Brigitte Endres-Niggemeyer
Fachhochschule, FBID, Hannover

Summarising Text for Intelligent Communication

Results of the Dagstuhl Seminar

Brigitte Endres-Niggemeyer is a professor of knowledge representation and computerized subject analysis at the Polytechnic of Hannover. In her research, she concentrates on AI and Cognitive science approaches to information processes. Currently, she is developing an empirically founded expert simulation system of professional summarizing, i.e., of abstracting, indexing, and classification.

Endres-Niggemeyer, B.: **Summarising text for intelligent communication.** Results of the Dagstuhl Seminar.
Knowl.Org. 21(1994)No.4, p. 213-223

As a result of the transition to full-text storage, multimedia and networking, information systems are becoming more efficient but at the same time more difficult to use, in particular because users are confronted with information volumes that increasingly exceed individual processing capacities. Consequently, there is an increase in the demand for user aids such as summarising techniques. Against this background, the interdisciplinary Dagstuhl Seminar *Summarising Text for Intelligent Communication*,
(Dec.1993) outlined the academic state of the art with regard to summarising (abstracting) and proposed future directions for research and system development. Research is currently shifting its attention from text summarising to summarising states of affairs. Recycling solutions are put forward in order to satisfy short-term needs for summarisation products. In the medium and long term, it is necessary to devise concepts and methods of intelligent summarising which have a better formal and empirical grounding and a more modular organisation. (Author)

1. Summarising - a Basic Cognitive Skill

In summarising (abstracting), a body of information, often represented by a text, is reduced in size and content to its important points (ALTE92). To do this, it is essential to analyse the input information, e.g. through understanding a source text or interpreting sensations, to rework and reduce the resulting mental representation and to produce the summary. Fig. 1 shows the basic organisation of this summarising process. The sub-processes of analysis, condensation and generation proceed from a source text to a target text, in this case the summary. Other interpretations of the summarising process are possible. For example, text does not have to be taken as the only type of source and target information (c.f. below Fig. 6 and Fig. 10).

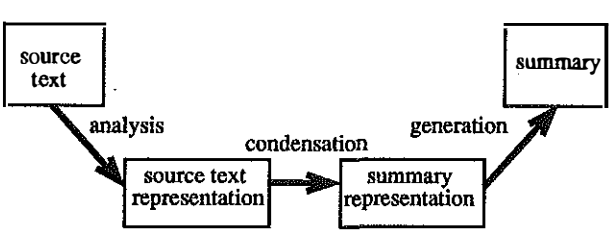

Fig. 1: Summarising texts

Of interest from an academic point of view is summarising as a basic cognitive skill. The ability to summarise allows us to filter out the most important information from the vast amounts we are confronted with each day. Summaries have a practical significance in many contexts of everyday life, including information systems. In such systems, a large number of abstracts are required as a way of providing brief information about the documents available in the system. The need for summarising skills increases with the number of information systems and their functional development. Hence the question as to how to generate higher quality summaries with automatic systems is becoming increasingly significant.

2. From 1958 till Today

2.1 Automatic abstracting: sentence extraction

Even in the early days of data processing, pragmatically-oriented attempts at automatic abstracting were made (overview see BORK68). The most common method was (and to some extent still is) sentence extraction. On the basis of core words, syntactic structures or "indicator phrases" with which the author announces important propositions, those sentences are selected from the original document that are likely to contain central elements of information. By stringing them together, an extract of the original is obtained, which can assume the function of an abstract. Fig. 2 shows the first auto-abstract in LUHN58.

Exhibit I

Source: The Scientific American. Vol.196,No.2, 68-94, February 1958
Title: Messengers of the Nervous System
Author: Amodeo S. Marazzi

Editor's sub-heading: The internal communication of the body is mediated by chemicals as well as by nerve impulses. Study of their interaction has developed important leads to the understanding and therapy of mental illness.

Auto-Abstract

It seems reasonable to credit the single-celled organisms also with a system of chemical communication by diffusion of stimulating substances through the cell, and these correspond to the chemical messengers (e.g. hormones) that carry stimuli from cell to cell in the more complex organisms. (7.0)

Finally, in the vertebrate animals there are special glands (e.g., the adrenals) for producing chemical messengers, and the nervous and chemical communication systems are intertwined: for instance, release of adrenalin by the adrenal gland is subject to control both by nerve impulses and by chemicals brought to the gland by the blood. (6.4)

The experiments clearly demonstrated that acetylcholine (and related substances) and adrenalin (and its relatives) exert opposing actions which maintain a balanced regulation of the transmissions of nerve impulses. (6.3)

It is reasonable to suppose that the tranquilizing drugs counteract the inhibitory effect of excessive adrenalin or serotonin of some related inhibitor in the human nervous system. (7.3)

Fig. 2: The first autoabstract in LUHN58

Machine-produced extracts are unsatisfactory because of their lack of textual coherence. Attempts were made to

smooth out this deficiency. However, concurrently the need was expressed for more intelligent processes:

"future automatic extracting methods must take into account syntactic and semantic characteristics of the language and the text: they cannot rely on gross statistical evidence." (EDMU69)

2.2 Human summarising in cognition science research

Cognitionscience theories (overview in FAYO91, most comprehensive individual theory in KINT83) describe the summarisation of texts with empirical-experimental methods. One popular experimental approach is presented to the test subjects as a memory test. This establishes which statements from what is often a relatively short text are memorised and repeated as a summary (SCHNO81 among others). Summarising appears here as a process of text understanding (encoding), followed by a corresponding decoding process, in which what has been remembered from the original text is reproduced from memory (c.f. Fig. 3).

