[TMI 99: *Proceedings of 8th International Conference on Theoretical and Methodological Issues in Machine Translation*, August 1999, Chester, UK]

MENTAL SPACES, SPACE BUILDERS AND BILINGUAL SUMMARIZATION OF NEWS REPORTS

Barbara Gawronska, Jaana Anttila and Dan-Ivar Jacobsson

Department of Languages, University of Skovde, P.O. Box 408, S-541 28 Skovde, Sweden

Fax Number: +46 500-464825
Phone Number: +46 500-464818
E-mail: *barbara.gawronska@isp.his.se, a96jaan@student.his.se, a96danja@student.his.se*

Abstract

The paper presents the general architecture of an experimental system for English-Swedish written and spoken summarization of news reports and focuses on the information extraction component. Information extraction and information structuring is based on the notion of mental spaces - one of the central notions in cognitive semantics. Speech act phrases, epistemic verbs, tense forms and certain adverbs and subjunctions are identified by the semantico-syntactic parsing procedure as marking shifts between different mental spaces and the textual information is structured accordingly. Less salient mental spaces are omitted in the textual representation. The summary generation component has access to language specific ways of formulating news reports. The text generator also provides the syntactic structures with prosodic markers that modify the default prosodic rules of the text-to-speech system that reads the summary.

1 Introduction: Text summarization versus machine translation

Most computational linguists of today agree as to the fact that information extraction combined with multilingual generation of text summaries can serve as a useful alternative for machine translation. It has been pointed out (Somers et al. 1990; Kitani et al. 1994; Glass et al. 1994; Myers & Mulgaonkar 1995; Beale et al. 1996) that it in many cases it is better for the user to be offered a short and comprehensible summary than a poor translation of a whole long text. This does not mean that the "traditional" machine translation is no longer of interest; rather, the different methods of transferring textual information into a target language shall be seen as simulations of different communication tasks. A machine translation system is thought to act as a human translator attempting to render the source text as exact as possible (the quality of the output depending not only on pure language skills but also on domain knowledge and the translation method), whereas an ideal multilingual summarization system should work as a well-educated secretary (Somers et al. 1990), providing the user with the very essence of particular texts and waiting for further requests (Shall the whole text be sent to a professional translator? Shall certain paragraphs be highlighted? Does the user need some background information? Does the user want to comment the text and/or - if it is a letter - to answer it in some particular language?). Thus, the design of translation system and summarization system must differ, though both require some kind of simulation of linguistic knowledge and domain knowledge.

A quite common assumption is that summarization systems can do without advanced syntactic and semantic components and mainly rely on rhetoric patterns and statistics. By this method, summarization is achieved by isolating text fragments containing words, phrases and collocations that are statistically salient with respect to a given domain and a given rhetoric pattern. These fragments may later be 'squeezed' by elimination of certain adverbial and attributive phrases, and the result is presented in the source language; as the next step, a machine translation component may apply. One of the most advanced summarization systems of today, the MIND at CRL (Beale et al. 1996), works to a great extent in this way. This method appears to be successful when applied to quite long texts, having homogenous main topics. When dealing with text documents consisting of parts belonging to different topic areas, although connected by some common denominator - as in the case of e.g. news texts - other summarization strategies are required. In news summarization, statistic based cues without quite elaborated syntactico-semantic parsing easily become misleading (Nirenburg et al. 1997). If the system is too eager to draw quick conclusions from the appearance of certain lexical items, headlines like "Asian stocks lose battle" or "Blair moves to educate an army of computer engineers" may be interpreted as referring to military operations (Gawronska & House 1998). Furthermore, the user may be interested in getting not only a summary of the news of today, but also in information about the background and development of a particular event. The system shall thus preferably be able to generate a short description of a course of events.

