

Employing Cognitive Notions in Multilingual Summarization of News Reports

Barbara Gawronska

Dept. of Languages, University of Skövde, Box 408
S-542 28 Skövde, Sweden
barbara.gawronska@isp.his.se

Abstract. The paper presents an approach to automatic text understanding inspired by speech act theory and cognitive semantics, especially by the notion of ‘mental spaces’ (Fauconnier 1985), and by Pustejovsky’s (1991a, 1991b, 1995) notion of ‘qualia’ and his definition of formal and telic hyponymy. This approach is employed in an experimental system for understanding of news reports and multilingual generation of news summaries. The system, aims at analyses of English news reports in the domain of the world’s news and generation of summaries in Swedish, Danish, and Polish. The paper focuses on the understanding component and on the possibility of using WordNet as the main lexical knowledge resource for English. An appropriate semantic analysis requires some modifications of the hyper-/hyponymy and holo-/meronymy relations that are encoded in WordNet. A combination of a cognitive analysis of certain lexical and phrasal categories with qualia-based re-formulations of WordNet hierarchies is proposed and tested.

1. Introduction

While information retrieval (IR, or automatic document classification), is dominated by statistical methods, deep text understanding requires semantic analysis and, consequently, semantic classification of lexical entries (for a survey of classic statistical and knowledge-based retrieval and abstracting methods, see Mani and Maybury 1999, Zechner 1995). In high-quality text understanding systems, statistical methods can hardly be successful when not combined with semantic, rhetoric and pragmatic information (Nirenburg 1996, Nirenburg and Mahes 1997). Most work on text summarization employs some kind of semantic and/or ontological classification, and some variant of logical analysis, although the approaches differ with regard to the terminology and to the details of their theoretical background (for a survey of knowledge-based approaches to text summarization, see Hahn and Reimer 1999, Mani and Maybury 1999). These differences often depend on the domain of the study. ‘Plot units’ (Lehnert 1999) are no doubt useful when analysing coherent narratives, where temporal and causative relations play a crucial role for the text structure. When summarizing human-human dialogues, the game-theoretical approach appears to be fruitful – as for example in the Verbmobil system (Wahlster 2000, Alexandersson et al 2000). For summarization of longer monologues, like politicians’ speeches, a

‘squeezing’ strategy, i.e. a strategy aimed at isolating fragments that are salient with respect to a certain domain and a given rhetoric pattern, is quite efficient (Beale et al. 1996). The differences in approaches to summarization are no doubt psycholinguistically motivated, since human memorization strategies and the focus of human attention differ in different cognitive tasks (Miller and Johnson Laird 1976). Understanding a spoken dialogue requires a different neurolinguistic activity than understanding a long written narrative. However, a lot of terminological discrepancies in the field of text understanding could be avoided, if some notions introduced by cognitive linguists, especially the notion of ‘mental spaces’ (Fauconnier 1985), were more widely employed. Cognitive linguistics with its focus on relations between linguistic structures and human perception and conceptualisation mechanisms (Jackendoff 1983, Langacker 1991, Lakoff 1993, Talmy 1988) provides useful tools for text understanding methodology and can be integrated with computational strategies. In the following, we will illustrate this claim by presenting an experimental system for multilingual text summarization, employing the notion of mental spaces and combining it with Pustejovsky’s (1991a, 1991b, 1995) notion of *qualia*.

2. Cognitive grammar in text understanding

2.1. Background: Implementing a prototype summarization system

The practical goal of the first phase of the project to be presented here was to construct a prototype of an interactive news summarization system, according to the outline in Figure 1.

