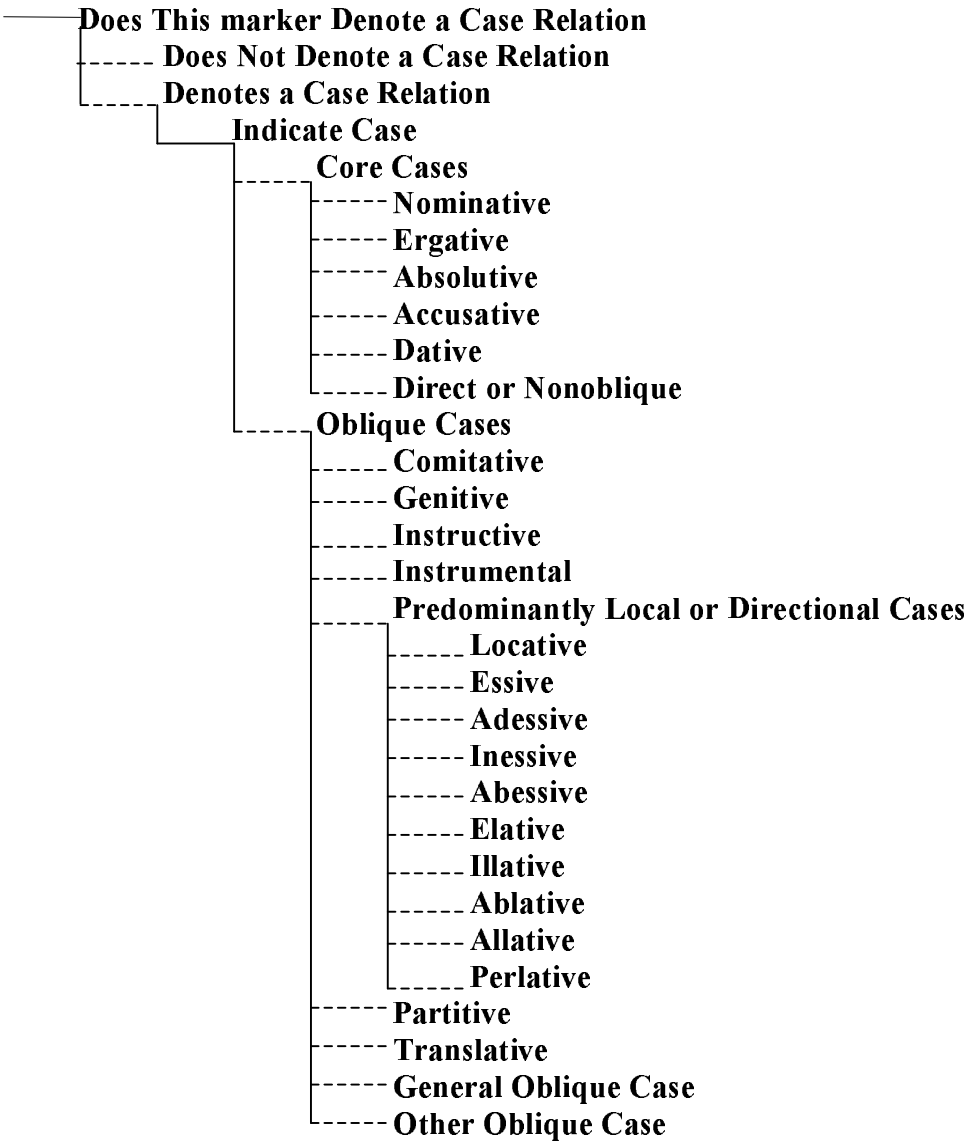
The same of course also applies to the following, which need not be spelled out here in detail:

Marking through Sentence Level Suprasegmentals (+)

Finally, note that there is also the possibility for "zero overt marking". This allows the author to make use of "zero morphs" where his or her analogy requires:

Zero Marking or Lack of Other marking

After the means of marking has been entered, the following question appears (required node): "Does This Marker Denote a Case Relation". That is, "case" is of course not a MEANS OF MARKING, and while it receives an interpretation as "case" primarily from its use to denote the grammatical role of a complement of the predicate (as well as its paradigmatic properties), it will often (if not always) have other uses as well, such as to denote length of time, distance, or its use in comparative and superlative constructions, to name just a few. Hence, this question allows the consultant of the grammar to determine (for example, with a search command) in what contexts e.g. the accusative in a certain language is used, or what form case markers take in a certain language. The cases are divided into "core" and "oblique" cases. The following list is intended to be exhaustive, but if the need arises, it may of course be expanded:


Does This marker Denote a Case Relation
Does Not Denote a Case Relation
Denotes a Case Relation
Indicate Case
Core Cases
Nominative
Ergative
Absolutive
Accusative
Dative
Direct or Nonoblique
Oblique Cases
Comitative
Genitive
Instructive
Instrumental
Predominantly Local or Directional Cases
Locative
Essive
Adessive
Inessive
Abessive
Elative
Illative
Ablative
Allative
Perlative
Partitive
Translative
General Oblique Case
Other Oblique Case

5.2.6. Presentation Structure

Our discussion of "Presentation Structure" is based primarily on the criteria in Lambrecht (1994) and Gundel et al. (1993), although we have supplemented these by a number of other criteria.

We first distinguish between the "Presentational Structure of a Proposition or Clause" and the presentational structure of a complement or adjunct, what we will term here "Pragmatic Role of a Sentence Constituent".

For the presentational structure of a proposition or clause, we then distinguish between the content and the form of this presentational structure. That is, what is focused and how is this marked?

We begin on the content side, where any one of the following may be focused:

- VP or Unmarked focus. This is what is generally referred to as "Predicate Focus" or "Comment" in a topic-comment construction.

- Focus of a Constituent or a Part Thereof. That is, an entire constituent can be focused, such as an entire NP, or just a part of this NP - the semantic head, a modifier, a determiner or a quantifier. The same criteria also hold for a focused predicate - is the semantic head focused, a light or auxiliary verb, etc.?

- Sentence or Clause Focus. We further distinguish here between Thetic Clauses and Presentational Clauses.

The next criterion which must be specified is the foreground or background information. Does the proposition/clause present information which is considered foregrounded (i.e., advances the plot of a story, etc.) or backgrounded? This information is important as the two are often distinguished in languages by the use of different verbal categories to present information which merely serves as background information as opposed to information which is considered foregrounded.

Finally, we would like to know the givenness of the information - is it presented for the first time (i.e., new)? Or is this information which has already been given at a previous time? If it is not completely new information, when was this information given for the first time - in the previous utterance, two utterances earlier, or three or more? Or is the givenness status of this information indeterminate as to when it was given for the first time?

After the content has been specified, we can then turn to the form or "Means of Marking", which was discussed in the previous section and need not be repeated here.

Here, a schematic diagram of the uppermost nodes of the presentational structure of a clause or proposition:

Presentation Structure of a Proposition or Clause

- **Content of Focus**
  - **VP or Unmarked Focus**
  - **Focus of a Constituent or Its Part**
    - **Focus of an NP or a Part Thereof (+)**
    - **Focus of a Predicate or Part Thereof (+)**
    - **Focus of a Sentential Modifier or Part Thereof (+)**
  - **Sentence or Clause Focus**
    - **Thetic Clause**
    - **Presentational Clause**
- **Means of Marking (+)**

We now turn to the treatment of the pragmatic role of a sentence constituent. Here, we follow the criteria in Gundel et al. (1993) with only slight modifications. A sentence constituent is considered to be placed without ambiguity onto a scale of seven pragmatic statuses. These are discrete levels (i.e., not "fuzzy"). While the criteria in Gundel et al. (1993) are intended as a hierarchy, where the placement of a sentence constituent on this scale autmatically implies that all criteria below this point will also be fulfilled, we do not require this. The main reason for this is that we have added the "Generic Class" to the scale, which Gundel et al. (1993) explicitly leave out (Gundel et al., 1993: 83, n. 14). We have also expanded the "activated" node considerably. For details on the exact status of the various levels, cf. Gundel et al., 1993.

The following diagram presents these criteria in "tree" format.

Pragmatic Role of a Sentence Constituent

- **Under Discussion for at Least Two Consecutive Utterances**
- **Activated**
  - **Activation Status**
    - **Activated by Current Speaker**
    - **Activated by Other Speaker**
  - **Topic Flow Status in Previous Utterance**
    - **Contrastive Topic or Corrective Contrast**
    - **Non Contrastive Topic**
      - **Former Topic is Taken Up Again**
      - **Discontinous or New Topic**
      - **One of a Number of Possible Topics (i.e., 3 or more)**
- **Familiar**
- **Uniquely Identifiable**
- **Referential**
- **Generic Class**
- **Type Identifiable**

6. The Lexicon

6.1. Problems with Traditional Lexica

The lexicon in the traditional sense plays a central role in the description of any language. Generally, this is merely a list of lexical entries where each entry consists of a citational form and an attempt to explain its most basic meaning(s) in simple prose. In the case of a bilingual dictionary, the main purpose is of course normally to give as close a general translation of the foreign-language entry as possible.

There are a number of problems when using this kind of lexicon in cross-linguistic research. First, depending on the language and the word class, this entry may be provided in any one of a number of different forms. For example, a verb can be listed in the lexicon as an infinitive (e.g. Spanish), a hypothetical verbal root (e.g. Sanskrit), a participle (e.g. Bengali), the first or third person singular of the present tense (e.g. Greek or Pali, respectively) or only through its characteristic consonants (Arabic, Hebrew). Bringing this information into a unified standard form is then left to the linguist him-/herself.

This problem can be avoided in the CRG on the one hand by using the English-language gloss as a guide, while explicitly requesting the generally accepted citation form on the other, so that both forms are available upon request.

Also, much information which would typically be of interest to the linguist is generally not presented in such a lexicon. For example, the entry for the English verb 'eat' in a popular dictionary (Fowler & Fowler, 1958:378) begins as follows:

"v.t. & i. (past *ate, eat,* pron[ounced] e$_{t}$: p.p. *eaten*, pron[ounced] ētn). Masticate & swallow (solid food); swallow (soup); ..."

Although this definition can certainly be termed "accurate", it omits much potentially important information, such as whether or not the use of this verb is restricted to agentive contexts only (as it is - by definition - restricted to animate subjects), whether it refers to the act of "mastication and swallowing" only of humans or also of animals, etc. It also tells us nothing about its use in the passive and in other verbal categories or about its inherent aspectual meaning or *Aktionsart*.

While this information may be considered trivial for the user of this dictionary, who by necessity will be familiar with English, the problem is much more acute when dealing with a language with which the user is not at all familiar. Here, the user is likely to be interested in a large number of traits of this particular lexeme, such as the following:

- Is this lexeme restricted to agentive contexts?

- Can other forms be derived from it (non-causative or detransitive (i.e., "unaccusative") forms, causative forms, etc.)?

- How are these valency changes expressed on the verb itself (inflection, derivation, no changes (i.e., zero derivation), etc.)?

- Is this verbal lexeme defective or can it occur in all verbal categories which the language possesses?

- Is there any suppletion in the paradigm of this verbal lexeme?

- How many complements does it take in the different verbal categories?

- How are these marked and what is the possible range of their semantic roles (agent, instrument, experiencer, patient, etc.)?

- If one of the complements is not known, what effects does this have?

- What is the inherent lexical aspect (or *Aktionsart*) of the stem?

- How is the morphosyntactic marking of the verb and its complements related to the pragmatic status of these complements?

- Does this verb ever occur with an expletive?

Needless to say, depending on the user's special interests, this list of *desiderata* could be continued indefinitely.

6.2. The Lexicon in the Cross-linguistic Reference Grammar (CRG)

In the CRG, these issues are of the utmost concern, as the project is dedicated to the description of all human languages. Thus, the traditional type of lexicon will not be sufficient for such an undertaking and must be expanded to take these types of questions into account.

At the same time, the fact that the grammar is encoded in electronic form, and that each lexeme will be used in a number of examples, each of which contains an exhaustive morphosyntactic and phonological description, means that all relevant information is both accessible and can be presented in a structured form that goes well beyond the means at the disposal of a traditional dictionary.

6.2.1. The Structure of the CRG Lexicon

The lexicon as such will be replaced by a number of "Lexica" or "Inventories", each of which is dedicated to one or more aspects of an individual entry. The data can be entered separately into the appropriate lexicon or can be abstracted to a large extent from the examples themselves and the accompanying grammatical description.

Let us begin by showing how this information can be partially extracted from the examples themselves. To begin with, each entry will typically contain a phonological entry (Level +2), a morphemic reconstruction (Level +1), the gloss (Level -1), where necessary a representation of complex morphology (Level -2), and a level for representing the various sentence consitituents (Level –3). The information in these levels are all entered manually by the author.

Furthermore, there are two automatically generated levels, the so-called "Templates" (Level -1', Level -2'), which present the abstract categories that the categories in Levels -1 and –2 belong to.

As an example, consider the following condensed English entry:

Level +2 Phonological transcription	aɪ̯	æm	ˈiː.tɪŋ	
Level +1 Morphemic reconstruction	aɪ̯	æm	iː.t	–ɪŋ
Level –1 Gloss	PRN.1.S.NOM	AUX.NPT.1.S	eat:V	-IPFV.PTCP
Level –2 Complex Morphology		PROG-eat-NPT.1.S		
Level –3 Sentence Constituents	NP(SUBJ)	V-FIN	V-INF	
Level –5 English translation	'I am eating.'			

With the automatically generated levels:

Level -1'	PRN.PERS.NUM.CAS	AUX.TNS.PERS.NUM	LEX:V	-ASP.WCU
Level -2'		ASP-LEX-TNS.PERS.NUM		

where V-FIN = 'finite verb', V-INF = 'nonfinite verb' and WCD = 'word-class or derivational marker'.

Even this brief entry - without the full grammatical description which is attached to it (not visible here) - presents a large amount of information for the user:

- What is realized in Level +2 as [æm ˈiː.tɪŋ] consists of an auxiliary plus a participial form of the lexeme /iːt/.

- The lexeme in question can easily be recognized by the computer as a lexeme, as it is written in small-case letters in Level –1 (the gloss), while grammatical information is encoded in small capitals.

- The fact that the lexeme is followed by ":V" in the gloss denotes that the hypothetical author has chosen to consider /iːt/ to be a basic lexical verb.

- The example shows that this lexeme can occur in environments in which it takes only one complement.

- The single complement in this example appears in the nominative.

- The single complement in this example is the subject of the clause.

- The constituent order is NP(SUBJ) – V-FIN – V-INF.

- The verbal lexeme can be used in the progressive.

- The progressive in English is formed by an auxiliary which precedes the lexical verb and by the suffix / –ŋ/, which is attached to this verbal lexeme.

To take all this information into account, the traditional notion of a lexicon has been supplemented in the CRG-Project by a number of "lexica" or "inventories". These will record all information which is encoded either in the examples themselves (including their grammatical description) or which have been entered manually into the appropriate partial lexicon. The lexica will be divided up according to various criteria of interest to the user.