2 The Newspeak project

The procedure for multilingual generation of text summaries presented below is an essential part of the experimental system Newspeak, currently used as a research tool. The summarization and generation procedure makes use of the components shown in fig 1:

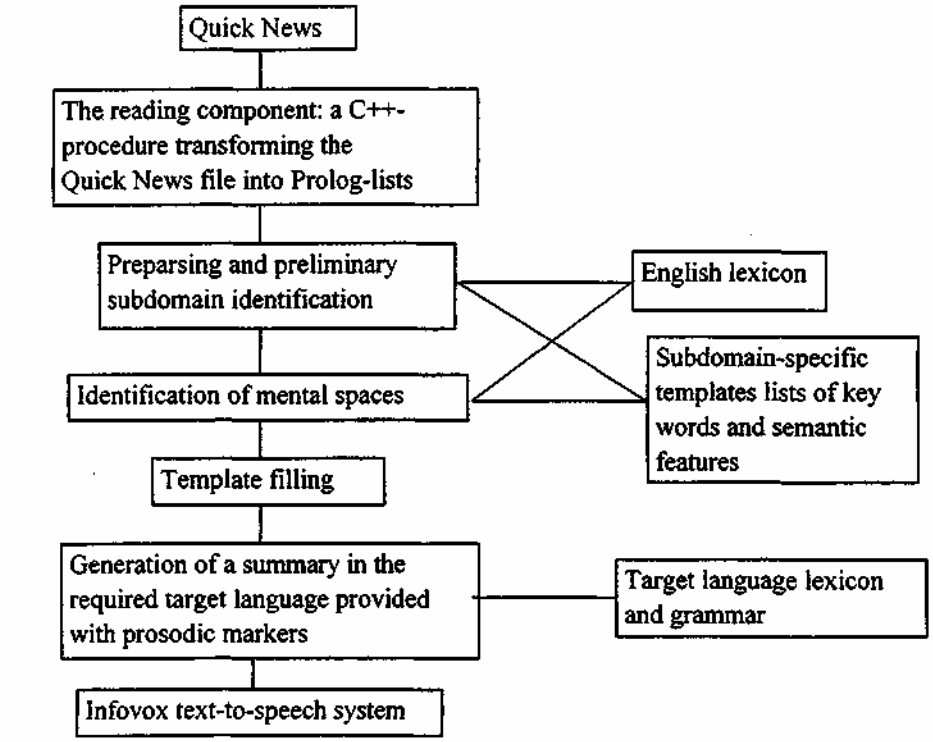

Figure1: the architecture of the summarization system

The system uses CNN's internet-based Quick News service in English. Its goal is to give the user a news summary in either English or Swedish and to answer the user's questions concerning certain event developments. The output may be given as a written text or be read by the system using the Infovox text-to-speech system (Blomberg et al. 1993; Gawronska & House 1998). Restricted domains are currently certain political events (state visits, strikes and riots) and disasters (natural disasters, transportation accidents and terrorist incidents). Before the summarization and generation procedure starts, the news file is read by a C++ program that transforms the text into Prolog lists and converts certain graphical items (e.g. proper names) into Prolog atoms. This task could also be performed by Prolog, but the C++ code works quicker and more efficient.

The next step is a preparsing procedure; the result is a quite flat syntactic structure, distinguishing between subject phrases, verb phrases and topicalized adverbial phrases. In course of preparsing, lexical items are matched against semantic descriptions in the English lexicon and lists of subdomain-specific keywords and bundles of semantic features. This part of the procedure enables first guesses concerning semantically heavy versus semantically 'empty' or less important parts of the sentence (e.g. dummy subjects are eliminated, and auxiliaries are identified as less important items; sentence parts that the system assumes to be less important are only partially parsed). At the same time, the system makes first assumptions as to the subdomain the text is about. A sample output from the preparsing component is shown below; the output corresponds to the first three sentences in text (1).