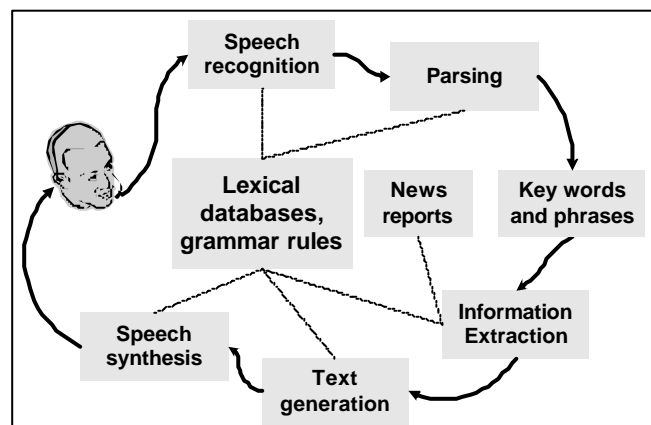


Fig. 1. The general architecture of the interactive summarization system

The system was thought to function as a ‘well-educated secretary’ (Somers et al 1990), who, given a question in language A, collects the relevant text(s) in language B, looks for the relevant information, renders it in language B, and awaits further questions. The main objectives of the project, in its initial stage, were the following:

- Evaluation of different methods of semantic classification in the lexicon
- Development of a summarization module that would be well-suited for the news domain
- A comparison between ‘traditional’ machine translation (MT) on the one side, and information extraction (IE) combined with multilingual text generation (MTG) on the other side
- Exploration of the interplay between textual structure, syntax, and prosodic markers.

In the following, we will concentrate on the three first tasks, and refrain from a presentation of the text-to-speech module, since the work on this part of the system has been published in another context.

2.2. ‘Mental spaces’ in news reports

Perhaps the most serious difficulty for both statistical and logical methods in the domain of world news is caused by the fact that a news report often contains many versions of an event, seen from the viewpoints of different sources/witnesses – as in sample 1 (phrases marking different view points are highlighted):

Sample text 1

(CNN, August 30th 2001): BEIT JALA, West Bank (CNN) – Israeli troops pulled out of Beit Jala before dawn on Thursday, leaving the Palestinian town quiet amid reports of fresh violence in other West Bank towns.

The Palestinians said the Israel Defence Forces had staged incursions into Hebron, killing one and injuring 16 others, and Tulkarem, killing one and injuring 10. The Israel Defence Forces (IDF) had no immediate comment on the accusation that troops had entered Tulkarem, and strongly denied there was an incursion at Hebron. The army said there was an exchange of fire in the town’s Abu Sneineh neighbourhood ...

Different versions of an event can be handled either by means of modal (epistemic and doxastic – Hintikka 1962) logic, or in the spirit of cognitive semantics – by means of the notion of ‘mental spaces’ (Fauconnier 1985, Sweetser and Fauconnier 1996, Dancygier and Sweetser 1996). A ‘mental space’ is a kind of possible world, but it differs from possible worlds in logic (Carnap 1947) by not obeying the set calculus laws: if a mental space M1 is included in a mental space M, the elements of M1 do not have to be elements of M. For example, in the text above, ‘incursions into Hebron’ are elements of a mental space M1 = the Palestinian version of the event course, but there is no possibility to decide whether they are elements of M = the ‘objective’ reality. Even if ‘incursions into Hebron’ have not taken place in M, M1 is nevertheless included in M, since the report exists. The relations among the different mental spaces evoked by the first two paragraphs of Sample text 1 are schematically shown in Figure 2. In the figure, we use a notation which is slightly simplified in comparison with Fauconnier’s (1985) original notation. For an analysis of narratives, following all subtle details of Fauconnier’s model, see Sanders and Redeker (1996).

The ‘mental space’ approach is, in its very essence, compatible with e.g. the work of Wilks (1985) and Lee and Wilks (1996), that focus on the representation of speech acts and the sender’s/receiver’s points of view in NLP- and AI-applications. Farwell and Helmreich (1997, 2001) also argue for the need of representing different points of