6.2.2. The Basic Lexica

The most basic of these is a simple "**Inventory of Lexical Roots**". This information is completely retrievable from the examples (Levels +1 and -1) but it may also be entered directly into the lexicon.

This information can be viewed either as an alphabetical listing of all lexical roots, or it can be divided into various classes, such as basic word classes, as this information is also presented in the gloss (Level –1). In the English example given above, the lexeme was glossed as "eat:V". Here, the sign ": V" denotes that the lexical entry 'eat' is considered to be predominantly predicative in nature.

In addition to the generally accepted word classes, an alternative will also be provided for precategorial lexical entries, i.e., lexical entries which are not necessarily preferably used in one particular function (e.g., predicative or referential functions), but which may have both (or other) functions. This "neutral" or precategorial class will be denoted by ":PRC" in the gloss. This will be especially important for a large number of languages in which many or even most lexical entries can function both as predicates and nominals with the same ease, such as many Polynesian languages, Munda languages (Austro-Asiatic) and sign languages, but also in many others to varying degrees.

A further type of inventory, which also may be either automatically generated by the computer on the basis of the actual examples or which can be directly entered by the author, is the "**Inventory of Word Forms**", i.e., all entries which may be considered words. The information for this inventory comes from comparing Level +2 with Level +1 and by adding the gloss (Level -1). The user can then directly compare this with the "**Inventory of Lexical Roots**", discussed above.

The "Inventory of Grammatical Markers"

Another source of information which the gloss provides is an entry for the "**Inventory of Grammatical Markers**" which a particular language possesses.

Consider once again the English example 'I am eating'.

Level +2 Phonological transcription	aɪ	æm	ˈiːtɪŋ	
Level +1 Morphemic reconstruction	aɪ	æm	iːt	–ɪŋ
Level –1 Gloss	PRN.1.S.NOM	AUX.NPT.1.S	eat:V	-IPFV.PTCP
Level –2 Complex Morphology		PROG-eat-NPT.1.S		
Level –3 Sentence constituents	NP(SUBJ)	V-FIN	V-INF	
Level –5 English translation	'I am eating.'			

Level -1'	PRN.PERS.NUM.CAS	AUX.TNS.PERS.NUM	LEX:V	-ASP.WCD
Level -2'		ASP-LEX-TNS.PERS.NUM		

The English sentence above contains both an auxiliary verb (/æm/) and a participial marker (/–ɪŋ/), which are easily recognized by the computer as grammatical markers through the use of the small capitals in the gloss. These would then both be entered automatically to the list of grammatical markers of the language, along with their glosses (Level -1). Alternatively, they may also be entered by the author directly into the appropriate lexicon.

These grammatical markers may of course be viewed alphabetically. The list would then look something like the following (non-exhaustive!) hypothetical list for English:

Grammatical Marker	Grammatical Meaning
æm	AUX.NPT.1.S
aɪ	PRN.1.S.NOM
–ɪŋ	-IPFV.PTCP
ju	PRN.2.S.NOM
ʌv	POSS:PREP
–s	-NPT.3.S
–s	-POSS
–t	-PT
tu	DIR:PREP

More typically, one is interested in markers for specific verbal categories such as tense/aspect/mood, or case markers, adpositions, etc.

It is here that Level –1', the so-called "Templates", plays a central role in the organization of the lexica. As Level –1' provides the more abstract grammatical categories to which the grammatical morphemes in Level –1 belong, this part of the lexicon can also be viewed according to subparts of the grammar, i.e., verbal morphology, case morphology, adpositions, etc.

How this is achieved: In Level –1' in the English example above, we see that the categories of the morphemes in the gloss (Level -1) have been given the more abstract labels of the categories they belong to. For example, 'Nonpast' (NPT) belongs to the category 'Tense' (TNS), while 'eat:V' is a verbal lexeme, or 'LEX:V'. Also, '1' is a marker for 'Person' (PERS), while 'singular' (S) is of course 'Number' (NUM), etc.

On this basis, the user may choose how s/he would like the information to be organized. The user will be presented with a number of possibilities for viewing this information, such as the grammatical markers belonging to the category 'Tense' (TNS), 'Case', (CAS), 'Adpositions' (ADP), etc., and will then be shown all of the grammatical markers and their respective glosses which belong to this category. Our mini-"Inventory of Grammatical Markers" could then be reorganized along these principles, giving something like the following:

Grammatical Marker	Grammatical Meaning

Grammatical Markers for Tense	Grammatical Meaning
æm	AUX.NPT.1.S
–s	-NPT.3.S
–t	-PT

Adpositions	Grammatical Meaning
ʌv	POSS:PREP
tu	DIR:PREP

Pronouns	Grammatical Meaning
aɪ	PRN.1.S.NOM
ju	PRN.2.S.NOM

Grammatical Markers for Aspect	Grammatical Meaning
–ɪŋ	-IPFV.PTCP

Grammatical Markers for Possession	Grammatical Meaning
ʌv	POSS:PREP
–s	-POSS

Auxiliaries	Grammatical Meaning
æm	AUX.NPT.1.S

6.2.3. The Paradigms

Level –1' is also of use in presenting the data along another principle which is of interest to the linguist – the "**Inventory of Paradigms**". The necessary information for the paradigms is encoded both in the gloss entered by the author (Levels –1 and –2) as well as in the automatically generated Levels –1' and –2'. These Paradigms are then merely the systematic representation of this data according to various criteria. The paradigms can of course also be entered manually.

For example, the English example above would contain an entry for the

"**Paradigms of Pronominal Forms**", /aɪ/ with the gloss "PRN.1.S.NOM",

an entry for the

"**Paradigms of Progressive Forms**"(æm ˈiː.t–ɪŋ, "AUX.NPT.1.S LEX:V-IPFV.PTCP"),

and an entry for the

"**Paradigms of Auxiliaries**" (æm "AUX.NPT.1.S").

Further categories include "**Paradigms of Cases**", "**Paradigms of Adjectives**", etc.

This information can be viewed in any number of different combinations. Some examples:

1. First according to lexeme, then according to grammatical category. For example, the user can first select the lexeme 'eat' and then the verbal category "Progressive". This would then present the user with a complete paradigm of the lexical verb 'eat' in the progressive. For English, this would look something like the following:

Language: ENG(USA)			
Lexeme: 'eat:v' (= /iːt/)			
Category: Nonpast Progressive (NPT.PROG)			
Entries		Gloss	
æm	ˈiːt–ɪŋ	AUX.NPT.1.S	LEX:V-IPFV.PTCP
aɹ	ˈiːt–ɪŋ	AUX.NPT.2.S	LEX:V-IPFV.PTCP
ɪz	ˈiːt–ɪŋ	AUX.NPT.3.S	LEX:V-IPFV.PTCP
aɹ	ˈiːt–ɪŋ	AUX.NPT.P	LEX:V-IPFV.PTCP
wʌz	ˈiːt–ɪŋ	AUX.PT.1.S	LEX:V-IPFV.PTCP
wɹ	ˈiːt–ɪŋ	AUX.PT.2.S	LEX:V-IPFV.PTCP
wʌz	ˈiːt–ɪŋ	AUX.PT.3.S	LEX:V-IPFV.PTCP
wɹ	ˈiːt–ɪŋ	AUX.PT.P	LEX:V-IPFV.PTCP

   This would also include entries for other tenses, e.g. 'Future', as well as other moods, etc. By clicking on additional criteria, the user can further narrow down the choice, choosing only 'NPT', or '1' for first person, etc.

2. Alternatively, one could view the category "Progressive" in English or in several languages, with no reference to any one particular lexeme, in order to get an overall idea of the construction. This would present the user with a complete paradigm for all forms in the category "Progressive" for the respective language(s) and for all lexemes.

3. Another possibility is to click on one lexeme in the "**Inventory of Lexical Roots**" – in one or several languages – and then click "**Inventory of Paradigms**" to see all entries in all categories in which this lexeme is found. This would then be a complete paradigm for a single lexeme.

As this information is only gradually generated by the examples, which will probably have a strong preferance for the 3rd person, singular, the author can also chose to enter this information directly into this part of the lexicon. Either way, whether through the examples or

entered directly, a complete paradigm for all the various grammatical categories can be generated and presented.

6.2.4. The Schema Inventories

Finally, a user will often be interested in seeing the different ways in which the various sentential constituents combine in one or more languages to form sentences. That is, what is the word order of subject, object and verb in one language as compared to another, or the relative order of these constituents in main clauses, in subordinated clauses in general, or in a particular kind of subordinated clause.

This information is partially retrievable from the examples themselves, and partially from the grammatical description entered by the author to accompany each example. In the English example above, 'I am eating', we had the constituent order

NP(SUBJ) V-FIN V-INF

which is entered manually into the example by the author (in Level –3).

After the author has entered the full example, this constituent pattern will be entered automatically into the "**Inventory of Sentence Patterns**" and can be viewed by the user and compared with any number of other schematic inventories and, if the user wishes, these may all be accompanied by the examples from which they were derived.

The other schematic inventories are not derived from the examples alone but also by the information provided in the corresponding grammatical description which accompanies every example. These include the "**Inventory of Dependent Clause Patterns**", which are further divided into the "**Inventory of Complement Clause Patterns**" and the "**Inventory of Modifying Clause Patterns**". Some of the other categories included here are the "**Inventory of Nominal Phrases**" and the "**Inventory of Adpositional Phrases**".

As an example of the last two categories, consider the following English sentence:

Level+2 Phonological transcription	ðʌ	tu	laɪd͡ʒ	bʊks		ɔn	ðʌ	tebl̩	aɪ	ɛksˈ.pɛn.sɪv
Level +1 Morphemic reconstruction	ðʌ	tu	laɪd͡ʒ	bʊk	-s	ɔn	ðʌ	tebl̩	aɪ	ɛksˈ.pɛn.sɪv
Level –1 Gloss	DEF. ART	two: ADJ	large: ADJ	book: N	-P	LOC: PREP	DEF. ART	table:N	COP. NPT.3. P	expensive: ADJ
Level –3 Sentence constituents	NP(SUBJ)								V-FIN	PRED.ADJ
Level –5 English translation	'The two large books on the table are expensive.'									

In the course of the grammatical description of this example, the author will be asked for the complements of the predicate 'are expensive'. To this s/he will click on the node 'One complement' and fill in the respective semantic information. S/He will then be asked to specify the form of this complement, in this case 'Nominal Phrase', and answer various questions on the form of this NP (cf. section 5.2.2. for details). At the end of this description,

the author will be asked to include the schematic representation of this NP, which s/he will do as follows:

```
Please indicate schematic pattern:
```

DET	QUANT	ADJ	N	ADP

where DET = 'determiner', QUANT = 'quantifier', ADJ = 'adjective', N = 'noun' and ADP = 'adpositional phrase'.

At another point, the author will similarly be asked to type in the schematic pattern of the adpositional phrase, which will look like this:

```
Please indicate schematic pattern:
```

PREP	DET	N

where PREP = 'preposition'.

With this, the databank will then have two new schematic entries, one for NP's and one for ADP's, both of which may be viewed and compared with other NP's and ADP's, either in one and the same language (for example, according to grammatical function, semantic role, etc.) or cross-linguistically.

The same principle is also used to encode the structure of complement clauses, modifying clauses, etc., so that these may be compared with the structure of main clauses, both language internally and cross-lingustically. Compare the following German sentence:

```
Example:
```

Level+5 Standard orthography	'Ich wußte, daß er es nicht wußte.'										
Level +2 Phonological transcription	ɪç	'vʊstə			das	eɐ	ɛs	nɪçt	'vʊstə		
Level +1 Morphemic reconstruc-tion	ɪç	'vʊs	.–t	–ə	das	eɐ	ɛs	nɪçt	'vʊs	.–t	–ə
Level –1 Gloss	1. S. NOM	know :V	-PT	-1.S	CMPL	3.S.M. NOM	3.S.N. NOM	NEG	know:V	-PT	-3.S
Level –3 Sentence constituents	NP (SUBJ)	V-FIN			CMPLCL(OBJ)						
Level –5 English translation	'I knew that he didn't know it.'										

In entering this example, the author has already been asked to give the schematic structure of the sentence as a whole (cf. Level -3, where CMPLCL = 'complement clause').

In the course of the grammatical description of this complex sentence, the author will be asked to give the schematic pattern of the complement clause itself, which would have the following form:

```
Please indicate schematic pattern:
```

CMPL NP(SUBJ) NP(OBJ) NEG V-FIN

With these two entries, the computer now has two new schematic patterns for the inventories, one for sentences, and one for complement clauses. If the user wishes, s/he can view and compare both simultaneously, which would then look something like this:

Inventory of Sentences

NP(SUBJ) V-FIN CMPLCL(OBJ)

Inventory of Complement Clauses

IN OBJECT FUNCTION:
CMPL NP(SUBJ) NP(OBJ) NEG V-FIN

If desired, the user can also simultaneously view the respective examples.

Of course, in our case the computer only has one entry for each type whereas normally all of the various schematic patterns at a particular level which have been entered for a particular language will be presented simultaneously.

However, even with this small databank, our imaginary user will now probably have noticed just from these two examples that main clauses in German can have a different word order (verb-medial) than certain types of complement clauses (verb-final).

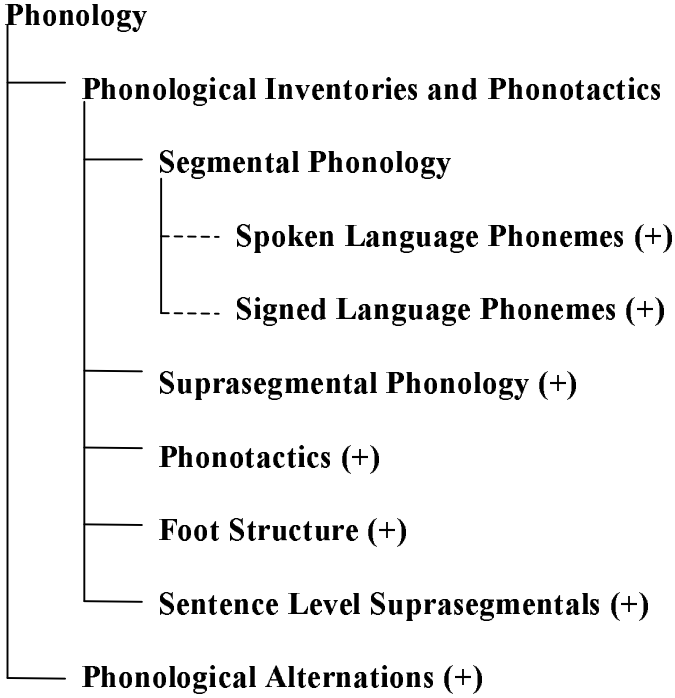
7. Phonology

The section of the grammar devoted to phonology has two purposes: 1. To provide an inventory of all distinctive phonemes of the language, both segmental and suprasegmental, and 2. To provide an inventory of distinctive phonological alternations, covering everything from metathesis, addition, and deletion to tonal changes. We would also like to know as much as possible about foot structure, intonational patterns and their meanings, and the phonotactics of the language. Of all sections of the grammar, "Phonology" most heavily depends on the criteria given in Comrie & Smith (1977), while of course not being an exact duplicate of this.