(1)
BOMB EXPLODES IN BILBAO BEFORE GUGGENHEIM DEBUT
A bomb exploded in Bilbao, Spain, early Friday morning after a warning to police from the Basque separatist group ETA. The explosion took place just one day before the inauguration of the city's new Guggenheim museum by King Juan Carlos and Queen Sofia, officials said. There were no injuries and little damage. ETA has killed some 800 people in a nearly 30-year campaign of violence for an independent Basque state.

```
temp([terrorists,explosion],
    s(np([salient,m([indefj),m([bomb])]),
      vp([salient,verb([exploded],m([explode]),[verb,fin]),
          in,placep([m([bilbao]),m([spain])]),timep([m([early]),m([friday]),m([morning])]),
             background([after,a,warning,to,police,from,the,basque,separatist,group,eta])])))

temp([terrorists,explosion],
    s(np([salient,m([def]),m([explosion])]),vp([verb([took],m([took]),[verb]),
      placep([m([place])]),background([just,one,timep(day),before,the,inauguration,of,the,citys,new,
      guggenheim,placep([m([museum])]),by,king,juan,carlos,and,queen,Sofia]),
      speech_act([officials,verb([said],m([say]),[verb])]])))

temp([terrorist,explosion],
   s(np([m([dummy])]),vp([salient,verb([were],m([were]),[verb,aux]),injuries(no),and,
      damage(little)])))
```

The central part of the summarization procedure in Newspeak is called 'mental space identification'. Its theoretical foundations make use of the notion of mental spaces introduced by Fauconnier (1985) and further elaborated by several adherents of cognitive

linguistics (Dancygier & Sweetser 1996; Sweetser & Fauconnier 1996). In writings on computational linguistics, the term 'beliefs' is sometimes used in a sense almost identical with Fauconnier's 'mental spaces' (Wilks 1985; Lee & Wilks 1996). However, we prefer to use Fauconnier's term, since the use of the notion 'belief'could suggest that the system has access to the beliefs of the sender(s) of the news, something which is impossible for obvious reasons.

3 Identification of mental spaces

Mental spaces are to be understood as cognitive 'worlds' that differ from logical possible worlds. The most salient difference lies in the fact that if a mental space M is included in another mental space M', the elements of M do not have to be elements of M'. One of Fauconnier's classical examples illustrating the claim above is the sentence *In Len's picture, a witch is riding a unicorn* (Fauconnier 1985:19). Here, the mental objects corresponding to the NPs *a witch* and *an unicorn* belong to the mental space M covering the 'picture reality'; M is included in M', which corresponds to (roughly expressed) the objective reality - but this does not imply that the members of M are also members of M'. In the communication process, different mental spaces are continuously established by various linguistic means, and the communication is successful if the receiver manages to identify the mental spaces constructed by the sender and to place appropriate discourse objects and relations in right mental spaces. This is the assumption that the text understanding model in Newspeak is based upon.

3.1 Types of linguistic space-builders

New mental spaces can be created by linguistic markers belonging to different levels of language structure. The picture-space in Fauconnier's example is established mainly by lexical means: nouns denoting 'narrative' artefacts like books, films, plays, pictures etc frequently function as space-builders when occurring within adverbial PPs. But there are more space-builder types:

- morphological: tense forms, conditional forms
- syntactic: conditional constructions
- syntactico-semantic: subordinated clauses expressing spatial and temporal relations, superordinated clauses containing speech-act verbs or mental-state verbs like *the government spokesman **said**, New York Times **reported**, officials **believe*** etc
- lexically-semantic: time and space expressions, sentence adverbs expressing probability.

These types of markers must be taken into consideration before deciding which phrases and/or sentences in the input news text shall be used for summary generation, and which shall be omitted or stored as possible background information. For example, the last sentence in text (1) *(ETA has killed some 800 people in a nearly 30-year campaign of violence...)* is not necessary when generating a summary, and a possible confusion when extracting the information about the effect of the terrorist action shall be avoided (the system must not believe that the action reported caused the death of 800 people, something that easily can happen in a system that relays on key words). According to our model, the background status of the last sentence is discovered by identification of the tense change (Present Perfect after a continuous use of Simple Past) and the time expression *30 years*. Both markers create the background mental space, in which the salient event space (Bilbao,

Spain, early Friday morning) is included. Only the objects and relations belonging to the later mental space shall be referred to in the summary.