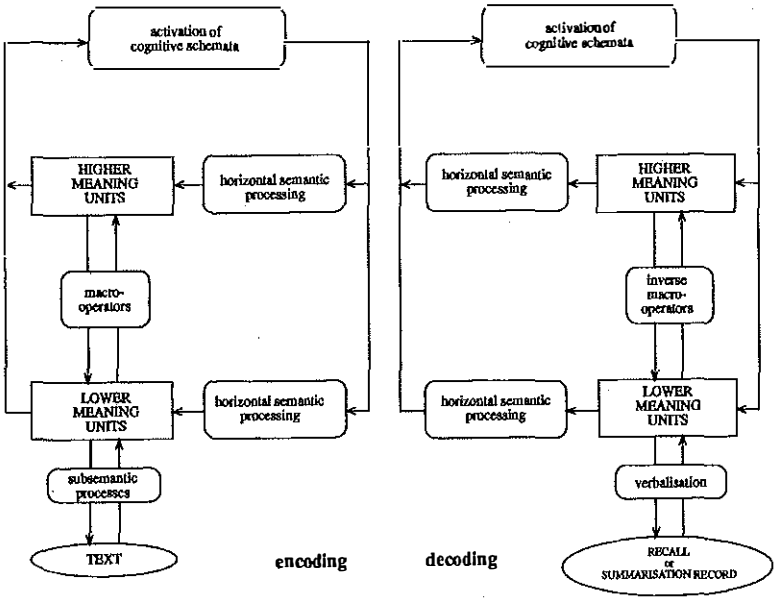

Fig. 3: Summarising teaching texts according to SCHNO81

Since longer source texts usually have a text type-specific meaning structure (macrostructure, superstructure), which offers a proven convention for the presentation of the relevant objects (e.g. an empirical study), the global structure of the text here becomes an important auxiliary factor in summarising (DIJK80, KINT83). The aim is to give a brief description of what is contained in the original, whilst respecting the macro- and the superstructure of the original. What is important in this context is that summarising is not understood as a simple selection and abstraction of statements from the original, for in the process of summarising, statements are reconstructed in an all-too-obvious way (RICK89). KINT83 already place summarising in the context of often dialogical everyday discourse. They explain discourse understanding and summarising as a cognitive act that combines linguistic knowledge and domain knowledge in a task-oriented manner.

2.3 Systems following the advent of cognition science

Early works concerned with automatic abstracting dealt with practical goals. The methods proceeded from the linguistic surface of documents. By contrast, approaches to automatic summarising inspired by cognition science (overview in ALTE92, KUHL89a) are oriented more towards human processes (c.f. Fig. 3): human beings first of all understand the text. Coherence gaps are filled from our own prior knowledge. What has been understood is now represented in memory. The memory representation serves as a basis for a textual summary. One central instrument of understanding and summarising are cognitive schemata. In the case of descriptive texts such schemata result, for example, from standard sequences of the plot ("plot units" - LEHN82a) or from genre-specific text structures (RUME77). However, they can also represent domain knowledge, e.g. knowledge about a goldfish pond. In their system design, the authors of SUSY (FUM84) refer to the cognitive science theory of text understanding and summarising as described in KINT83. SUSY generates its summaries accordingly by "understanding" the text and weighting its meaning components according to their relevance. The elements of meaning with the highest relevance values are included in the summary (c.f. Fig. 4).

With short descriptive and general texts, summarising (or information extraction - c.f. MUC-3, MUC-4) with the help of an event schema, as first realised in the FRUMP system (DEJO82), has so far proved to be the apparently most successful solution. What is essential here, is that the schemata include prior knowledge about the meaning structure of the stories to be analysed. It is this knowledge that allows us to examine the input texts specifically with regard to their most important aspects (events, actors, etc.). Thus schema- and expectation-oriented processes circumnavigate the obstacle of completely understanding a text.

A combination of expectation-driven processing and partial text understanding is the success principle behind the SCISOR system (RAU89, JACO93). The authors of FASTUS (APPE93) also come to the conclusion that in order to evaluate and summarise real documents a full understanding of the text is not always necessary, an active interpretation based on prior knowledge certainly is, however. Professional abstractors take the same view (ENDR92).

Research into summarising reveals a particularly high deficit where the need is most serious: in the case of long texts, which in a professional information context is more likely to be the rule than the exception. TOPIC (HAHN90, KUHL 89b) summarises a few texts that are as long as 2000 words. A thesaurus-like knowledge base and partial text parsing based on a theme-rheme approach are used to represent text components through nominal concepts with attributes, which are represented as frames.

The nominal structures are either displayed graphically as networks or texted with templates to produce indicative abstracts.

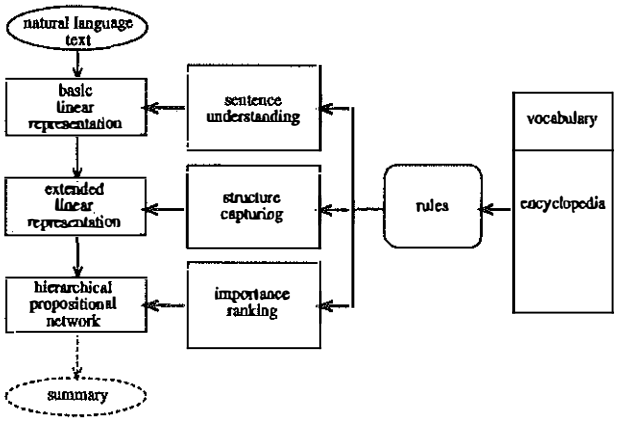

Fig. 4: The SUSY architecture (Source: FUM84)

2.4 The further development of information systems

As the application environment for automatic summarising, information systems are particularly important. They have remained stable in terms of their basic technical and organisational characteristics for a relatively long time. Currently, they are undergoing a thorough restructuring process as a result of new technical developments. The need for and the character of summarising skills changes along with the system environment:

- Full-text information systems confront their users with larger volumes of text than before. Even in order to cope with the increasing volume of information with the same processing capacity, users require more aids, including more, and higher quality, summaries. Abstracts are much more practically useful in cases where they have been tailored to meet the actual demand.

- Multimedia information systems enable automatic summarising systems to perform what for human beings in everyday summarising situations is a matter of course: summarise bodies of information that are stored on various media (written text, graphics, animated pictures, etc.). As soon as information systems place other media alongside the written text, it will be natural in each case to present long versions and summaries of documents through the most appropriate media.

- Information networks pave the way to distributed information systems which by far exceed the volumes an individual data base supplier can provide. At the same time, information within the network is heterogeneous, shows little structure in terms of information methods and it is also active, since it flows towards a user, rather than waiting to be retrieved from the data base. Thus the information problem shifts from one of acquisition to a problem of information filtering (BELK92), which involves an active search, but also selection and rejection of information junk (c.f. Fig. 5). For this, users need instruments which allow them to review information in the desired (also reduced) degree of detail, i.e. summaries.

Overall, what is required as a result of the increased information volume is a greater summarising capacity with a higher or more differentiated effectiveness (i.e. more human and machine intelligence). New forms of summarising can assume new functions, for example, by specifically supporting the assimilation of new knowledge or making use of new forms of presentation, such as hypermedia.