The first major division is between "Phonological Inventories and Phonotactics" and "Phonological Alternations". Let us begin with the phonological inventories.

7.1. Phonological Inventories and Phonotactics

These are divided into "Segmental Phonology", "Suprasegmental Phonology", "Phonotactics", "Foot Structure" and "Sentence Level Suprasegmentals". Already at this level, however, we must take the differences between the mechanisms used to produce sign languages and those used to produce spoken languages into account, as it will not be possible to proceed in the same manner to describe both. Thus, "Segmental Phonology", etc., is immediately divided into "Spoken Language Segmental Phonology" and "Signed Language Segmental Phonology". Compare the following diagram of the uppermost nodes under phonology:



As we would like a description of all of the various areas of phonology, these are all of the required type - i.e., the author must provide information on all areas. If these do not apply to a particular language, for example, information on tones in a non-tonal language, this information must also be entered into the databank. By using a required node, the author is thus required to include this information - i.e., the respective language has no tones - instead of simply not providing information here, so that the consultant of the grammar can be sure that a particular area of phonology does not apply to this particular language.

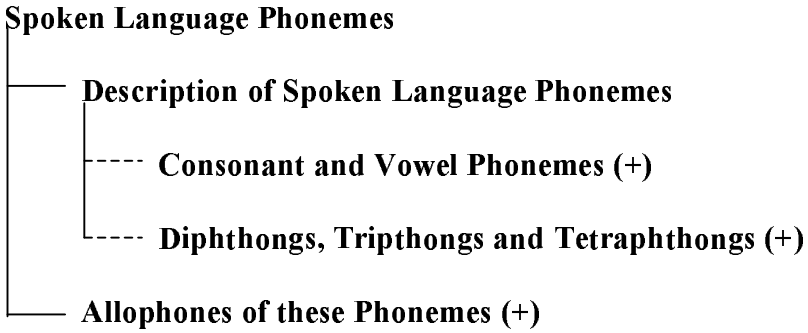
Once the author has begun to fill out a description of a particular area of phonology, e.g., segmental phonology, we have the exclusive nodes referring to the nature of the language, as the author is describing EITHER a spoken OR a signed language, but of course not both simultaneously.

7.1.1. Segmental Phonology

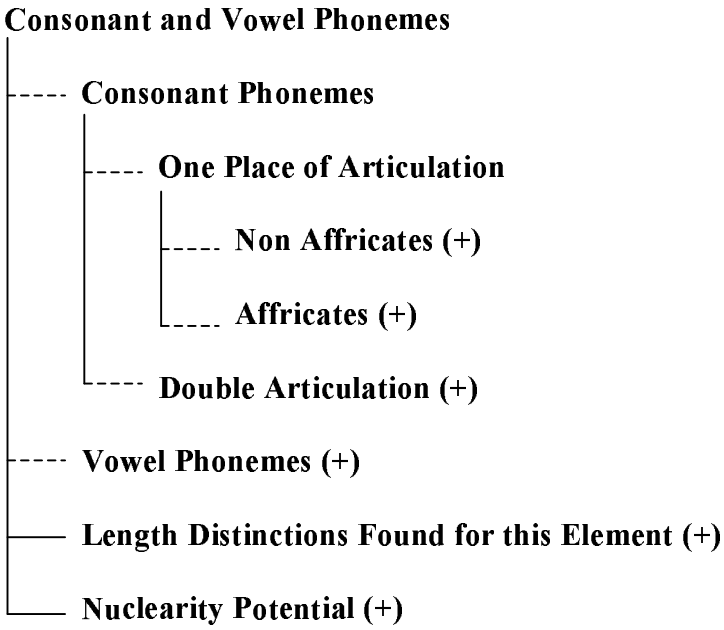
7.1.1.1. Spoken Language Phonemes

For the sake of discussion, let us take a more detailed look at the description of spoken language phonemes. After this, we will also show how the description of signed language phonemes is entered into the databank. This detailed discussion should serve to illustrate the main strategy involved in phonological description in the CRG, so that the remaining areas need not be described in such great detail.

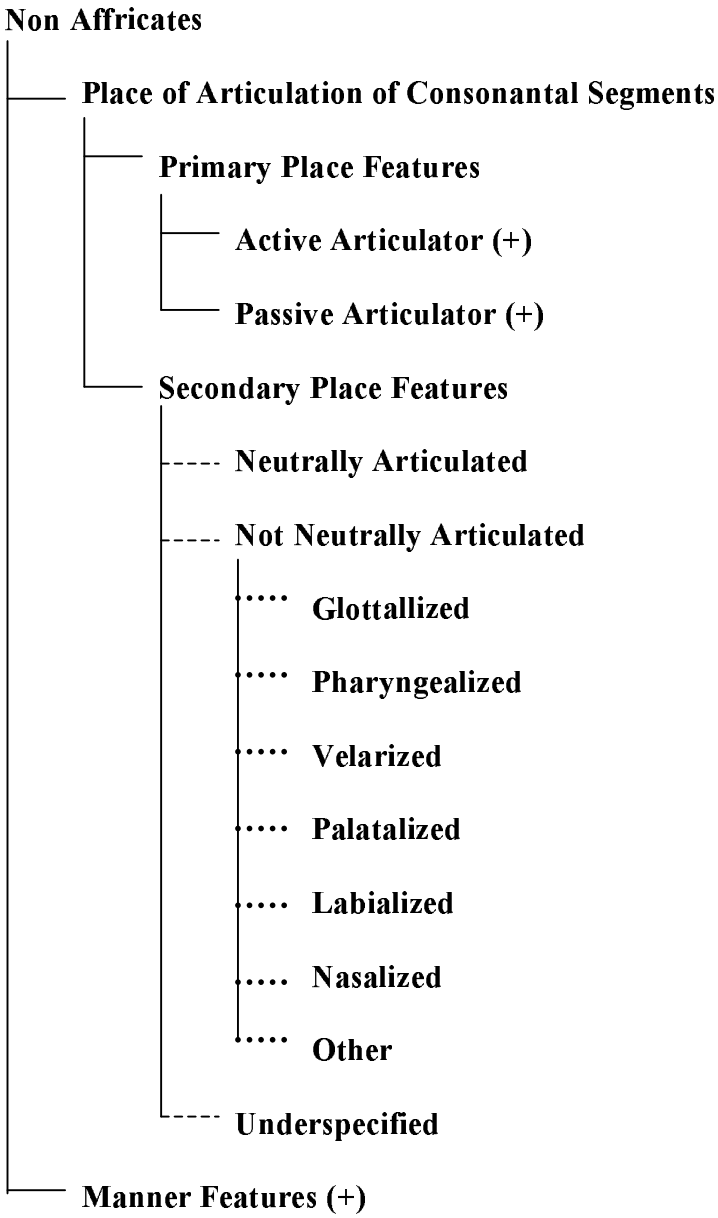
First, the uppermost nodes of spoken language phonology:



When describing a phoneme, there are two things we need to know: a description of the distinctive features of this phoneme, and its actual realization in different environments, that is, its allophones. Hence, both are required nodes. If we follow the description of the phoneme one step further, we are forced to indicate whether the phoneme is a consonant or vowel, or a diphthong, triphthong or (rare) tetraphthong. Let us assume for the moment that we are describing a consonant (say, a plosive). We then click on "Consonant and Vowel Phonemes" and find the following:

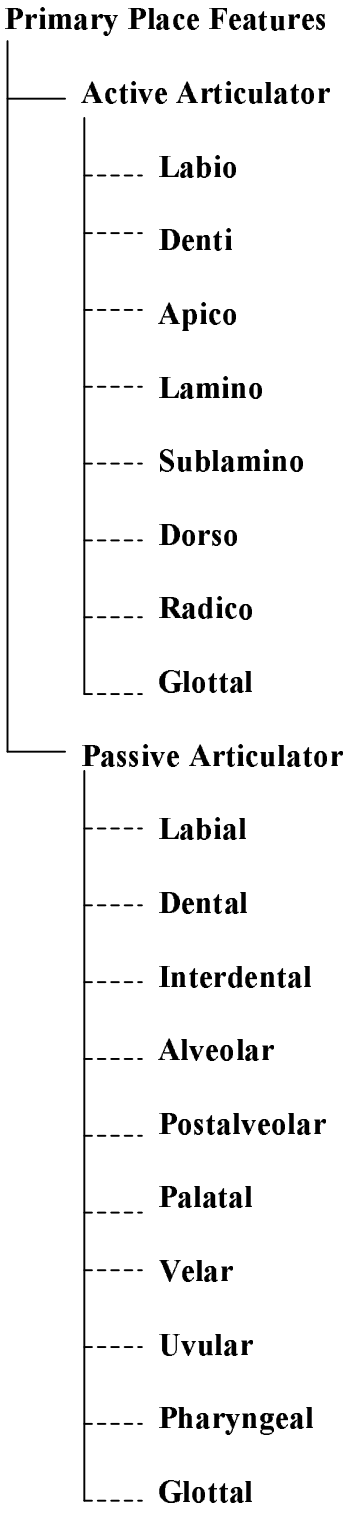


For both consonantal and vocalic phonemes, we would like to know the nuclearity potential of the phoneme (i.e., may or may not serve as a syllable nucleus), as well as any possible length distinctions (short, half short, long, extra long). If it is a consonantal phoneme, say a phoneme which the author chooses to represent with a //T//, then it may further be described as to the number of places of articulation (in this case, one) and whether or not it is an affricate (here, nonaffricate). If we follow the node "Non Affricates" further, this is what we find:

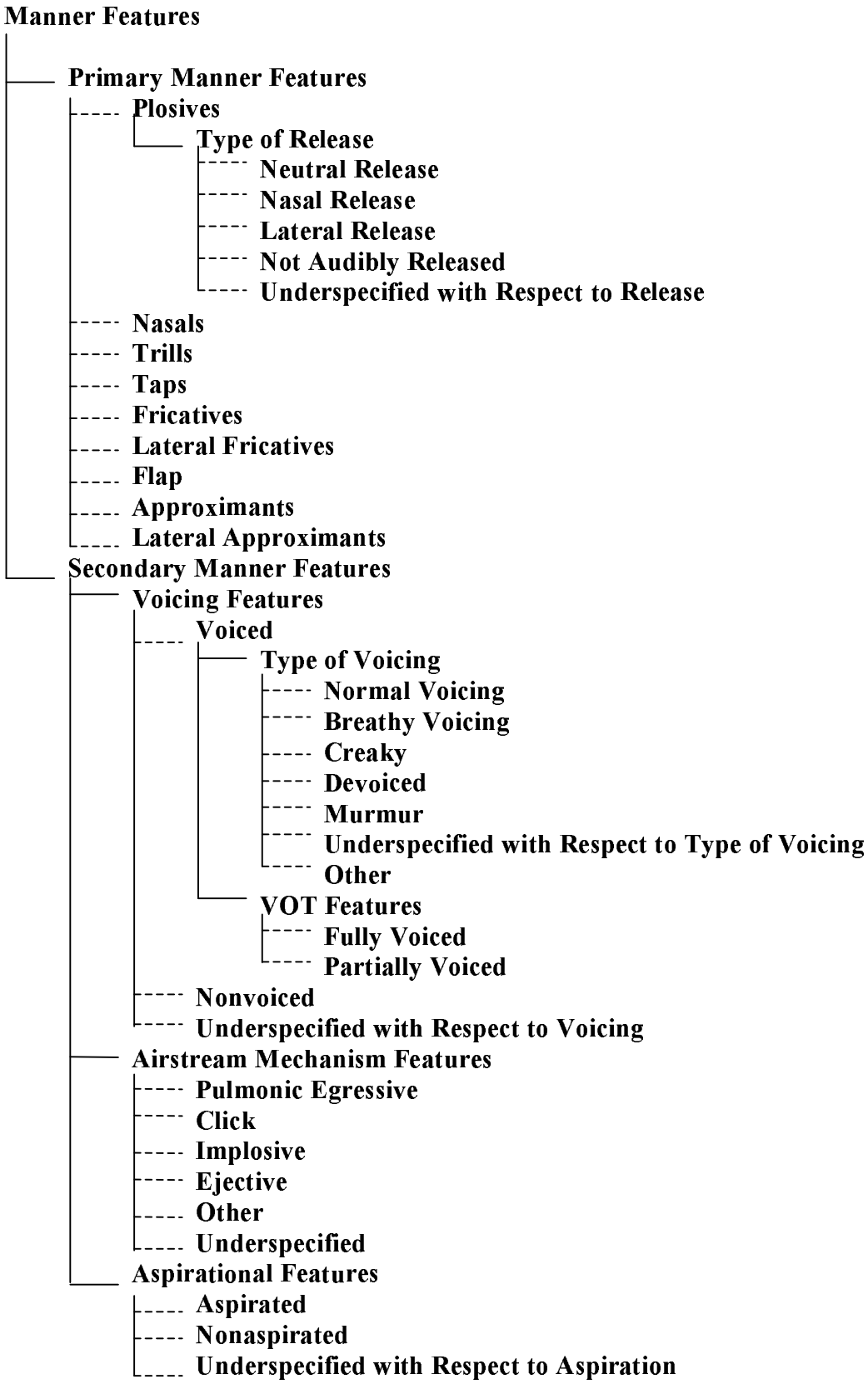


That is, any consonant with one place of articulation and which is not an affricate will have a place of articulation, with primary and secondary place features. The primary place features are the active and passive articulators, while the secondary place features include glottalization, labialization, etc. The manner features include information on voicing, airstream mechanisms and of course, plosive, nasal, etc. These criteria are shown in the following diagrams.