3.2 Speech act spaces and event spaces

In most news texts, the salient mental space, i.e. the space corresponding to the reported event, is embedded in a speech act space. The later (like *official said* in the sample text above) is in many cases of no importance for the user and can often be omitted in the summary, but it limits must be identified and its contents investigated for several reasons.

Firstly, the time of the speech event must be distinguished from the salient event time, as in sample text (2):

(2)
RUSSIAN ARMY HELICOPTER CRASHES, SIX FEARED DEAD
Six officers were feared dead after a Russian military helicopter crashed northwest of Moscow, Interfax news agency said on **Thursday**. A defense ministry spokesman confirmed by telephone that a Ka-27 helicopter had come down on **Wednesday** near Tver, about 90 miles northwest of the capital, but declined to say if there were any casualties.

Secondly, some elements of the event spaces embedded in different speech act spaces may be incompatible - as the different versions of the result of a military conflict in text (3).

(3)
GUERILLA FIGHTS IN LEBANON
Israeli warplanes and artillery attacked suspected guerrilla hideouts Friday following a series of clashes in south Lebanon. Four guerrillas were reportedly killed. Guerrillas of the Syrian-backed Amal group attacked Israeli and allied militia positions in the Israeli-occupied zone at daybreak, Lebanese security officials said. **Three guerrillas were killed** in the assaults, **said an Israeli army spokesman** in Jerusalem. **Amal said none of its fighters was killed.**

Furthermore, the speech act phrase may contain information which belongs to the event-space, but is syntactically connected to the speech act clause, as in sample (4). This information shall be extracted and placed in the event template used for summary generation.

(4)
PIPELINE EXPLOSION KILLS 500
At least 500 people were killed and 100 wounded when a gasoline pipeline exploded as thieves were stealing from it, officials said Monday in the latest update on **the disaster in Nigeria**. With dozens of charred, unidentified bodies littering the town and flames still occasionally flaring, authorities began burying those killed in a mass grave and relatives converged on Jesse, 300 kilometers (180 miles) southeast of Lagos, to search for remains of their loved ones.

Relations between speech act spaces and event spaces in text (2) (the helicopter crash) are relatively straightforward to formalize. They correspond to figure 2:

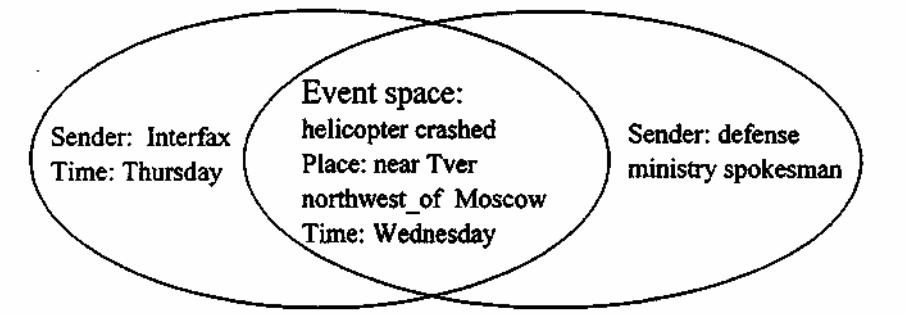

Figure 2: space relations in sample text (2)

The event spaces embedded in the two different speech act spaces are not incompatible. Thus, the speech act clauses can be omitted in the summary and the predefined template is filled as shown below:

slot(index,[**AM,073001,98,Nov,20,Friday**])
slot(time,timep([m([**Wednesday**])]))
slot(place,placep([m([**tver**]),**northwest_of,** m([**moscow**])]))
source(m([**helicopter**]))
slot(cause([**unknown**]))
slot(result([injuries(**unknown**),damage(m([**helicopter**])),dead(**hypothese(6)**)]))

The instantiated variables are marked by bold face; the plain text corresponds to the predefined Prolog structure.