view, or ‘beliefs’ (or if one prefers the cognitive linguistic terminology – ‘mental spaces’) in high quality NLP, especially in MT. Amid the terminological differences, the common ancestor of the pragmatics/cognition based approaches is Austin’s (1962) and Searle’s (1969) speech act theory, and the common denominator is Grice’s (1975) notion of conversational implicature. What is essential for text understanding in these approaches is not the truth-value of a predication, but the implicature/mental space evoked by an utterance. This mental space is evoked even if the recipient does not believe that the utterance expresses a true predication. In the following, we therefore prefer to use the term ‘mental space’, and sometimes ‘speech act space’ (the latter term referring to mental spaces evoked by speech act phrases), since, in the domain of news reports, these terms seem to be more adequate than the notions of ‘beliefs’ in modal logic. The utterances of news sources do not necessarily reflect their actual beliefs or their knowledge status. The cognitive model is thus, in this context, more appealing.

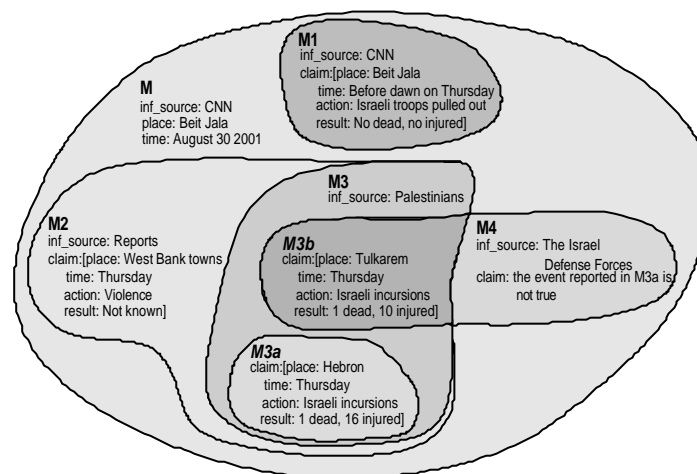


Fig. 2. ‘Mental spaces’ in the first two paragraphs of Sample text 1

2.3. The classification of mental spaces in news reports and its relevance for the text understanding process

Most mental spaces in the news reports are introduced by verb and phrases referring to speech acts (*X said, X reported*), but even adverbs, phrases with certain nouns, and complex prepositions can function in this way (*reportedly, amid the reports, according to*). The phrases referring to speech acts can be divided into different subgroups: neutral informative (*said*), affirmative (*confirmed*), negative (*denied*), and ‘utterance refusal’ (*declined to say, had no comments*). Other quite frequent introducers of new mental spaces are verbs and phrases – mainly negated – referring to epistemic states (*it is not known*), or to combinations of epistemic and emotional states (*feared*). Other groups are ‘declarations’ (*condemn, prize*), and ‘commissives’ (*promise, threaten*; Searle 1969). The last mentioned group introduces hypothetical mental spaces. Below, some examples of ‘space builders’ in news reports are shown:

Sample text 2:

Six officers were feared dead after a Russian military helicopter crashed northwest of Moscow, Interfax news agency said on Thursday. A defence ministry spokesman confirmed that a Ka-27 helicopter had come down on Wednesday near Tver, northwest of the capital, but declined to say if there were any casualties.

Sample text 3:

Three Moscow metro workers were injured 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.

A text understanding procedure aimed at summarization must be able to distinguish the different mental spaces from each other, and to make decisions as to their saliency. Redundant or less salient claims embedded in the speech act spaces should preferably not be rendered in the summary. The selection of claims to be rendered in the summary was performed according to the rules below. In the rules we use the functor 'claim' instead for 'event' for the reasons mentioned in section 2.2.: we cannot ascribe truth values to the events mentioned in the news reports. We can only state that a certain statement is mentioned in the text, and decide whether two statements are semantically compatible. Thus, in the rules below, $claim(x) \dot{\cup} claim(y)$ $\dot{\cup} y = \emptyset$ is to be read as 'y and x would exclude each other if they belonged to the same mental space'. The functor 'claim' applies not only to main clauses, but also to subordinated clauses, inclusive conditional clauses and indirect questions. All rules operate within the domain of a single news report.