The active and passive articulators are the following:



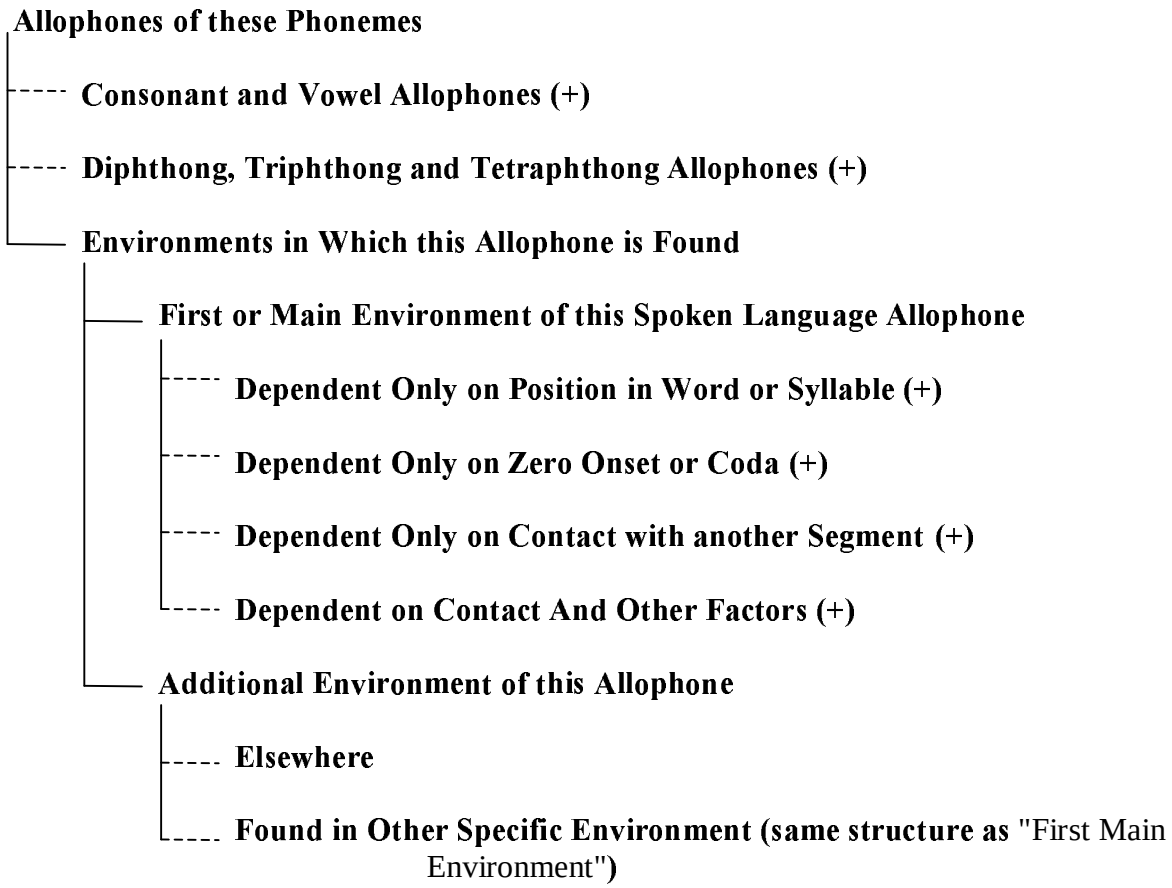
The following diagram shows the manner features:



Proceeding in this manner, it should by now be clear how the phoneme can be described in its entirety according to its distinctive features. The same basic principles are also applied to vowels, although the criteria are of course different. These include "Height Relation", "Front Back Relation", "Rounding", "Nasalization", "Advanced/Retracted", "Rhoticity", etc. and need not be given here in detail.

Again, the same principles as those used for the vowels are required for diphthongs, triphthongs and tetraphthongs. This is further complicated by the fact that we must also include this information - as well as information on length, etc. - for each constituent of the diphthong, etc.

Turning now to the allophones, we find the following uppermost nodes.



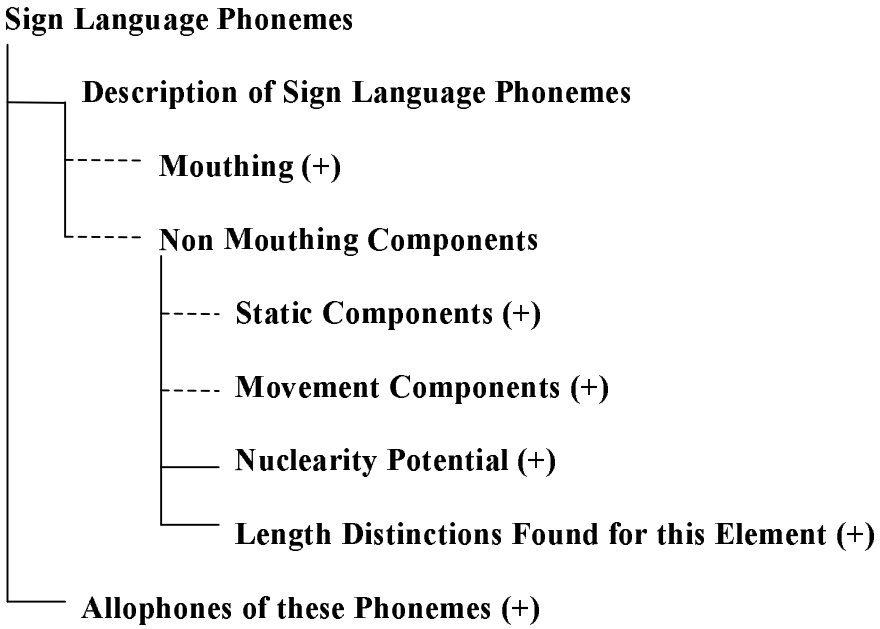
Thus, if a phoneme has allophones (this question is not shown in the diagram above), it must have at least two allophones which will be found in separate environments. In the case that there are more than two environments, "Additional Environment" can be clicked on more than once. Also, as allophones are generally described as having a specific form in one environment and another form "elsewhere", this alternative is also presented under "Additional Environment", although this environment may of course also be specific.

The criteria for determining this environment are as follows: the allophone may be dependent on the position of this phoneme in the word (for example, what is written as an "r" in Catalan is not pronounced word-finally, i.e., it has zero form), or this may depend on the position in the syllable (cf. the two pronunciations of "l" in English), or it may depend on contact with another segment. Here, there are two possibilities. This may depend on contact with certain segments, or it may depend on specific features of segments in general. Thus, the author may choose - according to his or her needs - either specific consonants, and enter these individually, or s/he may chose from among the features of the segments, discussed above in detail, which are all offered here individually.

Finally, there is the possibility that we have a combination of these alternatives, for example, the phoneme has a certain form in the coda after another certain segment, etc. This is taken care of under the node "Dependent on Contact and Other Factors".

7.1.1.2. Sign Language Phonemes

Let us now turn to the treatment of sign language phonemes. Below is a diagram of the uppermost nodes involved:



While the principles involved here are essentially the same - i.e., a full description of the distinctive features involved - the criteria offered are very different. Note however that the uppermost division is the same as with spoken languages - Phoneme and Allophones of this Phoneme.

Under phonemes, the first division is between "Mouthing" and "Non Mouthing Components". Mouthing here refers not to signs made with the mouth or to suprasegmental phonemes (see below), but rather to the use of the mouth to make the same forms as are made in the production of speech. Thus, while we are not interested here in the SOUNDS produced by the vocal organs, the FORMS these take can be described along the same lines as the production of sounds. Hence, this node has basically the same structure as does that of "Spoken Language Phonemes" given above.

If we turn to the "Non Mouthing Components", we notice that this is divided up into "Static Components" and "Movement Components", both of which can also be described in terms of their nuclearity potential and possible length distinctions, as with spoken language phonemes.

Let us now take a closer look at these criteria.

Static Components

- **Static Manual Elements**
  - **Description of Hand (+)**
  - **Wrist Orientation (+)**
  - **Position of Arm (+)**
  - **Place of Execution (+)**
- **Static Non Manual Elements (under construction)**

The criteria used here are primarily those used in HamNoSys (Prillwitz et al., 1989), although there have been slight adaptations made.[10]

"Description of Hand" includes a full description of the hand, including hand orientation, hand shape, and whether it is the dominant or non-dominant hand (or possibly both).
"Hand Orientation" includes the following information:

- PALM ORIENTATION. This has both a vertical component and a horizontal component, i.e., is the palm facing up or down, to what degree, towards which side (ipsilateral, contralateral)? Does the palm face the signer, the viewer, and to what degree?

- ORIENTATION OF THE BACK OF THE HAND. This is measured along the line from the wrist to the middle knuckle and is essential for a full description of the hand position. As with PALM ORIENTATION, this also involves a horizontal and a vertical component.

"Hand Shape" involves the following criteria:

- FORM CLASS. There are three of these in HamNoSys: FIST, FLAT HAND, INDIVIDUAL FINGER(S).

- Is this form class OPEN or CLOSED?

- POSITION OF THUMB OR F1

- POSITION OF REMAINING FINGERS. This involves a number of criteria:

- Do the fingers F2 through F5 (i.e., index finger through pinkie) have the same shape or not? If so, the fingers may all be described together. If not, each finger must be described individually.

- Are these fingers RELAXED or NOT RELAXED? If they are not relaxed, are they STRAIGHT, BENT OR CURVED (and at which joints?) or STRAIGHT FINGER BENT AT THE JOINT OF THE HAND?

[10] Although we assume that there are static segmental non-manual distinctive phonemes, there seems to be no literature as yet on this topic. Hence, we were not able to include any information here, as it is not clear to us which criteria could be relevant here. This information can easily be included at a later date.

- We also need to know the DEGREE OF SPREAD of the fingers from their respective neighbors.

"Wrist Orientation" is described according to the following criteria:

The wrist is either STRAIGHT IN LINE WITH THE FOREARM or it is BENT in relation to this. If it is bent, then we need the DIRECTION OF BEND (TOWARDS PULSE, AWAY FROM PULSE, TOWARDS F5 (PINKIE) or TOWARDS F1 (THUMB). We of course also need the degree of the bend.

The **"Position of Arm"** is easily dealt with and has the following three exclusive nodes:

- OUTSTRETCHED
- PULLED IN CLOSE TO BODY
- UNDERSPECIFIED

Here, we are merely concerned with whether the sign is formed far away from the body or close to it, or whether this information is not distinctive.

We now come to the last topic involved here, that of the **PLACE OF EXECUTION**. This refers to where on the body the manual sign is executed, whether there is contact with this body part or not, and which side of the body (IPSILATERAL, i.e., the same side as the dominant hand, or CONTRALATERAL, the side of the non-dominant hand). The locations are divided up as follows. For body parts which come in pairs, there is is also the additional question whether only one is intended or both.

- ABOVE HEAD
- LOCATIONS ON THE HEAD
  - HAIR
  - TOP OF HEAD
  - SIDE OF HEAD - EAR (ear lobe, etc.), TEMPLE
  - BACK OF HEAD
  - FACE (FOREHEAD, EYEBROW, EYE, BETWEEN EYES OR EYEBROWS, NOSE (and which part of nose, e.g. nostrils, tip, etc.), MOUTH (with further specifications for lips, teeth, tongue, etc.), BETWEEN CHEEK AND MOUTH, CHEEK, CHIN, JAWBONE, UNDERSIDE OF JAW, OTHER
- NECK
- THROAT
- LOCATIONS ON TORSO (CHEST, STOMACH, ABDOMEN, CROTCH)
- LOCATIONS ON ARMS (FOREARM, etc.)

- LOCATIONS ON HANDS (FINGERS, PALM, WRIST, BACK OF HAND, etc.)
- LOCATIONS ON LEGS (KNEES, SHINS, etc.)
- LOCATIONS ON FOOT (TOES, ANKLE, etc.)

We now come to the "Movement Components", whose uppermost nodes are as follows:

Movement Components

- **Manual Movement Components**
  - **Primary Movement (+)**
  - **Secondary Movement (+)**
  - **Movement Relations of Two Handed Signs (+)**
  - **Repetitions (+)**
- **Non Manual Movement Components (+)**
- **Possible Speed of this Movement (+)**
- **Possible Sizes of this Movement (+)**

The first distinction made here is between manual and non-manual movement - i.e., a movement must be one and only one of these two alternatives. Manual movements are of different types, only one of which will be described at a single time. The possibilities here are the following:

- PRIMARY MOVMENTS. These are movements of the hand in a specific direction or in a series of directions. The movements here are the following:

- MOVEMENTS WITH MORE THAN ONE SIGNLE GENERAL DIRECTION, i.e., X-MOVEMENTS, CROSS MOVEMENTS, SINGLE MOVEMENTS WITH RETURN TO BEGINNING POINT. All of these movements can (and must) be further specified (vertical, horizontal, etc.)

- MOVEMENTS WITH A SINGLE GENERAL DIRECTION - CIRCULAR MOVEMENTS, NONCIRCULAR MOVEMENTS (STRAIGHT LINE, SINGLE CURVE, SNAKE MOVEMENT OR SQUIGGLE, ZIGZAG MOVEMT, MOVEMENT AT WRIST), SPIRAL MOVEMENT (with the further specifications of SIZE OF THE SPIRAL IN THE COURSE OF TIME (i.e., does the spiral increase its size in time, become smaller or stay the same size), DIRECTION OF CIRCULAR MOVEMENT (CLOCKWISE, COUNTER-CLOCKWISE, etc.)

- SECONDARY MOVEMENTS. These include CHANGES (in hand orientation, position of arm, hand shape, place of execution) and MANUAL TRILLS, i.e., movement of all fingers while the hand remains stationary.

The manner of dealing with **allophones** here is essentially the same as that with spoken language. However, as the author was not able to locate any literature on this topic, this

section of the grammar must still be considered tentative and, as work in this area progresses and becomes more widely known, it can be built into the system at a later date. Hence, the main criteria here are whether the allophone depends on contact with specific segments in general or on the distinctive features of segments, in which case all of the features discussed above are offered. As the manner of proceeding here is essentially the same as with spoken-language allophones, we need not deal with this topic further here.

7.1.2. Suprasegmental Phonology

We now come to the suprasegmental phonemes of both spoken and signed languages. As the manner in which this data is encoded into the databank should be clear by now, in the following I will restrict my comments only to the basic criteria involved.

For SPOKEN LANGUAGES, the primary criteria here are the following:

- LENGTH DISTINCTIONS, and for which segments (vowels only, consonants only, both vowels and consonants, etc.)

- DISTINCTIVE TONAL FEATURES - all LEVEL TONES, CONTOUR TONES, ATONICS, GLOBAL RISE, GLOBAL FALL, UPSTEP, DOWNSTEP. Are these restricted to certain portions of the lexicon?

- DISTINCTIVE WORD ACCENT. If the language has distinctive word accent, does it have primary and secondary accent?

For SIGNED LANGUAGES, we have the following criteria. As work on this area appears to be rather limited, this area of the grammar is likely to be one of the first to be expanded in the near future. The criteria selected here result from the author's own experience with German Sign Language.

- DISTINCTIVE SIZE. Does the language have distinctive sizes and if so, what are these? Also, for which segments is size distinctive - movements, static components, both?

- DISTINCTIVE SPEED. Does the language have distinctive speeds and if so, what are these?

- DISTINCTIVE SIMULTANEITY OF SIGNS

- DISTINCTIVE SIGN LENGTH. If the language has distinctive length, which length distinctions does it show?

7.1.3. Phonotactics

Here, we are concerned with phenomena such as e.g. which consonant clusters, if any, are found in a language, which segments may occur word-initially, word-finally, in the onset, coda, etc.

Again, for this portion of the grammar we had no data at our disposal for signed languages. While it would have been possible to create a list of what we considered to be "logical possibilities", this would have been at best educated guess-work and we have decided to wait until research in this area has been brought to our attention. Of course, the electronic medium

chosen allows us to incorporate this information easily at a later date, so that we had no misgivings in chosing this option. Thus, this section is devoted entirely to the phonotactics of spoken languages.

7.1.3.1. Segmental Phonotactics

Here, we are primarily concerned with phenomena such as syllable structure, possible consonant clusters, restrictions on word-initial and word-final segments, moraic restrictions, consonant and vowel harmony, etc.