The relations between mental spaces in text (3) (guerrilla fights in Lebanon) are slightly more complicated. The main event space (Israeli attack against Amal guerrilla) is embedded in three different speech act phrases; two senders give different version of the result. In such cases, speech acts connected to different versions of an event shall be rendered in the summary. Template filling is thus preceded by sender identification, and the different event versions are connected to the different senders. Below we show fragments of the output of the sender identification procedure (the functor "m" is a marker of the lexical interlingua code):

Sender unmarked, main event space
Statement:
[salient,m([israeli]),m([warplanes]),m([and]),m([artillery])]
[verb([attacked],m([attack]),[verb,prt,fin]),verb([suspected],m([suspect]),[verb,prt,fin]),guerrilla,place
p([m([hideouts])]),timep([m([friday])]),verb([following],m([following]),[verb]),a,series,of,clashes,
,placep([in,south,m([lebanon])])]
Statement:
[m([four]),m([guerrillas])][verb([were],m([were]),[verb]),reportedly,verb([killed],m([kill])[verb,fin])]

Sender is: [m([lebanese]),m([security]),m([officials])]
Statement:
[salient,m([guerrillas]),p(m([of])),m([def]),m([syrian]),m([amal]),m([group]),attacked,m([attack]),m([israeli]),m([and]),m([allied]),m([militia]),m([positions]),p(m([in])),m([def]),m([israeli-occupied]),m([zone]),p(m([at])),m([daybreak])]

Figure 3 is an attempt to illustrate the mental space structure that serves as the basis for template filling:

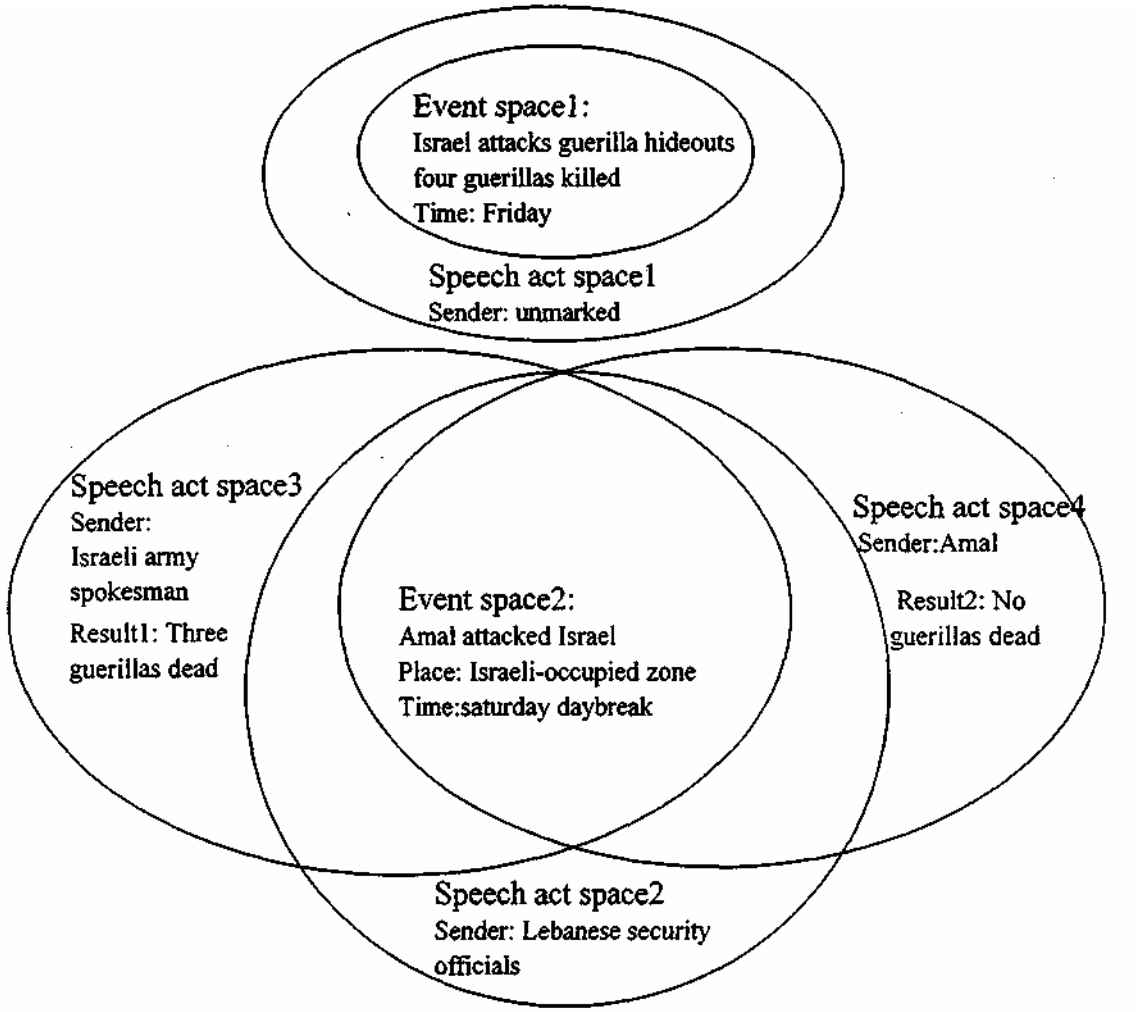

Figure 3:space relations in sample text (3)

The main problem when analyzing text (4) (pipeline explosion in Nigeria) is the fact that much relevant information is placed within the speech act clause *officials said Monday in the latest update on the disaster in Nigeria.* As country names are treated as very salient pieces of information, the noun *Nigeria* is extracted from the speech act phrase. Since the noun *disaster* in the superordinated clause and the event type in the subordinated clause *(explosion)* are related by hyponymy, a coreference link is established between the reported event and the NP *the disaster in Nigeria* and the phrase *in Nigeria* is subsequently placed in the salient event space. The hypo- and hyperonymy relations are identified by means of lexical links, connecting entries in the lexical data base.