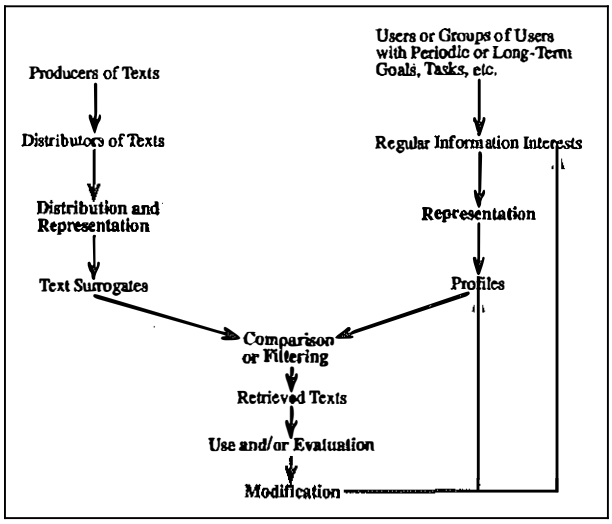

Fig. 5: Information filtering systems (Source: BELK92)

3. A selection of current approaches: the Dagstuhl seminar

The situation described above is reason enough to rethink current research into human and automatic summarising. At Schloss Dagstuhl, academics and practitioners from all interested disciplines came together in order to pool their knowledge and consider what direction the development of intelligent summarising systems might take.

At the opening session of the seminar, *Wolfgang Wahlster* mentioned a number of deficits in the scientific penetration of summarising. These include in particular:

- the definition of an (optimal) summary
- the assessment of summaries as compared to extracts
- the utilisation of textual, multimodal and interactive summaries
- the representations that are needed for different summaries
- the effects of limited resources when summarising, and economical summarising methods
- the evaluation of summaries

Karen Sparck Jones presented a description grid for summarising using the factors input, purpose and style as the main dimensions.

Both speakers agreed that the research situation is unsatisfactory. Since research efforts are spread among different disciplines, there is also a lack of a general overview of current approaches. The contributions of the seminar participants were a step towards filling this gap and are described in the following. A more detailed description can be found in ENDR94.

3.1 "No cognitive system summarises texts"

No cognitive system, according to *Gerhard Strube*, summarises texts. Instead, it assimilates a state of affairs from a source, which may equally be a text, a conglomorate of materials or direct observation, and presents it in brief. If a text form is used to reconstruct a state of affairs, the summary appears as a secondary text derived from the original (c.f. Fig. 1).

When states of affairs are presented as linguistic texts, the relatively familiar form of information organisation within the text can facilitate summarising, since in certain text types it is known where specific information is introduced. Where form is of no help, we have to rely on the processing of the described knowledge itself in order to produce a summary.

If cognitive systems - for example human beings - summarise states of affairs that are represented through texts or other material (e.g. through direct observation), computer systems should, for reasons of parity, be accorded the same field of action, since they are supposed - either wholly or in part - to solve the same tasks. Once information systems cease to insist on written texts as the form of presentation, it not only becomes more appropriate but also more expedient in the case of automatic summarising to loosen the link to the written text and introduce other forms of presentation. Researchers are already developing systems which deal with states of affairs in different representation forms.

Fig. 6 proposes an integrated cognitive view of summarising which takes both the automatic and the human aspect of the process into account. The summarising proceeds from a state of affairs represented in any form and presents a product that can also have any form. The outline also gives an overview of the methods with which research has approximated the subtasks of summarising. The most commonly used methods are stated first in each case.

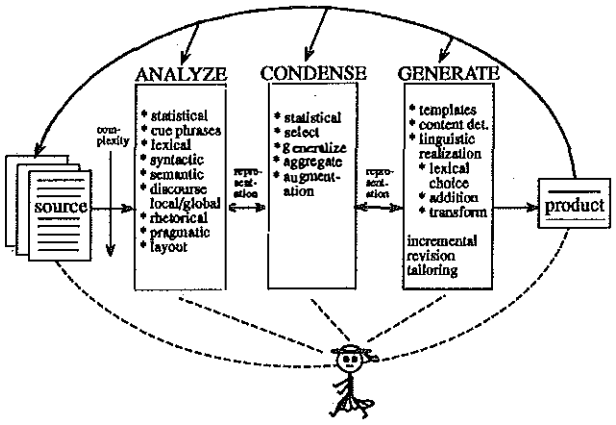

Fig. 6: Summarising - an integrated cognitive view of research

3.2 Summarising as text generation

Text generation systems proceed from a description of states of affairs that typically is not a linguistic text; often information from a data base is "texted". Text generation systems thus behave in the same way as the cognitive systems described above. Since current efforts concentrate on the generation of short texts, the result will always display central features of an abstract or summary. The generation therefore appears to be inherently summarising. Text generation methods have summarising side-effects which can also be made use of in other contexts.

Summarising simulated sequences of events

In order to summarise simulated sequences (as in the case of military manoeuvres, for example), *Mark Maybury* suggested techniques with different backgrounds:

- the use of the saliency of occurrences, or their attributes or meaning roles
- the semantic integration of several separate events in one overall event
- the linguistic integration of several separate statements in one overall statement (example: A and B fired a missile at the same time)
- recognising important events by the number of links to other events
- presentation techniques including deliberate choice of the most easily perceived medium

The summary is to include various types of information, including measurement values (c.f. Fig. 7). Together with linguistic and knowledge-based methods, traditional data reduction techniques are used. Metrics are used to first of all select from hundreds or thousands of individual occurrences those which are to be considered. The metrics take the frequency, uniqueness and domain-related saliency of the represented event as a basis. For example, they ensure that missiles that hit their target are valued higher in spite of their relative infrequency than the more frequent number of missile firings (c.f. Fig. 7). When generating the summary text, a theme sentence is first constructed, followed by the selected events in chronological sequence. User stereotypes (e.g. of logisticians) can be taken into account. The result is a summarising report consisting of several paragraphs.

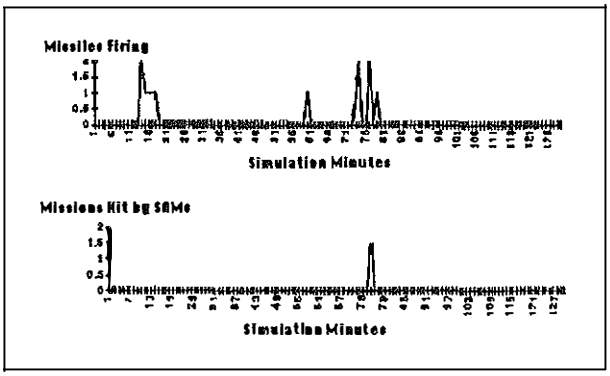

Fig. 7: Missile firings and missiles that hit their target (simulation time: 129 minutes) (Source: Maybury in ENDR94)

STREAK: summaries of baseball matches

Kathleen McKeown described two systems that generate summaries from quantitive and qualitative data: re-

ports about baseball matches and reports about the procedure that engineers propose for adapting the telephone network. The methods can be illustrated with the help of the more popular example of the baseball matches (the STREAK system).