Some principles for selection of claims to be rendered in the summary:

1) Informatives:

- Neutral, the sender is not marked for high status: *officials said, the news agency reported, reportedly...*
- $\forall x,p,z \{text(z) \wedge x \in z \wedge p \in z \wedge information_source(x) \wedge claim(p)\}$:
 $claims(x,p) \wedge \neg high_status(x) \wedge uses_speech_act_type(x,informative_neutral) \wedge \neg \exists s \neg \exists y$
 $\{s \in z \wedge y \in z \wedge s \setminus x \wedge y = \neg p \wedge information_source(s) \wedge claim(y)\}: (claims(s,y) \vee$
 $claims(s,p)) \rightarrow to_be_reported(p)$.
 (A claim p introduced by a neutral informative is rendered in the summary; the source is omitted if there are no denials or confirmations of p in the text and if the source is not marked for high status, like 'President')
- Neutral, the sender marked for high status, and 'declarations': *the President said...the government condemned...*
- $\forall x,p,z \{text(z) \wedge x \in z \wedge p \in z \wedge information_source(x) \wedge claim(p)\}$:
 $claims(x,p) \wedge high_status(x) \wedge (uses_speech_act_type(x,Type) \wedge (Type=informative_neutral$
 $\vee Type=declaration) \wedge \neg \exists s \neg \exists y \{s \in z \wedge y \in z \wedge s \setminus x \wedge y = \neg p \wedge information_source(s) \wedge$
 $claim(y)\}: (claims(s,y) \vee claims(s,p)) \rightarrow to_be_reported(x,p,Type)$.
 (The source is rendered if it is marked for high status)
- Affirmative; confirmations of explicit claims: *Israeli sources confirmed that...*
- $\forall x,y,q,z \{text(z) \wedge x \in z \wedge y \in z \wedge q \in z \wedge information_source(x) \wedge information_source(q) \wedge$
 $claim(y)\}: claims(x,y) \wedge uses_speech_act_type(x,informative_neutral) \wedge claims(q,y) \wedge$
 $uses_speech_act_type(q,informative_positive) \rightarrow to_be_omitted(q,y)$.
 (Confirmations of previous explicit claims are omitted in the summary)
- Affirmative; confirmations of claims that are not explicitly mentioned:

- $\forall x,p,z \{ \text{text}(z) \wedge x \in z \wedge p \in z \wedge \text{information_source}(x) \wedge \text{claim}(p) \}$:
 $\text{claims}(x,p) \wedge \text{uses_speech_act_type}(x, \text{informative_positive}) \wedge \neg \exists s \neg \exists y \{ s \in z \wedge y \in z \wedge s \neq x$
 $\wedge y \neq p \wedge \text{information_source}(s) \wedge \text{claim}(y) \}$: $(\text{claims}(s,y) \vee \text{claims}(s,p)) \rightarrow$
 $\text{to_be_reported}(x,p, \text{informative_positive})$.
 (Both the information source and the claim, including the type of the speech act phrase, are rendered in the summary, if the speech act is a confirmation of a claim not present in the news report)
 - Negative, or neutral followed by denied claims: *The president denied.....said that it is not true...*
 - $\forall x,y,p,q,z \{ \text{text}(z) \wedge x \in z \wedge y \in z \wedge p \in z \wedge q \in z \wedge$
 $\text{information_source}(x) \wedge \text{information_source}(q) \wedge \text{claim}(y) \wedge \text{claim}(p) \wedge y \neq p \}$:
 $\text{claims}(x,y) \wedge \text{claims}(q,p) \rightarrow \text{to_be_reported}(x,y, \text{informative}) \wedge$
 $\text{to_be_reported}(q,p, \text{informative})$. (Both the initial claim and its denial are rendered in the summary together with the information about the senders)
- 2) Utterance refusal, negated speech act phrases, hypotheses, commissives, interpretations: The Israeli sources neither denied or confirmed, the minister did not say, if..., ...declined to say...
- Utterance refusals related to an explicit previous claim:
 - $\forall x,y,p,s,z \{ \text{text}(z) \wedge x \in z \wedge p \in z \wedge y \notin z \wedge s \in z \wedge$
 $\text{information_source}(x) \wedge \text{claim}(y) \wedge \text{claim}(p) \wedge y \neq p \}$: $\text{claims}(s,p) \wedge \neg \text{claims}(x,p) \wedge$
 $\neg \text{claims}(x,y) \rightarrow \text{to_be_omitted}(x)$.
 (Utterance refusals or negated speech act phrases related to an explicit claim are omitted)
 - Utterance refusals not related to an explicit previous claim:
 - $\forall x ,p,z,y \{ \text{text}(z) \wedge x \in z \wedge p \in z \wedge \text{claim}(p) \wedge$
 $\text{information_source}(x) \wedge \text{claim}(y) \wedge y \neq p \}$: $\neg \text{claims}(x,p) \wedge \neg \text{claims}(x,y) \wedge$
 $\text{uses_speech_act_type}(x, \text{utterance_refusal}) \wedge \neg \exists s \neg \exists y \{ s \in z \wedge y \in z \wedge s \neq x \wedge y \neq p \wedge$
 $\text{information_source}(s) \wedge \text{claim}(y) \}$: $(\text{claims}(s,y) \vee \text{claims}(s,p)) \rightarrow \text{to_be_reported}(x,p,$
 $\text{utterance_refusal})$.
 (If a source refuses to confirm/deny a claim that has not been explicitly mentioned in the previous part of the text, the whole speech act is rendered, inclusive the type of the speech act)
 - hypotheses and commissives:
 - $\forall x, p, z \{ \text{text}(z) \wedge x \in z \wedge p \in z \wedge$
 $\text{information_source}(x) \wedge \text{claim}(p) \}$: $\text{claims}(x,p) \wedge \text{uses_speech_act_type}(x, \text{Type}) \wedge (\text{Type} =$
 $\text{commissive} \vee \text{Type} = \text{hypothese}) \rightarrow \text{to_be_reported}(x,p, \text{Type})$
 (Hypotheses and commissives are rendered together with their sources)
- 3) Epistemic spaces: e. g. no one knows if the device was planted deliberately or if it was leftover from New Year's Eve
- $\forall y, p, z \{ \text{text}(z) \wedge y \in z \wedge p \in z \wedge$
 $\text{claim}(y) \wedge \text{claim}(p) \}$: $y \neq p \wedge \neg \exists x \{ x \in z \wedge (\text{information_source}(x) \wedge$
 $(\text{claims}(x,p) \vee \text{claims}(x,y))) \}$ $\rightarrow \text{to_be_reported}(y, \text{hypothese}) \wedge \text{to_be_reported}(p, \text{hypothese})$.
 (If two claims would exclude each other in the same mental space, and if no source in the text takes responsibility for any of these claims, both claims are to be rendered as hypotheses).

2.4. The way from an input text to a summary

The core part of the procedure for identification and selection of mental spaces was implemented in Prolog. For interface between the input file and the understanding and generation procedures as well as for the interface between the generator and the final output, we used procedural programming languages. Figure 3 depicts the stages between the input text and the written and/or spoken output. The syntactic and semantic analysis was performed by means of three knowledge sources: an English lexicon, a set of subdomain specific templates and lists of subdomain specific key words and phrases. All knowledge sources have been constructed in a computer-aided way on the basis of corpus investigation.