We first distinguish between **segmental occurrence** and **segmental co-occurrence phenomena**. Under both we also distinguish between restrictions in WORDS and LEXICAL MORPHEMES. As the two have virtually identical structures, we will restrict our comments in the following to (co-)occurrence restrictions in words.

Under segmental occurrence restrictions we find the following criteria:

- BASIC SYLLABLE STRUCTURE, i.e, all possible forms of the onset, nucleus and coda, such as ZERO, C, CC, CCC, CCCC, CCCC for onset and coda. For the nucleus, this is somewhat different, containing VOWELS, CONSONANTAL NUCLEI (NASALS, APPROXIMANTS, etc.), DIPHTHONGS, TRIPHTHONGS and TETRAPHTHONGS

- RESTRICTIONS ON WORD-INITIAL/FINAL SEGMENTS, i.e., consonants are/are not allowed here, vowels are/are not allowed here, consonant clusters are/are not allowed. Also, if consonants (etc.) are allowed, are there some which are not allowed here?

- MORAIC RESTRICTIONS. If there are moraic restrictions, what is their nature (mono- and bimoraic only, etc.)? Are there moraic restrictions on the first or last syllables or on internal syllables? We would also like to know the MORAICITY OF CONSONANTS IN THE CODA, as these are not always moraic.

- ARE THERE SEQUENCES OF SYLLABIC VOWELS? If so, are these restricted in any way?

Under co-occurrence restrictions, the following criteria are requested:

- Are there CONSONANT CLUSTERS in the language? What are the clusters found in initial and final position? Are medial clusters equal to the set of initial and final clusters combined, or are there more medial clusters?

- CONSONANT and VOWEL HARMONY. Here we first need to know whether the language has harmony restrictions and if so, we need information on the type of harmony (i.e., palatal, etc.), type of dominance, domain of harmony, neutral segments in harmony, and of course, the complementary vowel and consonant sets.

- Restrictions between the NUCLEUS AND ONSET/CODA

Finally, we would like to know how medial units are assigned to syllables, i.e., the Maximal Onset Principle only, or is this dependent on morphological structure (i.e., between morphs in general, between grammatical and lexical morphs, etc.)?

7.1.3.2. Suprasegmental Phonotactics

Here we are primarily concerned with tonal co-occurrence restrictions, such as TONAL SEQUENCE RESTRICTIONS, CO-OCCURRENCE RESTRICTIONS BETWEEN TONE AND LENGTH, RESTRICTIONS BETWEEN TONE AND SYLLABLE STRUCTURE, and RESTRICTIONS BETWEEN TONE AND STRESS. We will present these criteria briefly here, without going into unnecessary details.

With respect to tonal co-occurrence restrictions, we need to know first of all whether the language has tonal co-occurrence restrictions before we proceed to describe them or if it even has distinctive tones. If it does, then we will need a full description of the tonal combination which is not allowed, the tonal combination which is actually found in its place, and of course the factors contributing to this restriction. These restrictions can be either CONTACT or DISTANCE RESTRICTIONS, and if they are distance restrictions, we need to know the intervals between them - how far apart are they, are there neutral tones which may occur between these, etc.?

With the restrictions between tone and syllable structure, we need to know if there are restrictions between certain tones and the onset, the nucleus and the coda and of course, what types of restrictions apply. Which tones are affected, which tones are not affected?

Similar comments apply to the restrictions on tone and distinctive stress, provided of course that the languge has distinctive stress. If it does, which tones may occur in stressed syllables, which may or may not occur in unstressed syllables?

Finally, as this small list cannot be considered exhaustive, there is an "Other" node where the author is free to add - in prose form - any restrictions which s/he feels necessary.

7.1.4. Foot Structure

With respect to foot structure, it would probably be best to simply display the "tree format" here, which should be self-explanatory. The same format can be used for spoken and signed languages. The criteria here are adapted from Gussenhoven & Jacobs (1998).

Foot Structure
- **Not Applicable**
- **Foot Structures in This Language**
  - **Indicate All Types of Feet Found**
    - **Monosyllabic Feet**
    - **Disyllabic Feet**
    - **Trisyllabic Feet**
    - **Tetrasyllabic Feet**
    - **Pentasyllabic Feet**
    - **Other**
  - **Feet Structure Determined by**
    - **Stress Pattern**
    - **Tonal Pattern**
    - **Other**[11]
  - **Bounded or Unbounded Language**
    - **Bounded Language**
      - **Quantity Sensitivity**
        - **Quantity Sensitive**
        - **Quantity Insensitive**
        - **Not Applicable (see note 11)**
      - **Left or Right Dominance**
        - **Left Dominant**
        - **Right Dominant**
        - **Other**
      - **Monosyllabic Root in Words with Odd Number of Syllables**
        - **Monosyllabic Root is Stressed**
        - **Monosyllabic Root is Destressed**
        - **Other**
        - **Not Applicable**
    - **Unbounded Language**

7.1.5. Sentence-Level Suprasegmentals

We now come to the last section of phonological inventories and phontactics, that of the sentence-level suprasegmentals. As usual, we first distinguish here between spoken and signed languages, due to the inherent differences in the nature of the inventories of these suprasegmentals. However the same basic principles hold here as well.

In both cases, we need to know whether the intonational unit is a complete intonational phrase or not. For example, in the sentence *Jim - you remember him? - was telling me that...* Here, the occurrence of *Jim* - which forms an intonational unit with *was telling me that...* has been

[11] This node is offered especially for signed languages, as it is not yet obvious which factors (speed? size? abruptness of stop?) should be taken as decisive in signed languages.

separated by a polar question. If we have a discontintuous intonational phrase, we must therefore first ask which intonational unit this forms a discontinuous phrase with.

The only real differences between spoken and signed languages are to be found in the actual expression of the "intonation". In spoken languages, these are of three possible kinds, which may of course be combined (and are hence included as optional nodes): LOUDNESS, PITCH and DURATION. In signed languages, there are a large number of possibilities, and we cannot be certain at present that our list is exclusive. The ones known to this author, taken partly from Hillenmeier et al. (1996b) include the raising of eyebrows, nodding, shaking the head, leaning to one side/forward/backward, puffing the cheecks, wide eyes, etc. Here, an attempt was made to complement those forms of suprasegmentals known to me with other logical possibilities. As research on this area continues, this section of the grammar can be updated.

The entering of the "intonational" pattern is also different for spoken and signed languages. In spoken languages, following the discussion in Ladd (1996), we require the author to indicate the left boundary tone, pitch accents, phrase tone and right boundary tone. In signed languages, on the other hand, we require only the beginning and end of each sentence-level intonational marking form, as there may be two or more simultaneous markers here (for example, raised eyebrows and shaking the head in German Sign Language to indicate a leading polarity question where an affirmative answer is expected).

The last two points are the same for both forms of language - "Meaning of Suprasegmental Pattern" and "Minimal Suprasegmental Pairs".
As possible alternatives for the meaning of the intonational pattern, we have offered illocutionary modalities and presentational structure, as these are by far the two most commonly found (and perhaps the only ones in spoken languages). As usual, we have added an "Other" node here, which will be necessary at least for signed languages, where "intonation" can also express negation, etc.

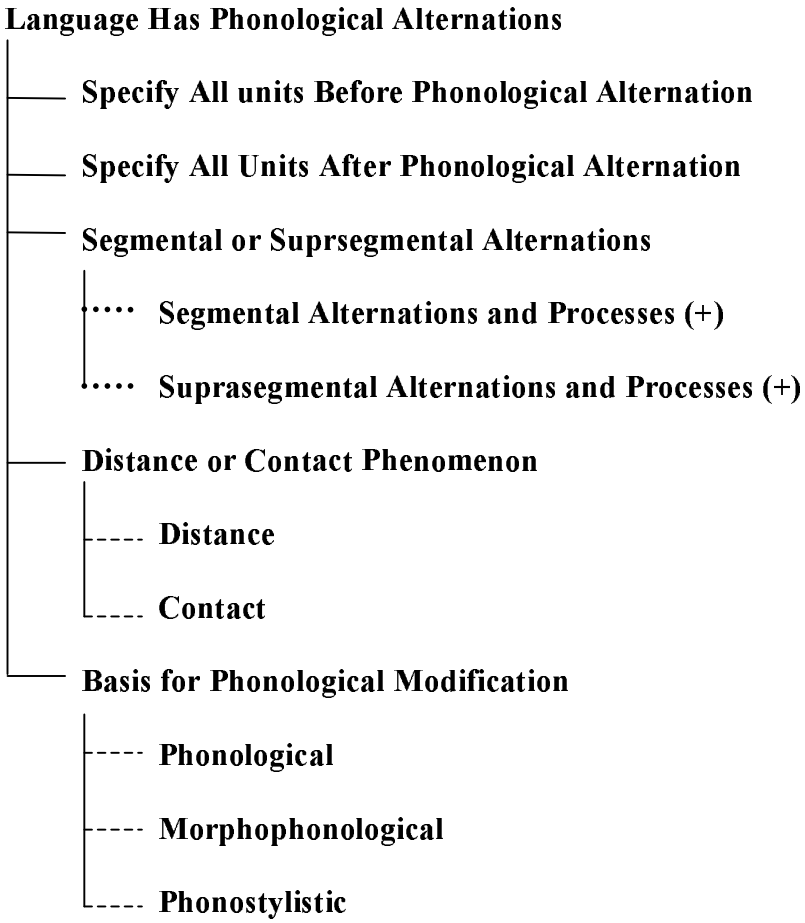
"Minimal Suprasegmental Pairs" is, as its name suggests, a node where we request the author to specify whether the given intonational pattern is part of a minimal intonational pair and, if so, to describe both the form and meaning of the other member of this pair along the criteria given above.

7.2. Phonological Alternations

This is the final node in the section on phonology. It is a required node, so that the author must enter information here. The first node here offers two mutually exclusive alternatives: Either the language has phonological alternations, or it has none. If it has alternations, we would of course like a detailed description of these.

The first part of the description is the most simple: The author here is requested to specify all units, both before and after the phonological alternations have taken place, so that the consultant can see the actual example for him-/herself.

The remaining uppermost nodes have the following form:



In this manner, the author is required to specify both whether the alternation is a distance or contact phenomenon as well as how this alternation is motivated, i.e., purely phonologically, morphophonologically, or phonostylistically.

Note also that a given example may contain cases of both segmental AND suprasegmental alternations, so that these are offered as optional nodes, i.e., both or either may be chosen as necessary.

Under segmental and suprasegmental alternations and processes, we again have the distinction between signed and spoken languages and again, we have no criteria for signed languages, as we had no relevant data to use as the basis for our choice. As repeatedly

mentioned throughout the present paper, this data can easily be added at a later date as research in this area progresses.

Concentrating thus on spoken languages, we have segmental and suprasegmental alternations and processes. Under segmental alternations and processes, we have the following alternatives: METATHESIS, DELETION AND ADDITION, CONTRACTION AND EXPANSION, and ASSIMILATION AND DISSIMILATION. Under suprasegmental alternations and processes we have CHANGE IN ACCENT and CHANGE IN TONE.

While metathesis may require no further comment, we should add a few words here on deletions and additions. We must of course first enquire as to whether it is an example of addition or deletion, as well as enquire as to what is added or deleted ("focus") and the domain in which this addition/deletion takes place. These two last criteria have the same alternatives: SEGMENT, SUBSYLLABIC (ONSET, NUCLEUS, CODA, RHYME), SYLLABLE, FOOT, PROSODIC WORD, PHONOLOGICAL PHRASE, INTONATIONAL PHRASE. Finally, there is the question as to what part of the domain this addition/deletion takes place in: INITIAL, MEDIAL, FINAL.

Similar questions concerning the exact status of contractions/expansions or assimilation/dissimilation must also be answered, the details of which need not interest us here, so that the tree has the following general form:

Segmental or Suprasegmental Alternations
- **Segmental Alternations and Processes**
  - **Spoken Language Segmental Alternations and Processes**
    - **Metathesis**
    - **Deletions and Additions**
      - **Deletion or Addition (+)**
      - **Focus: Deleted or Added Unit (+)**
      - **Domain (+)**
      - **Position of Focus in Domain (+)**
    - **Other Alternations and Processes**
      - **Contraction and Expansion (+)**
      - **Assimilation and Dissimilation (+)**
  - **Signed Language Segmental Alternations and Processes (under construction)**
- **Suprasegmental Alternations and Processes**
  - **Spoken Language Suprasegmental Alternations and Processes**
    - **Change in Accent**
    - **Tonal Change**
  - **Signed Language Suprasegmental Alternations and Processes** (under construction)

8. Summary

We have now covered all major areas of human language. There is much more detail in the format than what has been shown here, but to show all of this in detail would have been impractical, as the current paper would have at least doubled in size.

There are also many other aspects of language which have not been integrated into the current format but which could easily be integrated at a later date. One example of this: The current format is strictly synchronic in nature. We have not taken diachronic developments into consideration at all. While we firmly believe that this is a good starting point for an enterprise such as the present one, this type of information - and many others - can be integrated into the format at a later date if this is considered desirable. This could conceivably consist of an "Etymology" branch, which offers the most common paths of grammaticalization and which is then attached to the tree each time a lexical form is requested. As has been repeatedly stressed here, the electronic medium chosen, with its inherent flexibility, ensures that the present format can be expanded (or shortened) as the need arises, so that in the course of time it can be adapted to meet many needs which are not met here.

The next step of course is to test the present format with as many languages as possible. Indeed, this is the stated goal of the present format. We have constantly tried to take data from as many languages as possible into consideration, and have checked virtually all areas of grammar with examples taken straight from various grammars, but we are well aware of the fact that much has slipped by unnoticed.

Also, it is quite possible that some of the nodes in the actual tree format have on occasion "slipped" onto the wrong level. With several thousand nodes connected to each other in a very complex fashion, this is hardly to be avoided in a first edition. These types of errors, as well as an occasional logical error, can only be "weeded out" by applying the grammar format over and over again to as many typologically diverse languages as possible. This will be a long and slow process. Nevertheless, the format - even in its present form - is already capable of documenting large portions of the grammar of any language in a very detailed fashion and comparing these data with those of other languages. This of course requires that the data be entered into the format.

Additional work is necessary on some portions of the grammar format. This especially applies to the phonology of sign languages. Our hands were tied here to a large extent by the simple fact that much basic work still needs to be done here, and even very basic concepts such as the structure of the "syllable" are the topic of much controversy in sign languages (cf. e.g. the discussion on the structure of the syllable in sign languages in Brentari, 1995:624ff).

We, as non-experts in the phonology of sign languages, of course cannot decide these matters alone. We have included as much detail on these phenomena as seemed feasible, but it was often not possible to continue here and we have had to leave these sections for future research. This of course also applies to other sections of the grammar, although to a much lesser extent, so that the minimal basic assumptions we have had to make (cf. section 2) should allow us to deal with these sections of the grammar successfully.