The input is the data describing a match. An empirical study of the short descriptions of baseball matches produced by news agencies revealed three main points:

- The sentences are long and complex, because as much information as possible has to be packed into one sentence.
- In baseball reports, a distinction must be made between obligatory and facultative information. Facultative information is integrated into the text if and where the opportunity arises.
- Certain information appears in a specific position and other information can appear anywhere within the body of text. The latter accounts for 40% of the information content.

When generating summaries, which resemble the flash reports provided by agencies, the following problems have to be solved:

- It must be decided what information is to be included in the summary and what is to be left out.
- A maximum amount of information must be packed into a minimum amount of space.
- Lexical and syntactic means should be selected with a view to the possibility of adding further information.

The STREAK system produces a summary incrementally. First, an initial sentence is drafted, which only contains the basic information. This first version is then reviewed using several operators (c.f. Fig. 8). For example, nominalisations are introduced, since a number of attributes can be added to the nominal phrase, giving facultative information.

1. Initial draft (basic sentence pattern):
"Hartford, CT -- Karl Malone *scored* 39 points Friday night as the Utah Jazz defeated the Boston Celtics 118 94."

2. Adjunctization:
"Hartford, CT -- Karl Malone *tied a session high with* 39 points Friday night as the Utah Jazz defeated the Boston Celtics 118 94."

3. Conjoin:
"Hartford, CT -- Karl Malone tied a session high with 39 points *and Jay Humphries added 24* Friday night as the Utah Jazz defeated the Boston Celtics 118 94."

4. Absorb:
"Hartford, CT -- Karl Malone tied a session high with 39 points and Jay Humphries *came off the bench to add* 24 Friday night as the Utah Jazz defeated the Boston Celtics 118 94."

5. Nominalization:
"Hartford, CT -- Karl Malone tied a session high with 39 points and Jay Humphries came off the bench to add 24 Friday night as the Utah Jazz *handed* the Boston Celtics *their sixth straight home defeat* 118 94."

6. Adjoin:
"Hartford, CT -- Karl Malone tied a session high with 39 points and Jay Humphries came off the bench to add 24 Friday night as the Utah Jazz handed the Boston Celtics their *franchise record* sixth straight home defeat 118 94."

Fig. 8: Incremental generation of a sentence in the STREAK system (Source: McKeown in ENDR94)

Short biographies from a network representation

John Bateman argues with the help of a system which generates short biographies of artists from the knowledge base of an editor workbench. In the form of a semantic network, the knowledge base contains the facts about works, buildings etc. that are derived from input text sentences which can be analysed with the network, i.e. a kind of summary of the input texts. Before a new text is generated from the network (c.f. Fig. 9), the user determines the theme with his question. He thus declares some of the knowledge in the network to be interesting. The next step is to select a text type, which further restricts which facts are extracted from the knowledge base and in what order they are to appear in the text. The planning process considerably reduces the information that is included in the target texts.

Generally speaking, in text generating the amount of information is always limited in such a way that the generating component is not overwhelmed by meaning material that it is supposed to express in linguistic terms. Text generation is thus also a summarising process in which what can be said is reduced to what can be represented in a text.

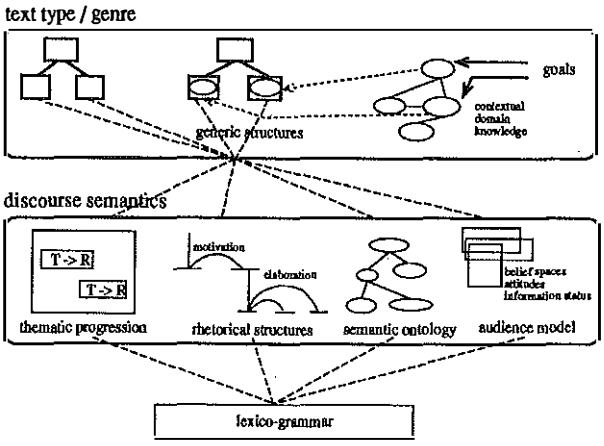

Fig. 9: Generation of short biographies in PENMAN/KOMET (Source: Bateman in ENDR94)

Summarising with PLUM and SPOKESMAN

Since it is difficult for automatic systems to cognitively understand and summarise texts adequately, those solutions that are technically feasible have an advantage.

This is the path taken by *Ralph Weischedel.* He proposed that summarising be made up of a task of information extracting, realised by PLUM, and a task of text generation, performed by SPOKESMAN (METE91). PLUM evaluates various information types (e.g. radar data or documents) by searching through them for predefined categories (objects, persons etc.) and depositing the results in an object-oriented data base. A user states what specific information interests him. His user-specific filter allows him to select from the data base the information that is to be integrated in a summary. In place of a traditional summary, what results is a response to a user question which may be composed of multimodal elements (i.e. tables, text, graphics etc.).

3.3 Summarising as a special case of text production

Abstracting from the perspective of text production

Annely Rothkegel, Sumiko Mushakoji and *Rosemarie Gläser* all refer concurrently to the three possible time relations between a text and a corresponding abstract:

- the reference text is written before the abstract, but is not present at the same time (standard example: an essay and the corresponding abstract in a bibliographic data base)
- reference text and abstract are written parallel to each other and are present at the same time (as with a magazine essay accompanied by an abstract)
- the abstract already exists before the reference text is written (as in the case of abstracts submitted for appraisement as possible conference papers)

In all three constellations, a common central texttheme links the long and the short text version. When writing abstracts, the core information on a theme is presented in a very concentrated form according to the text type. Otherwise the process follows a normal text production model. Fig. 10 illustrates this situation.

The empirical observations of *Sumiko Mushakoji* confirm these findings. The author abstracts studied by her are written by the same authors as essays and deal with the same subject. The fact that an abstract and an essay belong together is explained above all by the common subject. Otherwise, social practice is sufficient to state that the short version (abstract) is a summary of the longer version (the essay). It is not necessary to derive the abstract from the original essay, as summarising models usually do (c.f. Fig. 1). Empirical observations contradict the assumption that an author abstract is entirely derived from the corresponding essay:

- authors often fail to distinguish consciously between what is in the essay and what is in the abstract
- the macrostructure of abstracts often differs from that of the essay.