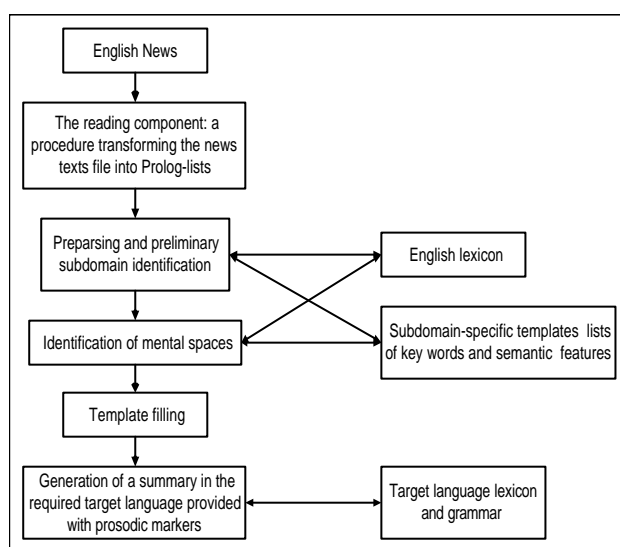


Fig. 3. The way from input to output

The small (ca 3000 items) but grammatically and semantically informative lexicon has been built by extracting knowledge from informants by a method inspired by the BOAS system (Nirenburg and Raskin 1988, Nirenburg 1996, Oflazer et al. 2001). Verbs were provided with quite elaborated valence frames, and nouns were classified by sets of semantic features that enabled both noun-noun and noun-verb connections in the lexical network. The lexicon and template development was no doubt time consuming, but – if the text did not contain more than 30% unknown words – the results were of a quite high quality. A sample output from the analysis and extraction component (a filled template corresponding to sample text 2) is shown in Figure 4. The functor ‘m’ marks the interlingua codes in the Prolog lexicon.

```

[sender([interfax,news,agency,on,thursday]),
  sact([say,speech]),
place,
  [placep([m([northwest])),of,m([moscow])],
trajector,
  trajector (m([helicopter]),means_of_transportation(air)),
dead,
  [m([six]),m([officers]),verb([feared],m([feared]),[verb])],

sender([m([indef]),m([defense]),m([ministry]),m([spokesman])]),
  sact([confirm,speech,pos]),
time,
  [m([wednesday])],
place,
  [[m([near]),m([tver])],[m([northwest])],[the,m([capital])]],
trajector,
  trajector (m([helicopter]),means_of_transportation(air))).

```

Fig. 4. The result of the analysis of Sample text 2, performed by the system prototype

The last step consisted of target language summary generation. The main advantage of multilingual generation compared to machine translation was a greater possibility of achieving more idiomatic lexical and phrasal choices. The target generators made use of standardized ways of presenting certain situations in Swedish/Danish/Polish news reports, which resulted in structures that were functionally, although often not syntactically, equivalent to the English input, for example:

Swedish:

Inga personer skadades, och det rapporterades enbart obetydliga materiella skador
 no persons injury+pass and it report+pass only less important material damage
 'There were no injuries and little damage'

The main problems that the prototype suffered from were the lexical limitations. We could also conclude that the templates used in the prototype were a little too rigid and that more flexibility and robustness would be desirable. However, we did not want to refrain from high quality demands.

3. Towards a more robust system – developing lexical resources

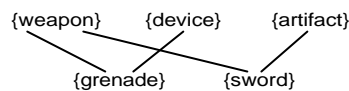
3.1. Introducing WordNet into the system

WordNet (Miller 1990, 1995) is a well-known lexical database aimed at encoding psycholinguistically relevant connections between lexical entries: synonymy, antonymy, hypo/hypernymy ('A is an X'), and meronymy ('A is a part of X'). The general problem with the use of WordNet in NLP is, however, its frequently discussed and criticized fine-grainedness (Véronis & Ide 1990, Ide and Véronis 1998, Montoyo et al. 2001, Mendes and Chaves 2001 and many others). The system of 'synsets' employed in WordNet makes it scarcely possible to distinguish between polysemy and homonymy. As pointed out by Mendes and Chaves (2001), who base their analysis on Pustejovsky (1991b, 1995) and Wierzbicka (1994), the hypo/hypernymy hierarchies in WordNet do not express the differences between

formal, telic, and constitutive *qualia*. This distinction can be illustrated by examples 19-21 in Mendes and Chaves (2001:111):

- (19) {dog} [[Formal=canine(x) ∧ domestic(x)]...]
- (20) {police dog} [[Formal=dog(x)...], [Telic=support(e,x,police_officer)],...]
- (21) {German shepherd} [[Formal=dog(x)...], [Const=has_fur(x,y) ∧ short(y) ∧ colour(y,z) ∧ dark(z)...]...]