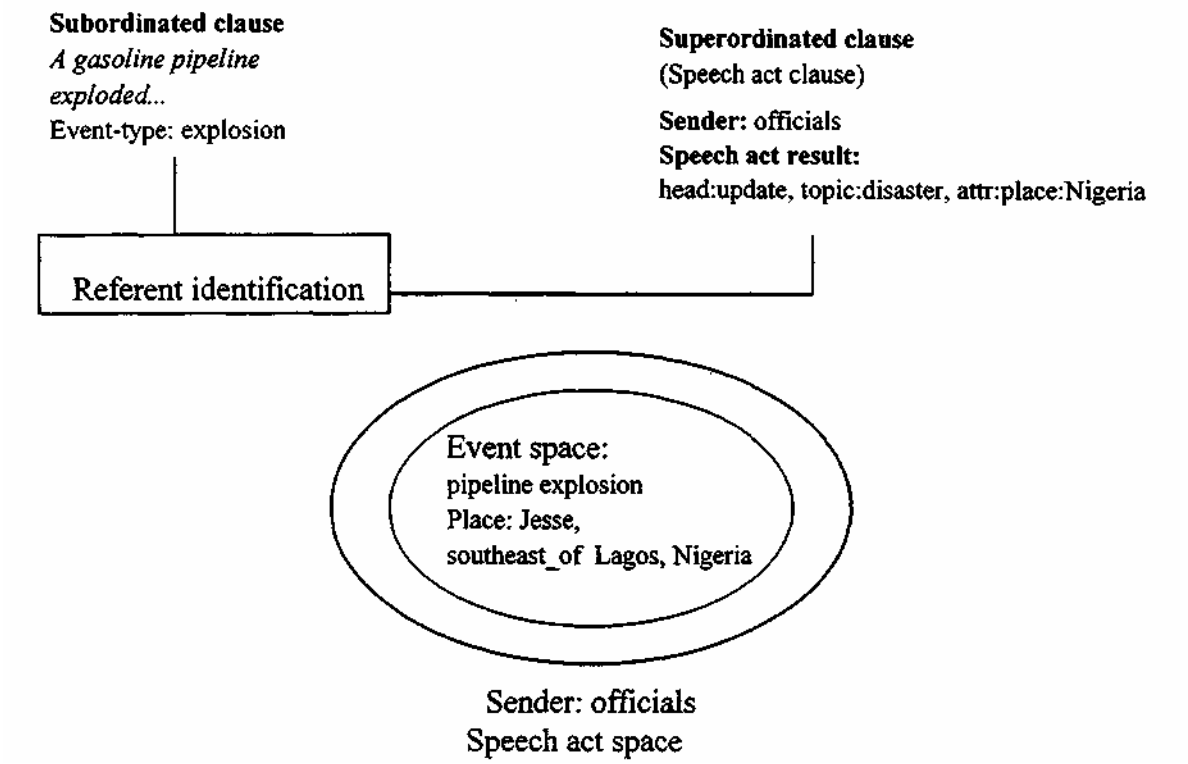

Figure 4: space relations in sample text (4)

The template containing the information extracted from text (4) has the following shape:
slot(index,[**AM,073001,98,Oct,19,Monday**]), slot(time,timep([m([**default**])])),
slot(place,placep([m([**jesse**]),**southwest_of**, m([**lagos**]), m([**nigeria**])]))
slot(source([m([**gasoline,pipeline**])],slot(cause([**stealing**]))
slot(result([injuries(**100**),damage(m([**gasoline,pipeline**])),dead(**500**)]))

3.3 Hypothetical mental spaces

In event categorization, the semantics of the arguments of the verb often plays a crucial role. Especially the arguments having the semantic roles of source/cause/agent are of importance when deciding about the main event type. Predications containing the same verb may result in different interpretations depending on the semantics of the arguments. Compare e.g. *a bomb exploded* (default reading: terrorist attack), *a pipeline exploded* (default reading: accident), *the police exploded a bomb* (default reading: discovery and prevention of a planned terrorist attack), *the president candidate exploded in laughter* (default reading: meeting, discussion). Structures like the examples above are currently handled in a quite successful way by matching the verbal arguments against bundles of semantic features specified in the lexicon. However, the semantic roles 'cause' and 'agent' are not always clearly specified in news texts. Frequently, especially within the natural disaster/accident/terrorist attack subdomain, only assumptions as to the nature of the cause and or the suspected agent are expressed. This requires identification of hypothetical space builders, i.e. those linguistic markers that introduce guesses and assumptions (Dancygier & Sweetser 1996). The most frequent hypothetical space builders in news texts are negated

epistemic predications *(nobody knows, it is unknown* etc.) combined with interrogative subjunctions, as in text (5):

(5)
Three Moscow metro workers were injuried Thursday when a bag left on a train exploded. According to the Interfax agency, a driver found the bag and took it into a duty room at Tretyakov station. It detonated 10 minutes later. (...) **No one knows if** a device was **planted deliberately** or if it was **leftover from New Year's Eve**.

The system distinguishes elements of hypothetical spaces from the event space elements by marking the relevant slots in the template by the constant 'hypothetical'.

4 Conclusions

Current experiments with Newspeak have shown that distinguishing between different mental spaces improves the quality of generated summaries and even the quality of speech synthesis. In text generation, the filled templates function as an interlingua representation. The generation module is implemented in DCG with the filled template right of the arrow and language-specific English and Swedish syntactic structures on the left of the arrow. The language specific syntactic components are provided by numerical values modifying the default prosody of the Infovox text-to-speech system, especially the focal accent (House et al. 1997). Since the mental space model used in template filling enables distinguishing between more and less salient discourse referent, and since referent identification becomes easier (normally, coreference links are established within the same mental space), the degree of accentuation based on giveness/newness of a referent (Steedman 1996) can be chosen with a greater accuracy. Furthermore, text generation from interlingual space and event structure gives the possibility of achieving a more idiomatic lexical and phrasal choice. The Swedish module makes use of standardized ways of presenting certain events in Swedish news and to syntactic structures that are functionally, although not syntactically equivalent to certain English phrases. This is difficult to achieve in syntax based machine translation. Some examples of functionally equivalent phrases that differ syntactically and lexically are shown below (the sentences are taken from Newspeak's summarization of sample text (1)).