A realistic approach is to apply two text production processes, which process the same source information - for example experimental results (c.f. Fig. 10) with different goals. On the basis of empirical observation, both production processes can be more readily characterised as a social reconstruction of states of affairs than as their simple representation.

For the study of summarising, it is convenient that summarising is largely a text production process which differs from others through standard parameters such as a target texttype, its communicative function etc. From this follows that general knowledge about text production can be applied to summarising; only the specific characteristics of summarising processes and summaries require further explanation.

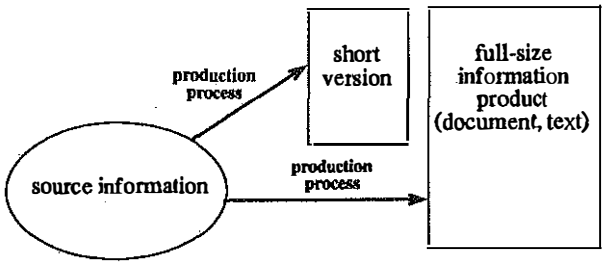

Fig. 10: Text production: source information, long version (original article) and short version (abstract)

Summarised text types as academic genres

From a specialised text-linguistic view, *Rosemarie Gläser* defines summaries as secondary text genres:

"All summarizing texts are derived text genres and depend on a previously existing primary original text."

In this, she falls back on the genre definition of SWAL90 and cites a process-oriented definition of summarising from WERL88:

"In a summary we present the information of a much longer text in much shorter reading or listening time. Through a summary we can inform others about the contents of, for instance, a book or a long chapter in the length of only one short paragraph or even of only one sentence. The original text is translated into a new text."

Since in the academic field a short version often exists prior to the long version, a summary cannot be too explicitly defined as a secondary text, even if it is often the case that a short version is derived from the long version of a text.

Through studying German and English abstracts produced by German-speaking physicians, *Ines Busch-Lauer* observed that the ability to write scientific essays does not necessarily go hand in hand with the ability to write abstracts. The observed physicians are only familiar to a certain extent with the general and linguistic or culture-specific characteristics of the text type abstract. The author draws the conclusion that abstracting should form a separate element of academic training.

3.4 Text type or genre support summarising

Using a discourse model for text analysis

Texttype-specific structure schemata facilitate the analysis, as they determine where what information appears in the text. *Elizabeth Liddy* described the implementation of a discourse model for news texts which builds on DIJK88. It uses a functional schema with the following components: CIRCUMSTANCE, CONSEQUENCE, CREDENTIALS, DEFINITION, ERROR, EVALUATION, EXPECTATION, HISTORY, LEAD, MAIN EVENT, NO COMMENT, PREVIOUS EVENT, REFERENCES and VERBAL REACTION. This functional model was augmented with features which assist the recognition of the text components: key words, tense markers and indicators for the duration of events.

The discourse structure as the basis for summarising

Texts are complex compound objects of meaning. The rules of syntax and compositional semantics determine how text segments can be aggregated, what the individual elements and the aggregate mean. Various coherence relations may exist between the aggregated elements, resulting in a different overall meaning of the text segment, i.e. also a different summary. *Jerry Hobbs* studied five of these coherence relations (explanation, ground-figure, parallelism, contrast and occasion) with a view to finding out how they influence or explain summarising.

3.5 Real discourses

Livia Polanyi pointed out the complexity of real discourses. They are less coherent and monological than we would like. Frequently they include semiotic acts that are related to various semantic models. Even a simple news bulletin often relates to several worlds, e.g. the world in which the newscaster and the listener interrelate and a world in the past which is being reported. When interpreting texts, at minimum the following contexts have to be reconstructed:

- the real world which speaker and listener share
- socially consructed worlds, in which actors are assigned activities, roles etc.
- modal contexts, which determine the attitude of the speaker to the content of his statements
- genres

In spite of this complexity, texts are conventionally structured. From a linguistic point of view, they can be described as dynamic hierarchical structures made up of discourse constituents and non-propositional discourse operators. Discourse constituents are for example sentences, lists, elaborations and speech events. A machine for interpreting discourse must, therefore, be able to interpret constructions made up of dynamic discourse constituents. It produces a large number of representations which reproduce the text meaning as conditions of the contexts in the source text. These representations are the input for the summarising process. What finds its way into the summary is determined by the weighted interaction between the intention of the speaker, the interests of the recipient and the discourse structure.

The formal-linguistic description of summarising given above is, according to *Livia Polanyi*, "AI-incomplete", since the complex model of domain knowledge that is needed for interpretation and is part of the human cognitive process is missing.

3.6 Summarising can be simulated in an empirically founded way

Harold Borko and *Brigitte Endres-Niggemeyer* demonstrated a cognitive process model of abstracting which is based on 36 natural (i.e. taken from day-to-day routine) working processes of 6 experts, one of whom was *Harold Borko*.

The model concentrates on the skills of professional summarising. It presupposes a normal understanding of the text. The principal components of the model are

- a toolbox containing 453 intellectual tools
- the empirically determined principles of human process organisation
- a large number of working processes from natural working environments

The working processes are broken down into working steps. In these, it is possible to study how people combine different intellectual tools in order to realise subgoals of summarising. Thus, the working steps become a propitious environment for an empirically founded inductive system design. The empirical observation lends a simulation system which is oriented towards human work organisation the following main characteristics:

- working steps as basic units of the activity
- cooperating agents (the intellectual tools)
- a blackboard model of system organisation
- a dynamic text representation which works out the document structure

Fig. 11 shows how in a typical working step empirically defined cooperating agents are gathered round their dedicated blackboards in such a way that they can simulate how Harold Borko recognised the theme of the document.