Despite all these caveats, it must be stressed here that, if it is possible to study human language in a consistent manner and to make assumptions as to the categories which human languages can possess and how these categories can be marked, then a grammar format such as the present one MUST be possible. It can of course only be as good as our current state of

knowledge. It is our hope, however, that such a format will help serve to point out those areas of grammar where our knowledge is still very limited, while at the same time suggesting possible solutions, merely by presenting the data in structured form, thereby suggesting analogies to other sections of grammar.

9. Bibliography of Literature Used in the Project

Aissen, Judith L. 1991. "Relational grammar". In: Droste & Joseph (eds.) 1991, 63-102.

Asher, R. E. & T. C. Kumari 1997. *Malayalam*. London/New York: Routledge (Descriptive Grammars).

Asher, R. E. 1985. *Tamil*. London etc.: Croom Helm (Croom Helm Descriptive Grammars).

Bakker, Dik et al. 1993. *EUROTYP Guidelines* (EUROTYP Working Papers).

Basbøll, Hans 1988. "Phonological theory". In: Newmeyer (ed.) 1988a, 192-215.

Bauer, Winifred 1993. *Maori*. London/New York: Routledge (Descriptive Grammars).

Beckman, Mary E. 1988. "Phonetic theory". In: Newmeyer (ed.) 1988a, 216-238.

Behrens, Leila & Dietmar Zaefferer (eds) 2002. *Theoretical views on the lexicon: competition and convergence in currrent lexicology.* Frankfurt/Main u.a.: Lang.

Bellugi, U. & S. Fischer 1972. "A comparison of sign language and spoken language". *Cognition* 1, 173-200.

Bhat, Darbhe N. Shankara 1994. *The adjectival category. Criteria for differentiation and identification.* Amsterdam/Philadelphia: Benjamins. (Studies in Language, Companion Series, 24).

Bhatia, Tej K. 1993. *Punjabi. A cognitive descriptive grammar*. London/New York: Routledge (Descriptive Grammars).

Birk, D. B. W. 1976. *The Malakmalak language, Daly River (Western Arnhem Land)*. The Australian National University, Canberra: Department of Linguistics/Research School of Pacific Studies (Pacific Linguistics, Series B, 45).

Bloomfield, Leonard. 1933. *Language.* Chicago: University of Chicago Press. [Reprint: 1984]

Böðvarsson, Á. 1983. *Íslensk Orðabók*. Reykjavík: Bókaútgáfa Menningarsjóðs.

Borg, Albert & Marie Azzopardi-Alexander 1997. *Maltese*. London/New York: Routledge (Descriptive Grammars).

Bornemann, Eduard & Ernst Risch 1987. *Griechische Grammatik*. Frankfurt/M. etc.

Boyes Braem, Penny 1992. *Einführung in die Gebärdensprache und ihre Erforschung.* Hamburg: Signum (Internationale Arbeiten zur Gebärdensprache und Kommunikation Gehörloser, 11).

Brandner, Ellen, 1996, "The was-w-construction and clausal typing", in: Edit Doron und Shuly Wintner (eds.) IATL 3, 25-55.

Brandner, Ellen, 1996a. Specifiers and Functional Projection (Dissertation, Universität Stuttgart), Sprachwissenschaften, Band 5. Marburg:Tectum Verlag.

Brandner, Ellen & Christian Strömsdörfer 1996. *Lexical entries and argument structure in a CRG database.* University of Munich: AVG (AVG Project Papers).

Brandner, Ellen 1996a. *Basic word order and clause structure in language description.* University of Munich: AVG (AVG Project Papers).

Brandner, Ellen 1996b. *A systematics for interrogatives description.* University of Munich: AVG (AVG Project Papers).

Brandner, Ellen, Roman Pichler & Dietmar Zaefferer 1996. *Remarks on the summaries of the query on reference grammars and the query on language description and field work.* University of Munich: AVG (AVG Project Papers).

Brentari, Diane 1995. "Sign language phonology: ASL": In: Goldsmith, John (ed.). *A handbook of phonological theory*, Oxford: Blackwell, 615-639.

Brentari, Diane 1998. *A Prosodic Model of Sign Language Phonology*, Cambridge, MA: MIT Press.

Bresnan, Joan & Ronald M. Kaplan 1982. "Lexical-functional grammar: A formal system for grammatical representation". In: Bresnan (ed.) 1982a, 173-281.

Bresnan, Joan (ed.) 1982a. *The mental representation of grammatical relations*. Cambridge, Mass. etc.: MIT Press.

Bresnan, Joan 1982b. "Control and complementation". In: Bresnan (ed.) 1982a, 282-390.
Breu, Walter 2000. *Fokusaspekt und Statusaspekt im Rahmen des ILA-Modells* (Paper für den Zürcher Aspektworkshop 2000).
Brunello, Milly 1996. *Assimilation und verwandte Begriffe*. University of Munich: AVG (AVG Project Papers).
Buscha, Joachim 1995. "Zur Theorieneutralität von Referenzgrammatiken". In: Heidrun Popp (ed.) *Deutsch als Fremdsprache. An den Quellen eines Faches. Festschrift für Gerhard Helbig zum 65. Geburtstag.* München: Iudicium.
Bußmann, Hadumod ^{2}1990. *Lexikon der Sprachwissenschaft*. Stuttgart: Kröner.
Bybee, Joan et al. 1994. *The evolution of grammar. Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
Campbell, Lyle 1985. *The Pipil language of El Salvador*. Berlin, etc.: Mouton Publishers (Mouton Grammar Library, 1).
Carlson, Robert 1994. *A grammar of Supyire*. Berlin/New York: Mouton de Gruyter (Mouton Grammar Library, 14).
Chao, Yuen-ren 1965. *A grammar of spoken Chinese*. Berkeley: University of California Press.
Chierchia, Gennaro & Sally McConnell-Ginet 1990. *Meaning and grammar. An introduction to semantics*. Cambridge, Mass. etc.: MIT Press.
Childs, G. Tucker 1995. *A grammar of Kisi. A Southern Atlantic language* Berlin/New York: Mouton de Gruyter (Mouton Grammar Library, 16).
Chomsky, Noam 1957. *Syntactic structures*. The Hague: Mouton.
Chomsky, Noam. 1973. "*Conditions on transformations*". In: Stephen R. Anderson and Paul Kiparsky (eds.) *A Festschrift for Morris Halle*. New York: Holt, Rinehart, and Winston, 232-286.
Chomsky, Noam 1981. *Lectures on Government and Binding*. Dordrecht: Foris Publications (Studies in Generative Grammar, 9).
Chomsky, Noam 1982. *Some concepts and consequences of the theory of Government and Binding*. Cambridge, Mass. etc.: MIT Press (Linguistic Inquiry Monographs, 6).
Clements, George N. and Engin Sezer. (1982). Vowel and consonant disharmony in Turkish. In: Van der Hulst & Smith (eds.) 1982, 213-555.
Cole, Peter 1985. *Imbabura Quechua*. London etc.: Croom Helm (Croom Helm Descriptive Grammars).
Comrie, Bernard. 1976. *Aspect. An introduction to the study of verbal aspect and related problems*. New York, etc.: Cambridge University Press.
Comrie, Bernard. 1985. *Tense*. Cambridge: Cambridge University Press.
Comrie, Bernard 1998. "Ein Strukturrahmen für deskriptive Grammatiken". In: Zaefferer (ed.) 1998a, 7-16.
Comrie, Bernard, William Croft, Christian Lehmann & Dietmar Zaefferer 1993. "A framework for descriptive grammars". In: André Crochetière, Jean-Claude Boulanger & Conrad Ouellon (eds.) *Proceedings of the XVth International Congress of Linguists,* Quebec City, 159-170.
Comrie, Bernard & Norval Smith 1977. "Lingua Descriptive Studies: Questionnaire". *Lingua* 42, 1-72.
Cowan, J. M. 1976. *Arabic-English dictionary*. Ithaca, NY: Spoken Language Services.
Cowell, Mark W. 1964. *Reference grammar of Syrian Arabic.* Washington D.C.: Georgetown University Press.
Cruttenden, Alan 1986. *Intonation*. Cambridge: Cambridge University Press.
Crystal, David ^{3}1992. *A dictionary of linguistics and phonetics*. Oxford: Blackwell.
Dahl, Östen 1979. "Typology of sentence negation". *Linguistics* 17, 79-106.

Davidson, D. 1967. "The logical form of action sentences". In: N. Rescher (ed.). *The logic of decision and action*. University of Pittsburgh, 81-95.

Davies, John 1989. *Kobon*. London/New York: Routledge (Croom Helm Descriptive Grammars).

Derbyshire, Desmond C. 1979. *Hixkaryana*. Amsterdam: North Holland .

Dik, Simon C. (ed.) 1991. *Advances in functional grammar*. Dordrecht: Foris.

Dimmendaal, Gerrit Jan 1983. *The Turkana language*. Dordrecht: Foris.

Dixon, Robert M. (ed.) 1976. *Grammatical categories in Australian languages*. New Jersey: Humanities Press Inc. (Linguistic Series, 22, Australian Institute of Aboriginal Studies, Canberra).

Dixon, Robert M. 1982. *Where have all the adjectives gone? and other essays in semantics and syntax*. Berlin etc.: Mouton.

Dixon, Robert M. 1994. *Ergativity*. Cambridge: Cambridge University Press. (Cambridge Studies in Linguistics, 69).

Dolphyne, Florence Abena 1988. *The Akan (Twi-Fante) language. Its sound systems and tonal structure.* Ghana: Ghana University Press.

Donaldson, Bruce C. 1993. *A grammar of Afrikaans*. Berlin/New York: Mouton de Gruyter (Mouton Grammar Library, 8).

Dowty, David 1979. *Word meaning and Montague grammar*. Dordrecht: Reidel.

Dowty, David 1982. "Grammatical relations and Montague Grammar". In: Pauline Jacobson & Geoffrey K. Pullum (eds.), *The nature of syntactic representation*, 79-130.

Dowty, David 1991. "Thematic proto-roles and argument selection". *Language* 67, 547-619.

Driem, George van 1987. *A grammar of Limbu*. Berlin, etc.: Mouton de Gruyter (Mouton Grammar Library, 4).

Driem, George van 1993. *A grammar of Dumi*. Berlin/New York: Mouton de Gruyter (Mouton Grammar Library, 10).

Droste, Flip G. & John E. Joseph (eds.) 1991. *Linguistic theory and grammatical description*. Amsterdam etc.: Benjamins, 247-274.

DuFeu, Veronica 1996. *Rapanui*. London/New York: Routledge (Descriptive Grammars).

Ebbinghaus, Horst & Jens Heßmann 1994. "Formen und Funktionen von Ablesewörtern in gebärdensprachlichen Äußerungen". *Das Zeichen* 30. Hamburg: Signum.

Ebert, Karen H. (ed.). 1993a. *Studies in clause linkage. Papers from the First Köln-Zürich Workshop.* Zürich: Seminar für Allgemeine Sprachwissenschaft, Universität Zürich. (Arbeiten des Seminars für Allgemeine Sprachwissenschaft, 12).

Ebert, Karen H. 1993b. "Kiranti subordination in the South Asian areal context". In: Ebert (ed.) 1993a, 83-110.

Ebert, Karen H. 1995. "Ambiguous perfect-progressive forms across languages". In: Pier Marco Bertinetto, Valentina Bianchi, Östen Dahl & Mario Squartini (eds.) *Temporal reference, aspect and actionality.* Vol. II: *Typological perspectives.* Torino: Rosenberg & Sellier, 185-203.

Engberg-Pedersen, Elisabeth 1993. *Space in Danish sign language. The meaning and morphosyntax of use of space in a visual language.* Hamburg: Signum (Internationale Arbeiten zur Gebärdensprache und Kommunikation Gehörloser, 19).

Ethnologue: Languages of the World 2001. http://www.ethnologue.com/web.asp

Evans, Nicholas D. 1995. *A grammar of Kayardild. With historical-comparative notes on Tangkic*. Berlin/New York: Mouton de Gruyter (Mouton Grammar Library, 15).

Faraclas, Nicholas G. 1996. *Nigerian Pidgin*. London/New York: Routledge (Descriptive Grammars).

Fellbaum, Christiane. 1993. English Verbs as a Semantic Net. In: Five Papers on WordNet. www.cogsci.princeton.edu/~wn/papers/

Fellbaum, Christiane, Derek Gross & Katherine Miller 1993. *Adjectives in WordNet*. In: Five Papers on WordNet. www.cogsci.princeton.edu/~wn/papers/

Foley, William & Robert D. Van Valin 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.

Foley, William A, & Robert D.Van Valin 1979. "On the viability of the notion of 'subject' in universal grammar". In: *Proceedings of the 3rd annual meeting of the Berkeley Linguistics Society*. Berkeley, 293-320.

Fortescue, Michael 1984. *West Greenlandic*. London etc.: Croom Helm (Croom Helm Descriptive Grammars).

Fowler, Henry W. & Francis G. Fowler (eds.) ^{4}1958. *The concise Oxford dictionary of current English.* London: Oxford University Press. [Reprint of 1951 edition]

Gary, Judith Olmsted & Saad Gamal-Eldin (o.J.). *Cairene Egyptian colloquial Arabic*, London etc.: Croom Helm (Croom Helm Descriptive Grammars).

Gazdar, Gerald et al. 1985. *Generalized Phrase Structure Grammar*. Cambridge, Mass. etc.: Harvard University Press.

Givón, Talmy 1993. *English grammar. A function-based introduction*. Parts I & II. Amsterdam/Philadelphia: Benjamins.

Goldsmith, John A. 1990. *Autosegmental and metrical phonology*. Oxford: Blackwell.

Grimshaw, Jane 1990. *Argument structure*. Cambridge, Mass.: MIT Press.

Gruber, J. 1965. *Studies in lexical relations*. Doctoral dissertation. Cambridge, Mass.: MIT. [Published 1976 as *Lexical structures in syntax and semantics*. Amsterdam: North Holland.]

Gundel, Jeanette K., Nancy Hedberg & Ron Zacharsky 1993. "Cognitive status and the form of referring expressions in discourse." *Language* 69, 274-307.

Gussenhoven, Carlos & Haike Jacobs 1998. *Understanding phonology*. London, etc.: Arnold (Understanding Language Series).

Haegeman, Liliane ^{2}1994. *Introduction to Government and Binding Theory*. Oxford/Cambridge, Mass.: Blackwell (Blackwell Textbooks in Linguistics, 1).

Haider, H. 1993. *Deutsche Syntax – generativ.* Tübingen: Narr.

Haspelmath, Martin & Ekkehard König (eds.) 1995. *Converbs in cross-linguistic perspective. Structure and meaning of adverbial verb forms - adverbial participles, gerunds.* Berlin/New York: Mouton de Gruyter.

Haspelmath, Martin 1993. *A grammar of Lezgian*. Berlin/New York: Mouton de Gruyter (Mouton Grammar Library, 9).

Hawkins, Peter 1984. *Introducing phonology*. London: Hutchinson & Co. [Reprinted 1992 by Routledge]

Hayes, Bruce 1988. "Metrics and phonological theory". In: Newmeyer (ed.) 1988b, 220-249.

Heath, Jeffrey 1984. *Functional grammar of Nunggubuyu.* Canberra: Australian Institute of Aboriginal studies.