Especially the distinction between formal and telic qualia is of importance for text understanding. A good example for the difference between telic and formal hypernyms is shown in Mendes and Chaves (2001:110):



Here, *grenade* and *sword* inherit the telic role from *weapon*, while the formal roles are obtained from *device* and *artifact*.

Telic hypernyms are much more frequently used – both in spoken and written utterances – as anaphoric phrases than formal hypernyms: it is much more natural to refer to a *grenade* as a kind of *weapon* than as a kind of *artifact*. The organisation of our prototype lexicon was based mostly on *qualia* and on telic hyponymy relations. Inspired by the work of Mendes and Chaves (2001), and also by the ideas by Montoyo and al. (2001), who propose labelling WordNet synsets by terms used in standard news agencies classification system, we decided to enrich the WordNet structures by identifying the telic hypernyms that were most salient in the domain of news reports. Figure 5 shows fragments of the original WordNet hierarchy. Nodes that correspond to the highest number of words grouped within one category in our prototype lexicon are highlighted.

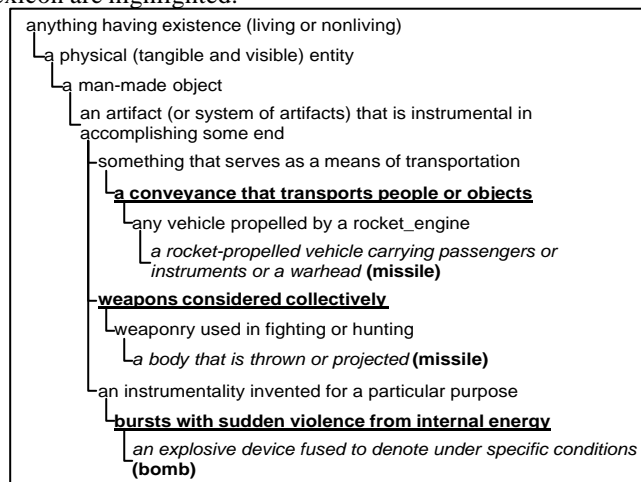


Fig. 5. The place of the nouns *bomb* and *missile* in the Word Net hierarchy

As can be seen, *bomb-missile* are not sister nodes in WordNet, and there are different distances between the nouns in question and the nodes that are of crucial importance

for news reports understanding. Replacing the prototype lexicon by WordNet thus required identification of those semantic categories that had proved successful in the prototype. The results of implementing connections between WordNet and the prototype lexicon are shown in Table 1. As is clear from the data in Table 1, the relevant nodes are mostly telic and sometimes constitutive hypernyms.

Table 1. The results of extraction and enrichment of WordNet (WN) noun hierarchies

WN synset definition	WN entries representing the synset	Sample hyponyms in WN	Semantic category in the prototype
a group of people who work together	organisation	army, military, troops, party	group of people
group of people willing to obey orders	force personnel	army, military, troops, police ...	armed forces
a conveyance that transports people or objects	vehicle	tank, missile, rocket..	means of transportation
a vehicle used by the armed forces	military vehicle	tank	destruction as function
the act of terminating a life	kill killing	suffocation	cause of injuries and/or death
any road or path	way	street	place-path
a human being	human, person...	passenger	human
a structure that has a roof and walls ...	building edifice	restaurant	place, three-dimensional (3D)
a speech act that conveys information	informing making known	report	speech act, informative
a speech act that conveys information	informing making known	report	source of information
a large indefinite location on the surface of the Earth	region	region, town, country	place, 2- or 3 D, +conventional borders
a means for communicating	medium	radio, newspaper	source of information
the solid part of the earth's surface	dry land, earth ground, land	peninsula	place, 2D, +natural borders
a person who rules, or guides other people	leader	President, Pope	human +high status
a person who does harm to others	offender wrongdoer	murderer	human +pejorative
impairment of normal physiological function	illness malady	flu, anthrax	disease
the act of changing the unity or wholeness	change	explosion	explosion
bursts with sudden violence from internal energy	explosive device	bomb	destruction as function

Table 1. contains the most salient categories. In the actual database, some more relations are implemented, among them some meronymy relations that are not coded

in WordNet, like ‘a street is a part of a town’, ‘a soldier is a part/member of armed forces’ etc.