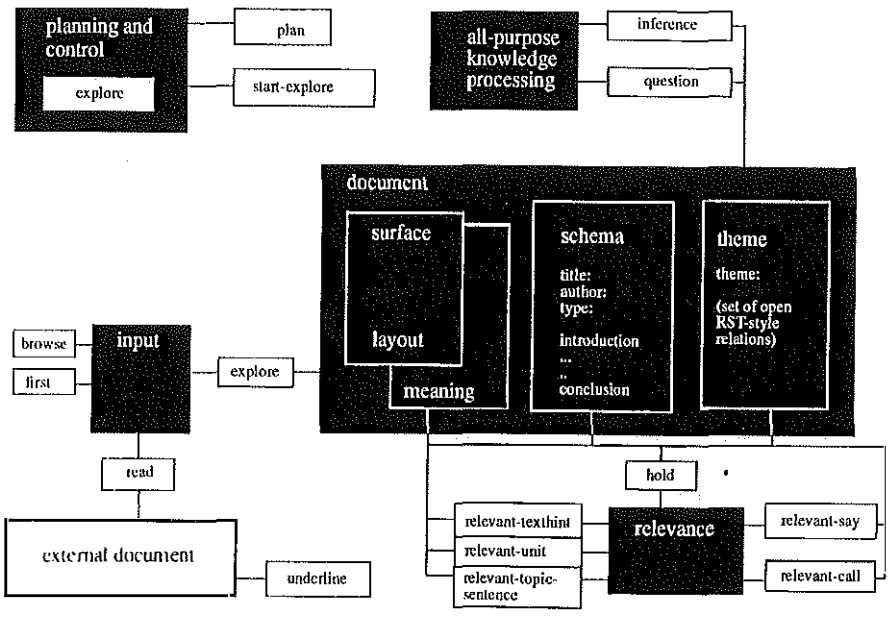

Fig. 11: Blackboard representation of a working step: recognising the text theme

3.7 Still topical: statistical approaches to sentence extraction

Systems that are on the verge of commercial viability as well as new approaches are nearing the skills of summarising by using statistical methods to extract sentences from the original document which appear to carry salient elements of meaning. In addition to this, in many systems statistical techniques are used alongside other methods of information reduction (for examples see above).

The ANES system

The ANES system devised by *Lisa Rau* and her colleagues is comprised of two components:

- a reading component, which presents tokens, sentences and paragraphs in an internal representation that determines the word frequency in the document and the word weighting
- the extraction component, which produces the summary. It checks whether information is to be extracted from the document, weights the sentences and determines which sentences are included in the summary.

ANES uses word frequencies in order to establish relevant sentences. The expectation values are determined in a training phase using a representative corpus specifically for this purpose. Against these, the "signature words" stand out, which determine the weight of the sentence in which they occur. In selecting sentences, not only its weight, but also its position in the document and the occurrence of anaphoric expressions are taken into account.

Much to the disappointment of the authors, the ANES extracts did worse in the system test than the introductory sections of newspaper articles, which generally have a summarising character. It is therefore intended to continue to develop ANES with a view to producing target-oriented extracts of variable length and proceeding beyond what can be found in introductory passages of press articles.

User-oriented summarising by means of text classification

A classification algorithm can also determine which sentences are to be included in a summary. Following the suggestion of *Ellen Riloff*, texts can be classified as relavant or irrelevant with regard to a particular information requirement by comparing them with a representative training set that the user has assessed as relevant. The semantic parser CIRCUS produces a number of instantiated concept nodes as a text analysis. The nodes which refer to one sentence are summarised as a "case" and stored. They serve as the basis for comparison when classifying the text. Texts are summarised by extracting those sentences which find strong correlates among the comparative group, as these sentences are the most closely associated with the information interest.

3.8 Summaries for intelligent users with limited capacity

As *Hans Strohner* reminded us, summaries are intended for intelligent users. In the same way as the processing of instruction texts, the understanding of summaries goes hand in hand with various types of inferences:

- sensomotory inferences in the case of processing with the sensory organs
- syntactic inferences, which support semantic and pragmatic processing
- code inferences, which in particular establish word meanings
- reference inferences, which assign reference objects from the outside world to concepts
- semantic inferences, which combine and instantiate concepts
- pragmatic inferences, which link the text knowledge which results from understanding to the mental model of the interaction partner

These inferences take place during reception. An abstract or a summary can only be understandable providing it does not demand any processing effort on the part of the recipient that is beyond his or her current possibilities.

3.9 What belongs in a summary?

What elements of the original text belong in a summary and how they are best presented depends, according to *Nicholas Belkin*, on the use for which it is intended. Recognising the intentions of the user is important for organising a summary, e.g. in terms of content, degree of detail and structure, in such a way that it helps the user as far as possible to realise his or her goals. How summaries are used, or in what situations people refer to summaries, can be examined empirically. Worthy of study is, ultimately, the interaction between the author and the reader, since the intentions of both are important if an abstract is to enhance communication between them.

Raya Fidel also pointed out the need to study the requirements of users more closely in order to be in a better position to appraise and organise summaries. Since the borderline between summarising and indexing as a special form of summarising is becoming more and more blurred, she argues in favour of taking on results and methods that originally referred to indexing.

Karen Sparck Jones made a constructive, rather than empirical analysis of what belongs in a summary. Since a text contains linguistic, factual and communicative information, we need linguistic, domain knowledge-related and communication-related summarising strategies. Summarising strategies of these three types were applied to ten trial texts ranging from one paragraph to a maximum length of one page. The focus tracking according to SIDN83 provided a natural summarising strategy. The discourse component that was most frequently focused was selected as the text theme. The Rhetorical Structure Theory (MANN88) proved difficult to apply stringently, however offered a convincing summarising method: by selecting the nucleus of each relation. As a top-down strategy for processing domain knowledge, scripts according to DEJO82 were adapted. Density functions selected those parts of scripts as the content of summaries that are most completely instantiated and therefore appeared important.

The intentional structure of texts according to GROS86 provides a natural summarising strategy because it enables us to include the hierarchically highest intentions in the summary.

3.10 Abstracts can be evaluated

Since abstracts or summaries are intended to serve different purposes of various users, which are usually unknown, it is notoriously difficult to assess their quality. An evening discussion round, however, arrived at the following suggestions for evaluating abstracts:

1. An intrinsic evaluation which refers only to the text characteristics of abstracts is not sufficient. It is more informative to assess abstracts extrinsically according to their usefulness in solving user problems. Relevant criteria refer especially to the input which goes into a summary, its intended purpose and the user.

2. Empirical or experimental user studies allow us to establish how and why abstracts are used. For this, it is possible to adapt methods from the field of information retrieval. Methods of qualitative field research including in-depth case studies also look encouraging.

3. Explorative qualitative methods at first appear particularly promising for the evaluation of abstracts. Later, metrics can be developed which correspond to the views of quantitative research with respect to the comparability and reproducibility of results.

4. Research strategies for automatic summarising

4.1 Recycling existing methods to cover the short-term demand for summaries

Short-term practically-oriented research strategies are necessary in order to cover the demand for automatic summarising brought about by the transition to full-text information systems and information networks. Following the model of computer-aided translation, existing summarising methods (in particular methods of sentence extraction) can be adapted in such a way that they are practically successful in appropriate applications. To this end they must - if need be - be combined with other techniques (e.g. postediting).

Since for this purpose an overview of the available methods is needed, it would be a good idea to set up a technology base which could help to recycle existing systems.