Heine, Bernd 1997. *Possession. Cognitive sources, forces, and grammaticalization*. Cambridge: Cambridge University Press (Cambridge Studies in Linguistics, 83).

Hewitt, Brian G. 1989, *Abkhaz*. In collaboration with Z. K. Khiba. London/New York: Routledge (Croom Helm Descriptive Grammars).

Hillenmeyer, Margit et al. ^{7}1996a. *Deutsche Gebärdensprache I. Grundkurs für Anfänger*. München: VHS München.

Hillenmeyer, Margit et al. ^{3}1996b. *Deutsche Gebärdensprache III. Aufbaukurs für Fortgeschrittene 2* . München: VHS München.

Hillenmeyer, Margit et al. ^{3}1996c. *Deutsche Gebärdensprache IV. Aufbaukurs für Fortgeschrittene 3.* München: VHS München.

Hillenmeyer, Margit et al. 81997. *Deutsche Gebärdensprache II. Aufbaukurs für Fortgeschrittene 1*. München: VHS München.
Hinds, John 1986. *Japanese*. London/New York: Routledge (Croom Helm Descriptive Grammars).
Hingerl, Gerhard 1996. *Eine deduktive Datenbank für die Linguistik*. Diplomarbeit. München: Institut für Informatik der LMU.
Hirst, Daniel & Albert Di Cristo (eds.) 1998. *Intonation systems. A survey of twenty languages*. Cambridge: Cambridge University Press.
Hogg, Richard & C. B. McCully 1987. *Metrical phonology: a coursebook*. Cambridge: Cambridge University Press.
Holes, Clive 1990. *Gulf Arabic*. London/New York: Routledge (Croom Helm Descriptive Grammars).
Hualde, José Ignacio 1992. *Catalan*. London/New York: Routledge (Descriptive Grammars).
Hulst, Harry van der & Jeroen van de Weijer 1995. "Vowel harmony". In: John A. Goldsmith (ed.) *The handbook of phonological theory*. Cambridge, Mass.: Blackwell, 495-534.
Humboldt, Wilhem von 1810/11. "Einleitung in das gesamte Sprachstudium". In: Hoffmann, Ludger (ed.) 1996. *Sprachwissenschaft. Ein Reader*. Berlin/New York: de Gruyter: 11-17.
Huttar, George L. & Mary L. Huttar 1994. *Ndyuka*. London/New York: Routledge.
Jackendoff, R. S. 1972. *Semantic interpretation in Generative Grammar*. Cambridge, Mass.: MIT Press.
Jacobs, Joachim 1994. "Kontra Valenz". *Fokus* 12. Trier: Wissenschaftlicher Verlag Trier.
Jacobs, Joachim, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.) 1995. *Syntax. Ein internationales Handbuch zeitgenössischer Forschung*. 2. Halbband. Berlin: de Gruyter (Handbücher zur Sprach- und Kommunikationswissenschaft, 9).
Jacobson, Pauline & Geoffrey K. Pullum (eds.) 1982, *The nature of syntactic representation*. Dordrecht etc.: Reidel.
Jacobson, Steven 1982. *A grammatical sketch of Siberian Yupik Eskimo*. Alaska: University of Alaska.
Jakobson, Roman & Linda R. Waugh 1986. *Die Lautgestalt der Sprache*. Berlin/New York: de Gruyter.
Joseph, Brian D. & Irene Philippaki-Warburton 1987. *Modern Greek*. London etc.: Croom Helm (Croom Helm Descriptive Grammars).
Kamp, Hans & Uwe Reyle 1993. *From discourse to logic. Introduction to model-theoretic semantics of natural language, formal logic and discourse representation theory.* Dordrecht etc.: Kluwer.
Karlsson, Fred 1983. *Finnish Grammar*. Provoo: Söderström.
Kaye, Jonathan D. 1982. "Harmony processes in Vata". In: Van der Hulst & Smith (eds.) 1982, 385-464.
Keating, Patricia A. 1988. "The phonology-phonetics interface". In: Newmeyer (ed.) 1988a, 281-302.
Keenan, Edward L. 1976. "Towards a universal definition of 'subject' ". In: Li (ed.) 1976, 303-334.
Keenan, Edward L. 1985. "Relative clauses". In Shopen (ed.) 1985, II, 141-170.
Keenan, Edward L. & Bernard Comrie 1977. "Noun phrase accessibility and universal grammar". *Linguistic Inquiry* 8, 63-99.
Kenstowicz, Michael 1994. *Phonology in generative grammar*. Cambridge, Mass.: Blackwell.
Kibrik, A. E. 1977. *The methodology of field investigations in linguistics*. The Hague: Mouton.
Kibrik, Aleksandr E. 1997. "Beyond subject and object: Toward a comprehensive relational typology". *Linguistic Typology* 1, 279-346.

Kiparsky, Paul 1982. *Explanation in phonology*. Dordrecht: Foris Publications.
Kiparsky, Paul 1984. "Some consequences of lexical phonology" In: Colin Ewen & John Anderson (eds.) *Phonology Yearbook* 2, 85-138.
Kiparsky, Paul 1988. "Phonological change". In: Newmeyer (ed.) 1988a, 363-416.
Kiparsky, Paul 1998. "Aspect and event structure in Vedic". In: Rajendra Singh (ed.) *The yearbook of South Asian languages and linguistics.* New Delhi / Thousand Oaks / London: Sage Publications.
Klatt, E. et al. 1983. *Langenscheidts Taschenwörterbuch Englisch*. Berlin etc.: Langenscheidt.
Kouwenberg, Silvia 1994. *A grammar of Berbice Dutch Creole*. Berlin/New York: Mouton de Gruyter (Mouton Grammar Library, 12).
Kroeger, Paul 1993. *Phrase structure and grammatical relations in Tagalog.* Stanford, Ca.: CSLI Publications.
Kyle, J. G. & B. Woll 1985. *Sign language. The study of deaf people and their language*. Cambridge: Cambridge University Press.
Labov, William 1994. *Principles of linguistic change.* Vol. I: *Internal factors*. Oxford/Cambridge, Mass.: Blackwell.
Ladd, D. Robert 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.
Ladusaw, W. & D. Dowty 1988. "Towards a nongrammatical account of thematic roles". In: W. Wilkins (ed.). *Thematic relations*. Academic Press (Syntax and Semantics, 21), 62-71.
Lambrecht, Knud 1994. *Information structure and sentence form. Topic, focus, and the mental representations of discourse referents*. Cambridge etc.: Cambridge University Press (Cambridge Studies in Linguistics, 71).
Langacker, Ronald W. 1987a. *Foundations of cognitive grammar.* Vol. 1: *Theoretical prerequisites*. Stanford, Ca.: Stanford University Press.
Langacker, Ronald W. 1987b. "Nouns and verbs". *Language* 63, 53-94.
Langacker, Ronald W. 1991. *Concept, image, and symbol. The cognitive basis of grammar*. Berlin/New York: Mouton de Gruyter (Cognitive Linguistics Research, 1).
Larson, R. K. 1988. "On the double object construction". *Linguistic Inquiry* 19, 335-391.
Laver, John 1994. *Principles of phonetics*. Cambridge: Cambridge University Press.
Leumann, Manu et al. 1963. *Lateinische Grammatik*. München: Beck.
Leumann, Manu et al. 1965. *Lateinische Syntax und Stilistik*. München: Beck.
Levin, B. & M. Rappaport Hovav 1995a. *Unaccusativity. At the syntax - lexical semantics interface*. Cambridge, Mass.: MIT Press.
Levin, B. & M. Rappaport Hovav 1995b. "The elasticity of verb meaning". In: M. Hovav & A. Mittwoch. *Proceedings of the 10th IATL Conference*, Haifa.
Levin, B. 1992. *English verb classes and alternation: A preliminary investigation*. Chicago, Ill.: Chicago University Press.
Lewandowski, Theodor ^{2}1976. *Linguistisches Wörterbuch*. Bd. 1-3. Heidelberg: Quelle & Meyer.
Lewis, G. L. 1967. *Turkish grammar.* Oxford: Clarendon.
Li, Charles N. & Sandra S. Thompson 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.
Li, Charles N. (ed.) 1976. *Subject and topic*. New York: Academic Press.
Lichtenberk, Frantisek 1983. *A grammar of Manam*. Honolulu: University of Hawaii Press.
Liddell, S. & R. Johnson 1989. "American Sign Language: The phonological base". *Sign Language Studies* 64, 195-277.
Liddicoat, Anthony 1994. *A grammar of the Norman French of the Channel Islands. The dialects of Jersey and Sark.* Berlin/New York: Mouton de Gruyter (Mouton Grammar Library, 13).

Lightner, Theodore M. 1966. "On the description of vowel and consonant harmony". *Words* 21, 244-250.
Link, G. 1979. *Montague-Grammatik*. München: Fink.
MacDonald, Lorna 1990. *A grammar of Tauya.* Berlin/New York: Mouton de Gruyter (Mouton Grammar Library, 6).
Maisch, Günter & Fritz-Helmut Wisch 1987. *Gebärdenlexikon*. Bd. 1: *Grundgebärden*. Hamburg: Verlag Hörgeschädigte Kinder.
Maisch, Günter & Fritz-Helmut Wisch 1988. *Gebärdenlexikon*. Bd. 2: *Mensch*. Hamburg: Verlag Hörgeschädigte Kinder.
Maisch, Günter & Fritz-Helmut Wisch 1989. *Gebärdenlexikon*. Bd. 3: *Mensch*. Hamburg: Verlag Hörgeschädigte Kinder.
Maisch, Günter & Fritz-Helmut Wisch 1994. *Gebärdenlexikon*. Bd. 4: *Aufbaugebärden*. Hamburg: Verlag Hörgeschädigte Kinder.
Mallinson, Graham 1986. *Rumanian*. London etc.: Croom Helm (Croom Helm Descriptive Grammars).
Mally, Gertrud 1993. *Redewendungen der Deutschen Gebärdensprache: Münchner Dialekt*. Hamburg: Signum.
Marantz, Alec 1984. *On the nature of grammatical relations*. Cambridge, Mass. etc.: MIT Press (Linguistic Inquiry Monographs, 10).
Marantz, Alec 1995. "The Minimalist Program". In: Gert Webelhuth (ed.). *Government and Binding theory and the Minimalist Program*. *Principles and parameters in syntactic theory*. Oxford/Cambridge, Mass.: Blackwell (Generative Syntax, 1), 349-382.
Martin, Samuel E. 1975. *A reference grammar of Japanese.* New Haven: Yale University Press.
Martin, Samuel E. 1992. *A reference grammar of Korean.* Tuttle: Rutland.
Matisoff, James A. 1973. *The grammar of Lahu*. Berkeley: University of California Press (University of California Publications in Linguistics, 75).
McCawley, J. D. 1973. *Grammar and meaning*. Tokyo: Taishukan Publishing Co.
Mel'chuk, Igor A. & Leo Wanner 1994. "Lexical co-occurence and lexical inheritance. Emotion lexemes in German: A lexicographic case study." *Lexikos* 4, 86-161.
Mel'chuk, Igor A. 1987. "Un affixe derivationnel et un phrasème syntaxique du russe moderne. Essai de description formelle". *Rev. Étud. slaves* LIX/3, 631-648.
Mel'chuk, Igor A. 1990. "Slovoobrazovanie v lingvisticheskix modeljax tipa "Smysl <-> Tekst". Predvaritel'nye zamechanija" ('Word formation in linguistic models of the "Meaning <-> Text" type. Preliminary considerations'). In: *Metody formalne v opisie języków slowianskich*. Festschrift Jurij Apresjan. Bialystok, 47-74.
Mel'chuk, Igor A. 1993. "The inflectional category of voice: towards a more rigorous definition." In: Bernard Comrie & Maria Polinsky (eds.) *Causatives and transitivity*, Amsterdam: Benjamins, 1-46.
Mel'chuk, J. 1984. *Dictionnaire explicatif et combinatoire au français contemporain*. Les presses de l'Université Montréal.
Merkle, Ludwig ^{5}1993. *Bairische Grammatik* (1st ed. 1975). München: Hugendubel.
Merlan, Francesca 1989. *Mangarayi*. London/New York: Routledge (Croom Helm Descriptive Grammars).
Merlan, Francesca C. 1994. *A grammar of Wardaman. A language of the northern territory of Australia*. Berlin/New York: Mouton de Gruyter (Mouton Grammar Library, 11).
Miller, George A. et al. 1993. *Introduction to WordNet: An on-line lexical database.* In: Five Papers on WordNet. www.cogsci.princeton.edu/~wn/papers/
Nedjalkov, Igor 1997. *Evenki*. London/New York: Routledge (Descriptive Grammars).
Neukom, Lukas 2001. *Santali*. München: Lincom (Languages of the World/Materials, 323)

Newmeyer, Frederick J. 1988a. *Linguistics: The Cambridge survey*. Vol.I.: *Linguistic theory: Foundations.* Cambridge: Cambridge University Press.

Newmeyer, Frederick J. 1988b. *Linguistics: The Cambridge survey*. Vol.II: *Linguistic theory: Extensions and implications*. Cambridge: Cambridge University Press.

Nickles, Matthias 2001. "*Systematics* - Ein XML-basiertes Internet-Datenbanksystem für klassifikationsgestütze Sprachbeschreibungen" CIS-Bericht-01-129.

Noonan, Michael 1992. *A grammar of Lango*. Berlin/New York: Mouton de Gruyter (Mouton Grammar Library, 7).

Oppenrieder, Wilhelm 1991. *Von Subjekten, Sätzen und Subjektsätzen*. Tübingen: Niemeyer.

Padden, Carol A. 1988. "Grammatical theory and signed languages". In: Newmeyer (ed.) 1988b, 250-266.

Paducheva, E. V. 1989. "K poiskam invarianta v znachenii glagol'nyx vidov: vid i leksicheskoe znachenie glagola" ('On the search for invariants in the meaning of verbal aspects: aspect and the lexical meaning of a verb'). NTI ser. 2, *Informacionnye Processy i Sitemy* 12, 24-31.

Paducheva, E. V. 1993. "O tak nazyvaemoj konkurencii sovershennogo i nesovershennogo vida v russkom jazyke" ('On the so-called competition of perfective and imperfective aspects in Russian'). *Wiener Slawistischer Almanach* 31, 259-273.

Pandharipande, Rajeshwari V. 1997. *Marathi*. London/New York: Routledge (Descriptive Grammars).

Papaspyrou, Chrissostomos 1990. *Gebärdensprache und universelle Sprachtheorie. Versuch einer vergleichenden generativ-transformationellen Interpretation von Gebärden- und Lautsprache sowie der Entwurf einer Gebärdenschrift*. Hamburg: Signum (Internationale Arbeiten zur Gebärdensprache und Kommunikation Gehörloser, 8).

Payne, John R. 1985. "Negation". In Shopen (ed.) 1985, I, 197-242.

Perlmutter, David M. 1982. "Syntactic representation, syntactic levels, and the notion of subject". In: Jacobson & Pullum (eds.) 1982, 283-340.