E: A bomb exploded in Bilbao, Spain, early Friday morning.
S: En bomb exploderade i den spanska staden Bilbao tidigt på fredagsmorgonen
a bomb explode-PAST in DEF Spanish city Bilbao early on Friday-morning-DEF

E: There were no injuries.
S: Inga personskador rapporterades
no person-injuries report-PAST-PASSIVE

E: ETA is suspected for being responsible for the attack.
S: Förmodligen ligger ETA bakom bombdådet.
Presumably lay-PRES ETA behind bomb-outrage-DEF

The summarization process works currently quite well within the subdomain of disasters (recall and precision about 70%). The politics subdomain requires an extension of the world knowledge component and elaboration of templates (the structure of events within politics is considerably less predictable). This will be the next stem in the development of the system. Furthermore, the existing generation component will be extended and tested for a non-Germanic language, preferably Greek or one of the Slavic languages.

References

Beale, S., Nirenburg, S. & Mahesh, K. 1996. HUNTER-GATHERER:
Three Search Techniques Integrated for Natural Language Semantics. *AAAI/IAAI,* volume 2pp. 1056-1061.

Blomberg, ML, Carlsson, R., Elenius, K., Granström, B., Gustavsson, J., Hunnicutt, S., Lindell, R., Neovius, L. 1993. An Experimental Dialogue System: Waxholm. In *Proceeding of Eurospeech -93.* Berlin. Pp 1867- 1870.

Dancygier, B. & Sweetser. E. 1996. Conditionals, Distancing, and Alternative Spaces. In Goldberg, A. E. (ed.): *Conceptual Structure, Discourse and Language.* Stanford, California: CSLI Publications

Fauconnier, G. 1985. *Mental Spaces: Aspects of Meaning construction in Natural Language.* Cambridge, Mass.: MIT Press.

Gawronska, B. & House, D. 1998. Information extraction and text generation of news reports for a Swedish-English bilingual spoken dialogue system. In: R.H. Mannell & J. Robert-Ribes (eds.): *ICSLP '98 Proceedings,* volume 4, pp. 1139-1142.

Glass, J., J. Polifroni & S. Seneff. 1994. Multilingual language generation across multiple domains. *ICSLP '94 Proceedings,* Yokohama, pp. 983-986.

House, D., Hermes, D.J. & Beaugendre, F. 1997. Temporal-alignment categories of accent-lending rises and fells. *Proceedings Eurospeech '97. Rhodes, Greece,* pp. 879-882.

Kitani, T., Eriguchi, Y & Kara, M. 1994. Pattern matching in the Textract information extraction system. *Proceedings of CoLing '94, Kyoto,* pp. 344-348.

Lee, M. & Wilks, Y. 1996. An ascription-base approach to Speech Acts. *Proceedings of CoLing'96,* pp. 699-704.

Myers, G. K. & Mulgaonkar, P.G. 1995. Automatic extraction of information from printed documents. *Proceedings of the Fourth Annual Symposium on Document Analysis and Recognition,* Tsukuba Science City, Japan, pp. 159-162.

Nirenburg, S. & Mahesh, K. 1997. Knowledge based systems for Natural language Processing. The Computer Science and Engineering Handbook 1997, pp. 637-653.

Somers, H. Tsuji, J. & Jones, D. 1990. Machine translation without a source text. *Proceedings of CoLing '90.* Helsinki. Vol. 3, 271-277.

Steedman, M. 1996. Representing discourse information for spoken dialogue generation. *ICSLP '96 Proceedings,* Philadelphia, pp. 89-92.

Sweetser. E. & Fauconnier, G. 1996. Cognitive links and domains: basic aspects of mental space theory. In: Fauconnier, G. & Sweetser. E. (eds): *Spaces, worlds and grammar.* The University of Chicago Press: Chicago and London, pp. 1-28.

Wilks, Y. 1985. Relevance, points of view and speech acts: An artificial intelligence view. *Technical Report MCCS-85-25,* New Mexico State University.