4.2 Development of new concepts and techniques

Summaries and especially abstracts are one type of utility texts. Their organisation has always been partly determined by the situation in which they are used. The fact that in the past summaries were usually realised as written texts is explained by the conditions of distribution, since all other forms of representation were more problematical. If technical developments make a graphic, spread or multimedial presentation of summaries possible, it makes sense to redesign summaries under the new conditions. This constellation has obvious advantages for the scientific penetration of abstracting: the greater variety of technical possibilities uncouples the process of summarising from its medial realisation forms and thus invites more general statements about human and automatic summarising.

In view of the forthcoming changes with regard to information systems, conceptual efforts with different orientations seem necessary:

1. New concepts and techniques need to be developed in order to describe and implement summarising. In this context, a modularisation through empirically and formally grounded models is important.

2. Descriptions of summarising must become scientifically more useable. A suitable theory of summarising must be able to deal with different medial and functional conditions, it must, however, particularly be able to describe how the summarising activity itself is carried out under different initial conditions. It must be informed enough to be able to distinguish between central and marginal observations. Ad-hoc solutions which are oriented towards specific realisation environments often operate with auxiliary statements, which have some factors in their favour, but which do not lead to a keener understanding of the phenomenon. A strong tie to the medium may, for example, suggest the superficial view that summarising with pen and paper differs in essence from summarising with a text system.

3. Summarising and summaries need to be reconsidered under the influence of new technical possibilities and growing demands. Since in future information systems summaries will acquire additional functions, new forms need to be developed which take these into account. What is required are summaries which

- make use of other modes of presentation than the written text (e.g. graphics or spoken language)
- serve additional purposes apart from the standard one of pointing out important information (e.g. aiding navigation in full texts or facilitating understanding as an advance organiser)
- no longer necessarily stand as an integrated text in one position in the data base but are just as easily distributed among the source data and can first of all operate locally before playing a role in the context of the overall text
- do not necessarily reproduce a text, but just as easily a body of information or documents from an information system, where possible from the point of view requested by the user
- where possible are only produced when they are needed, because then they can take account of the different purposes different users have for them
- do not have to be monological, but may also be produced in a dialogue with the user

4.3 The link between human and automatic summarising

Good abstracting is measured against qualified human performance. People are flexible in the way they approach things, but they also have a fund of knowledge and a processing routine, which automatic systems have so far not achieved.

In the early days of automatic abstracting it was legitimate to avoid the comparison with human performance. It was difficult enough and relatively successful to produce text extracts which resembled abstracts in that they respected the length restrictions and extracted their linguistic and conceptual material from the original document with a method that attempted to establish the importance of the information through simple means (e.g. word frequency). As the extract from 1958 shows (c.f. Fig. 2), it is indeed possible to achieve practical success with relatively simple methods.

As soon as technical resources in the field of computer science and the level of knowledge about summarising permit better solutions, these possibilities will have to be used in order to approximate the summarising performance of qualified humans. This is meaningful for the simple reason that summaries produced both automatically and by humans have to be used by humans in the same context. Automatic summaries lacking sufficient intellectual quality can have a particularly negative effect if, under normal conditions, users are unable to recognise the errors they contain.

On closer examination, the gap between the way humans and machines produce summaries is not insurmountable. Just as human beings often successfully use less demanding superficial or formal methods instead of thoroughly understanding a text, systems can also be successful, at least in appropriate sub-areas, without performing a "real" summarising action. Similar to humans, they can, for example, rely on easy-to-process indicators for the relevance of individual statements without actively analysing the statements themselves. Since humans summarise events presented in a multimedial form without further difficulty, it will be necessary to proceed from the relatively well researched abstracting of texts and extend the area of consideration to summarising under the natural conditions of potential multimediality.

4.4 Practical research considerations

Studying the nature of summarising

Anyone wanting to know more about summarising should first of all study what constitutes the core of summarising, namely purposefully reducing the presentation to the most important information. In the past, however, summarising was often subsumed under understanding (c.f. the description of SUSY - Fig. 4) or divided between the understanding process and the production process (c.f. the model presented by SCHNO81 - Fig. 3). This is a disadvantage when it comes to studying summarising because summarising is not considered as a separate task.

This problem is avoided by a subdivision of the summarising process into three main components (c.f. Figs. 1 and 6):

- understanding the source information
- the actual summarising (condensation, generalisation, selection etc.), which derives the internal representation of the target information from that of the source information
- production of the target representation

This conceptualisation is also more useful because it takes into consideration the fact that understanding and presenting information are general problems that not only arise in summarising and therefore do not necessarily have to be solved in the context of studying summarising. Automatic summarising systems can, therefore, integrate and adapt general system solutions for interpreting and producing information through different media (text, image etc.). What remains is to examine how "actual" summarising functions and how general modules of input interpretation and output presentation need to be adapted to the special task of summarising.

Defining and evaluating summaries on the basis of their use

The current understanding of summaries barely goes beyond the following generalised statements:

- summaries should be brief or readable in a short time
- they should restrict themselves to the salient propositions of the source information.

However, a sufficiently exact definition of what constitutes summarising and summaries is needed in order to

- produce good summaries
- evaluate summaries
- develop automatic systems for summarising

What elements of the source information are so important that they should be included in a summary is best determined depending on what the summary is to be used for: the summary of an opera libretto must necessarily be different if a mother is deciding whether the opera is suitable for her twelve-year-old daughter and if the same person as a literary scholar is examing how the opera can be classified ideologically.

In order to arrive at a more precise concept of summarising, it is necessary to observe empirically for what purposes summaries and abstracts are used and how competent people produce these task-oriented summaries. An empirically grounded modelling can realistically modularise the intellectual process of summarising and describe types of summaries.

Looking for knowledge where it is: the need for interdisciplinary cooperation

People refer to various sources of knowledge when they summarise something. They allow themselves to be guided by the text type, they use the layout of printed material, they take the reaction of their interlocutor into consideration when summarising in a dialogue, they refer to their specialised knowledge and ask what information is essential for the concrete application etc. Without a comparable

amount of knowledge and processing, systems cannot compete with the summarising quality achieved by competent human beings. Research into automatic abstracting must therefore be founded on the knowledge of different specialised fields. Without combined approaches that take account of the understanding of various specialised fields (textual studies, research into technical languages, computational linguistics, AI, information science), it is difficult in view of the current state of the art to imagine adequate theories and system solutions.

References

(ALTE92) Alterman,R. (1992): Text summarization. 1579-1587 in: Encyclopedia of Artificial Intelligence. New York: Wiley.