Peterson, John (in prep.). "Nepali converbs: A holistic approach".

Peterson, John (in press). "Evidentials, inferentials and mirativity in Nepali". In: Balthasar Bickel (ed.). *Person and evidence in Himalayan languages. Special Issue of Linguistics of the Tibeto-Burman Area.* (Fall 2000)

Pichler, Roman & Christian Strömsdörfer 1995. *Summary of the query on reference grammars*. University of Munich: AVG (AVG Project Papers).

Pichler, Roman & Dietmar Zaefferer 1996. *Remarks on the results of the queries*. University of Munich: AVG (AVG Project Papers).

Pichler, Roman & Vladimir Tourovski 1994. *Criteria for reference grammars.* University of Munich: AVG (AVG Project Papers).

Pichler, Roman 1994a. *Comments on the FDG Sursurunga stack.* University of Munich: AVG (AVG Project Papers).

Pichler, Roman 1994b. *Notes on hyperprinciples.* University of Munich: AVG (AVG Project Papers).

Pichler, Roman 1994c. *Dokumentation der Münchner Software-Entwicklung.* University of Munich: AVG (AVG Project Papers).

Pichler, Roman 1995. *Referenzgrammatikkriterien*. University of Munich: AVG (AVG Project Papers).

Pichler, Roman 1996. *Remarks on the notion subject*. University of Munich: AVG (AVG Project Papers).

Pichler, Roman 1996. *Summary of the query on language description and field work*. University of Munich: AVG (AVG Project Papers).

Pinker, S. 1989. *Learnability and cognition: The acquisition of argument structure*. Cambridge, Mass.: MIT Press.

Poizner, Howard, Ursula Bellugi & Edward Klima 1990. *Was die Hände über das Gehirn verraten*. Hamburg: Signum.
Pollard, C. & I. Sag 1994. *Head-driven phrase structure grammar.* Stanford/Chicago.
Pollard, Carl J. & Ivan A. Sag 1987. *Information-based syntax and semantics*. Stanford, Ca.: Center for the Study of Language and Information.
Pollard, Carl J. & Ivan A. Sag 1994. *Head-Driven Phrase Structure Grammar*. Stanford, Ca.: Center for the Study of Language and Information.
Pompino-Marschall, Bernd 1995. *Einführung in die Phonetik*. Berlin/New York: de Gruyter.
Press, Ian 1986. *A grammar of modern Breton*. Berlin etc.: Mouton de Gruyter (Mouton Grammar Library, 2).
Prillwitz, Siegmund & Guido H. G. Joachim 1993. *International bibliography of sign language*. Hamburg: Signum (Internationale Arbeiten zur Gebärdensprache und Kommunikation Gehörloser, 21).
Prillwitz, Siegmund et al. 1985. *Skizzen zu einer Grammatik der Deutschen Gebärdensprache.* Hamburg: Forschungsstelle Deutsche Gebärdensprache.
Prillwitz, Siegmund et al. 1989. *HamNoSys. Version 2.0. Hamburger Notationssystem für Gebärdensprachen. Eine Einführung.* Hamburg: Signum (Internationale Arbeiten zur Gebärdensprache und Kommunikation Gehörloser, 6).
Primus, Beatrice 1987. *Grammatische Hierarchien.* München: Fink.
Primus, Beatrice 1995. "Syntactic relations". In: Jacobs et al. (eds.) 1995, 686-704.
Pullum, Geoffrey K. & Arnold M. Zwicky 1988. "The syntax-phonology interface". In: Newmeyer (ed.) 1988a, 255-280.
Quirk, Randolph et al. 1985. *A comprehensive grammar of the English language.* London: Longman.
Reichenbach, Hans 1947. *Elements of Symbolic Logic*. London: Macmillan.
Rennison, John R. 1997. *Koromfe*. London/New York: Routledge (Descriptive Grammars).
Rice, Keren 1989. *A grammar of Slave*. Berlin/New York: Mouton de Gruyter (Mouton Grammar Library, 5).
Rijkhoff, Jan 1992. *The noun phrase. A typological study of its form and structure*. Amsterdam: Diss. University of Amsterdam.
Ringen, Catherine O. 1988. *Vowel harmony. Theoretical implications*. New York/London: Garland.
Roberts, John R. 1987. *Amele*. London etc.: Croom Helm (Croom Helm Descriptive Grammars).
Saltarelli, Mario 1988. *Basque*. London etc.: Croom Helm (Croom Helm Descriptive Grammars).
Sampson, Geoffrey 1985. *Writing systems*. Stanford/Ca.: Stanford University Press.
Sasse, Hans-Jürgen 1988. "Der irokesische Sprachtyp". *Zeitschrift für Sprachwissenschaft* 7/2, 173-213.
Sasse, Hans-Jürgen (ed.) 1991a. *Aspektsysteme*. Köln: Institut für Sprachwissenschaft, Universität Köln (Arbeitspapier Nr. 14 (Neue Folge)).
Sasse, Hans-Jürgen 1991b. "Aspekttheorie". In: Sasse (ed.) 1991a, 1-35.
Sasse, Hans-Jürgen. 1991c. "Aspect and aktionsart. A reconciliation". In: Carl Vetters & Willy Vandeweghe (eds.) *Perspectives on aspect and aktionsart, Belgian Journal of Linguistics* 6, 31-45.
Schachter, Paul 1976. "The subject in Philippine languages: Topic, actor, actor-topic, or none of the above". In: Li (ed.) 1976, 491-518.
Schachter, Paul 1987. "Tagalog". In: Bernard Comrie (ed.) *The world's major languages*, New York: Oxford University Press, 936-958.
Schaub, Willi 1985. *Babungo*. London etc.: Croom Helm (Croom Helm Descriptive Grammars).

Schütz, Heribert & Dietmar Zaefferer 1996. "Eine linguistische Wissensbank", Forschungsbericht PMS-FB-1996-17. München: Institut für Informatik der LMU.

Schwyzer, Eduard 1950. *Griechische Grammatik*. Bd. 2: *Syntax und Semantik*. München.

Senft, Gunter 1986. *Kilivila. The language of the Trobriand islanders*. Berlin, etc: Mouton de Gruyter (Mouton Grammar Library, 3).

SGB (Schweizerischer Gehörlosenbund Region Deutschschweiz) 1992. *Illustration der Grundgebärden der deutschschweizerischen Gebärdensprache (DSGS)*. Zürich.

Shopen, Timothy (ed.) 1985, *Language typology and syntactic description.* 3 vols. Cambridge: Cambridge University Press.

Smith, Carlota S. 1991. *The parameter of aspect*. Dordrecht, etc.: Kluwer (Studies in Linguistics and Philosophy, 43).

Smith, Norval 1992. "Vowel harmony and consonant harmony". In: William Bright (ed.). *International encyclopedia of linguistics*. Vol. 4. New York/Oxford: Oxford University Press, 232-234.

Sohn, Ho-min 1994. *Korean*. London/New York: Routledge (Descriptive Grammars)

Sridhar, Shikaripur N. 1990. *Kannada*. London/New York: Routledge (Descriptive Grammars).

Stokoe, William C. 1960. *Sign language structure: An outline of the visual communication system of the American deaf*. University of Buffalo (Occasional Paper 8).

Strömqvist, Sven 1994. "Spracherwerb, Sprachverarbeitung und Sprachwandel". *Das Zeichen* 29. Hamburg: Signum.

Strömsdörfer, Christian 1994a. *Softwareentwicklung & Datenbanken*. University of Munich: AVG (AVG Project Papers).

Strömsdörfer, Christian 1994b. *SHOEBOX: Übersicht über die Programmerkmale*. University of Munich: AVG (AVG Project Papers).

Strömsdörfer, Christian 1994c. *GRAMMARS: Overview of the program features.* University of Munich: AVG (AVG Project Papers).

Strömsdörfer, Christian 1994d. *Grammatische Deskriptoren für Gebärdensprachen*. University of Munich: AVG (AVG Project Papers).

Strömsdörfer, Christian 1994e. *Entwicklungsdokumentation zu GRAMMARS*. University of Munich: AVG (AVG Project Papers).

Strömsdörfer, Christian 1995. *A hierarchical system of descriptive terms for the phonology of natural human languages*. University of Munich: AVG (AVG Project Papers).

Strömsdörfer, Christian 1996. *Grammars. Beta version 0.004*. University of Munich: AVG (AVG Project Papers).

Strömsdörfer, Christian, & Theo Vennemannn, 1995, "Ziele der syntaktischen Typologie", in: Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, Theo Vennemann (Hgg.), Syntax. Ein internationales Handbuch zeitgenössischer Forschung, 2. Halbband. Berlin: de Gruyter 1995, 1031-1043.

Strömsdörfer, Christian, & Theo Vennemannn, 1995a, "Das Verhältnis des Syntaxwandels zur Theorie der Sprachzustände", in: Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, Theo Vennemann (Hgg.), Syntax. Ein internationales Handbuch zeitgenössischer Forschung, 2. Halbband. Berlin: de Gruyter 1995, 1126-1135.

Sulkala, Helena & Merja Karjalainen 1992. *Finnish*. London/New York: Routledge (Descriptive Grammars).

Takizawa, Tsuyoshi & Christian Strömsdörfer 1994. *Einscannen existierender Grammatiken: Manam*. University of Munich: AVG (AVG Project Papers).

Takizawa, Tsuyoshi 1996. *Explikation von Phonem, Silbe und Mora*. University of Munich: AVG (AVG Project Papers).

Talmy, Leonard 1985. "Lexicalization patterns: Semantic structure in lexical forms". In: Shopen (ed.) 1985, III, 57-149.

Taylor, Charles 1985. *Nkore-Kiga*. London etc.: Croom Helm (Croom Helm Descriptive Grammars).
Taylor, John R. 1996. *Possessives in English. An exploration in cognitive grammar*. Oxford: Clarendon.
Thomas, David D. 1971. *Chrau grammar*. University of Hawaii Press (Oceanic Linguistics, Special Publication No. 7).
Tiersma, Pieter Meijes 1985. *Frisian reference grammar*. Amsterdam: Foris.
Tourovski, Vladimir & Christian Strömsdörfer 1994. *AVG, Grammars, Shoebox, Rook: A comparison of four grammar authoring systems.* University of Munich: AVG (AVG Project Papers).
Tourovski, Vladimir 1993. *Der Stellenwert der Restriktivität in ihrem grammatischen Umfeld.* University of Munich: AVG (AVG Project Papers).
Tourovski, Vladimir 1994. *Rook: Overview of the program features.* University of Munich: AVG (AVG Project Papers).
Tourovski, Vladimir 1995. *Formal & functional parameters: A final report*. University of Munich: AVG (AVG Project Papers).
Trask, R. L. 1996. *A dictionary of phonetics and phonology*. London/New York: Routledge.
Underhill, Robert 1976. *Turkish grammar.* Cambridge, Mass.: MIT Press.
Van der Hulst, Harry & Norval Smith (eds.) 1982. *The structure of phonological representations*. Part II. Dordrecht: Foris Publications.
Van Valin, Robert D. 1993. "A synopsis of role and reference grammar". In: R. D. Foley Jr. (ed.). *Advances in role and reference grammar*. Amsterdam: Benjamins.
Van Valin, Robert D. & Randy J. LaPolla 1997. *Syntax. Structure, meaning and function*. Cambridge: Cambridge University Press.
Vendler, Z. 1967. *Linguistics in philosophy*. Ithaca, NY: Cornell University Press.
Vennemann, Theo 1982. "Grammatical relations". In: The Linguistic Society of Korea (ed.). *Linguistics in the morning calm.* Selected papers from SICOL 1981. Seoul: Hanshin, 233-267.
Vennemann, Theo 1986. *Neuere Entwicklungen in der Phonologie*. Berlin etc.: Mouton de Gruyter.
Vennemann, Theo 1988. *Preference laws for syllable structure and the explanation of sound change*. Berlin etc.: Mouton de Gruyter.
Vennemann, Theo 1994. *Universale phonologische Strukturen und Prozesse*. Tübingen: Niemeyer.
Wali, Kashi & Omkar N. Koul 1997. *Kashmiri. A cognitive descriptive grammar*. London/New York: Routledge (Descriptive Grammars).
Willet, Thomas 1988. "A cross-linguistic survey of the grammaticization of evidentiality". *Studies in Language* 12/1, 51-97.
Williams, E. 1994. *Thematic structure in syntax*. Cambridge, Mass.: MIT Press.
Wisch, Fritz-Helmut 1994. "Wert und Unwert des Gebärdenlexikons". *Das Zeichen* 27. Hamburg: Signum.
Yip, M., J. Maling & R. Jackendoff 1987. "Case in tiers". *Language* 63, 217-250.
Zaefferer, Dietmar 1994. *Ein System funktionsbezogener Beschreibungskategorien für deskriptive Grammatiken.* University of Munich: AVG (AVG Project Papers).
Zaefferer, Dietmar et al. 1994a. *AVG: A critical assessment of the current system*. University of Munich: AVG (AVG Project Papers).
Zaefferer, Dietmar et al. 1994b. *Formal and functional descriptive categories.* University of Munich: AVG (AVG Project Papers).
Zaefferer, Diemar 1995a. "Comparative surveys in syntactic typology: A critical guide". In: Jacobs et al. (ed.) 1995, 2. Halbband, 1109-1115.

Zaefferer, Dietmar 1995b. *Options for a cross-linguistic reference grammar database*. University of Munich: AVG (AVG Project Papers).

Zaefferer, Dietmar 1995c. "Typologie der Lautsysteme". Unveröffentlichtes Vorlesungsskript. Universität München.

Zaefferer, Dietmar 1997. "Neue Technologien in der Sprachbeschreibung. Der Paradigmenwechsel von linearen P-Grammatiken zu vernetzten E-Grammatiken", In: Zeitschrift für Literaturwissenschaft und Linguistik 106, 76-82.

Zaefferer, Dietmar 1998a. "Description and comparison. How much theory do we need to make language descriptions globally comparable?". In: Bernard Caron, ed. *Proceedings of the XVIth International Congress of Linguists*, Paris.

Zaefferer, Dietmar, (ed.) 1998b. *Deskriptive Grammatik und allgemeiner Sprachvergleich*. Tübingen: Niemeyer.

Zaefferer, Dietmar 2001. "Modale Kategorien", In: Martin Haspelmath, Ekkehart König, Wulf Oesterreicher, Wolfgang Raible (Hgg.), *Sprachtypologie und sprachliche Universalien* (HSK 20.1). Berlin: de Gruyter 2001, 784-816.

Zaefferer, Dietmar 2001a. "Deconstructing a classical classification. A typological look at Searle's concept of illocution type" In: *Revue Internationale de Philososphie 216 Searle - with his replies* (2/2001), 209-225.

Zubizarreta, Maria L. 1982. "The formal interaction of harmony and accent: the tone pattern of Japanese". In: Van der Hulst & Smith (eds.) 1982, 159-212.

Zubizarreta, Maria L. 1987. *Levels of representation in the lexicon and in the syntax*. Dordrecht: Foris.