(ANTO89) Antos,G.; Krings,H.P. eds.(1989): Textproduktion. Ein interdisziplinärer Forschungsüberblick. Tübingen: Niemeyer.

(APPE93) Appelt,D.E.; Hobbs,J.R.; Bear,J.; Israel,D.; Tyson,M. (1993): FASTUS: A finite-state processor for information extraction from real-world text. 1172-1178 in IJCA93.

(BARA84) Bara,B.G.; Guida,G. eds. (1984): Computational models of natural language processing. Amsterdam: North-Holland.

(BATO89) Batori,S.; Lenders,W.; Putschke,W. eds. (1989): Computational Linguistics - Computerlinguistik. An International Handbook of Computer Oriented Language Research and Applications. Berlin: de Gruyter.

(BELK92) Belkin,N.; Croft, W.B. (1992): Information filtering and information retrieval: Two sides of the same coin? Comm. ACM 35(12): 29-38.

(BORK68) Borko,H. ed. (1968): Automated language processing. New York: Wiley.

(BRAD83) Brady,M.; Berwick,R.C. eds. (1983): Computational models of discourse. Cambridge MA: MIT Press.

(CHAR91) Charolles,M.; Petitjean,A. eds. (1991): Le résumé de texte. Paris: Klincksieck.

(DEJO82) DeJong,G. (1982): An overview of the FRUMP system. 149-175 in LEHN82b.

(DIJK80) Dijk,T. van (1980): Macrostructures. Hillsdale NJ: Erlbaum.

(DIJK88) Dijk, T. van (1988): News analysis: Case studies of international and national news in the press. Hillsdale NJ: Erlbaum.

(EDMU69) Edmundson, H.P. (1969): New methods in automatic abstracting. Journal ACM 16(2): 264-285.

(ENDR92) Endres-Niggemeyer,B. (1992): Abstrahieren, Indexieren und Klassieren. Ein empirisches Proze_modell der Dokumentrepräsentation. Habilitationsschrift, Universität Konstanz, Fak. fnr Verwaltungswissenschaft.

(ENDR94) Endres-Niggemeyer,B.; Hobbs.J.; Sparck Jones,K. eds. (1994): Summarizing Text for Intelligent Communication. Dagstuhl Seminar Report 79, 13.12.-17.12.93 (9350).

(FAYO91) Fayol,M. (1991): Le r_sum_: un bilan provisoire des recherches de psychologie cognitive. 163-172 in CHAR91.

(FUM84) Fum,D.; Guida,G.; Tasso,C. (1984): A propositional language for text representation. 121-150 in BARA84.

(GROS86) Grosz,B.J.; Sidner,C.L. (1986): Attention, intention, and the structure of discourse. Computational Linguistics 12(1986)175-204.

(HAHN90) Hahn,U. (1990): TOPIC parsing: Accounting for text macro structures in full-text analysis. Information Processing & Management 26(1):135-170.

(IEEE91) IEEE Computer Society (1991): Seventh IEEE Conference on AI Applications. Los Almitos CA: IEEE Computer Society Press.

(IJCA93) IJCAI-93: Proceedings of the 11th International Joint Conference on Artificial Intelligence. San Mateo CA: Kaufmann.

(JACO93) Jacobs,P.; Rau,L. (1993): Innovations in text interpretation. Artificial Intelligence 63(1993)143-191.

(KINT83) Kintsch,W.; Dijk,T.A.van (1983): Strategies of Discourse Comprehension. Orlando FLA: Academic Press.

(KNOR93) Knorz,G.; Krause, J.; Womser-Hacker, C. eds. (1993): Information Retrieval 93: Von der Modellierung zur Anwendung. Konstanz: Universit_tsverlag.

(KUHL89a) Kuhlen,R. (1989): Information Retrieval: Verfahren des Abstracting. 688-696 in BATO89.

(KUHL89b): Kuhlen,R.; Hammwöhner,R.; Thiel,U. (1989): TWRM-TOPOGRAPHIC. Informatik Forschung und Entwicklung 4:89-107.

(LABE77) Laberge, D.; Samuels, S.J. eds. (1977): Basic processes in reading. Hillsdale NJ: Erlbaum.

(LEHN82a) Lehnert,W.G. (1982): Plot units: A narrative summarization strategy. 375-412 in LEHN82b).

(LEHN82b) Lehnert,W.G.; Ringle,M.H. eds. (1982): Strategies for natural languages processing. Hillsdale NJ: Erlbaum.

(LUHN58) Luhn,H.P. (1958): The automatic creation of literature abstracts. IBM Joumal, April, 159-163.

(MAND81) Mandl,H. ed. (1981): Zur Psychologie der Textverarbeitung. Mnnchen: Urban & Schwarzenberg.

(MANN88) Mann,W.C.; Thompson,S.A. (1988): Rhetorical Structure Theory: Toward a functional theory of text organization. Text 8(3): 243-281.

(METE91) Meteer,M.W. (1991): SPOKESMAN: Data-driven, object-oriented natural language generation. 435-442 in IEEE91.

(MUC-3) Third Message Understanding Conference (MUC-3), Proceedings (1991). San Mateo CA: Morgan Kaufmann.

(MUC-4) Fourth Message Understanding Conference (MUC-4), Proceedings (1992). San Mateo CA: Morgan Kaufmann.

(RAU89) Rau, L.F.; Jacobs,P.S.; Zernik,U. (1989): Information extraction and text summarization using linguistic knowledge acquisition. Information Processing & Management 25(4):419-428.

(RICK89) Rickheit,G.; Strohner,E. (1989): Textreproduktion. 220-256 in ANTO89.

(RUME77) Rumelhart, D.E. (1977): Understanding and summarising brief stories. 265-303 in LABE77.

(SCHN81) Schnotz,W.; Ballstaedt,S.-P.; Mandl,H. (1981): Kognitive Prozesse beim Zusammenfassen von Lehrtexten. 108-167 in MAND81.

(SIDN83) Sidner, C.L. (1983): Focusing in the comprehension of definite anaphora. 267-330 in BRAD83.

(SPAR93) Sparck Jones, K. (1993): What might be in a summary? 9-26 in KNOR93.

(SWAL90) Swales,J.M. (1990): Genre analysis. English in academic and research settings. Cambridge MA: Cambridge University Press.

(WERL88) Werlich,E. (1988): Summaries/Summarizing minutes. 42-52 in Werlich,E. ed. (1988): Student's guide to text production. Berlin: Comelsen.

Prof.Dr.Brigitte Endres-Niggemeyer, FHS Fachbereich Bibliothekswesen, Informationund Dokumentation, Hanomagstr.8, D-30449 Hannover