3.2. The summarization process in interaction with the extended lexical resources – an example

The extension of lexical resources increased the recall rate from ca 68% to about 86%. It is worth noting that before the connections between the prototype lexicon and WordNet had been implemented, attempts to parse news reports with Word Net as the only lexical resource were strongly discouraging. On average, we got 21 possible interpretations of a sentence. Preferring the telic hypernyms in WordNet and focusing on the semantics of nouns rather than on the semantics of verbs gives about 90% unambiguous results. Below, we show how the current system works by presenting the analysis of Sample text 1.

```
% BEIT JALA, West Bank (CNN) - Israeli troops pulled out of Beit Jala before dawn on
% Thursday, leaving the Palestinian town quiet
[[main_information_source(1, CNN),
  information_source(1, CNN),
    place(region_conventional_borders('West Bank', 'Beit Jala')),
      claim(1,trajector(armed_forces([nationality('Israel'), troops])),
        action(motion), direction(out_of(tr,lm)),
        place(region_conventional_borders(['West Bank', 'Beit Jala'])),
        time([before,dawn,on,'Thursday']),
        damage(none),dead(none),injuries(none))],
  % amid reports of fresh violence in other West Bank towns.
  [information_source(2, reports),
    claim(2,tr(none),action(destruction_as_function(violence)),
      place(region_conventional_borders(['West Bank', towns])),
      time(['Thursday']),
      damage(none),dead(none),injuries(none))],
  % The Palestinians said the Israel Defense Forces had staged incursions into Hebron, killing
  % one and injuring 16 others, and Tulkarem, killing one and injuring 10.
  [information_source(3, nationality('Palestinians'),
    claim(3,trajector(armed_forces([nationality('Israel'), forces])),
      action(military_action), direction(into(tr,lm)),
      place(['Hebron']),
      time(['Thursday']),
      damage(none),dead(one),injuries(16)),
    claim(4,trajector(armed_forces([nationality('Israel'), forces])),
      action(military_action), direction(into(tr,lm)),
      place(['Tulkarem']),
      time(['Thursday']),
      damage(none),dead(one),injuries(10))],
  % The Israel Defense Forces (IDF) had no immediate comment on the accusation that troops
  % had entered Tulkarem, and strongly denied there was an incursion at Hebron.
  [information_source(4, armed_forces([nationality('Israel'), forces])),
    claim(5, negation(claim(3)))]].
```

Fig. 6. The result of the analysis of sample text 1 using re-classified WordNet nodes and fragments of the prototype lexicon. The original text is shown as Prolog comments

The first information source in the text (CNN) is identified on the basis of the formatting convention: a CNN news report is, in the convention of today, always introduced by one or two geographical names and the name of the news agency. Information source 2 ('reports') is found thanks to the fact that the noun *report* is placed under the node 'a speech act that conveys information' in WordNet. In our classification (cf. Table 1), this node is labelled both as 'Speech Act' and as 'Information Source'. The identification of the remaining information sources is due

to speech act verbs and phrases (*said, denied, had no comment*) in interaction with the WordNet node labelled ‘Group of people’ and its hyponyms. The phrase which follows ...*had no immediate comment* is not represented in the output, according to the rules dealing with utterance refusal phrases presented in section 2.3.

The slots labelled ‘place’ are in the current example filled mainly using contextual cues. The only geographical name in the text that can be found in WordNet is *West Bank*. The others (*Beit Jala, Hebron, Tulkarem*) are classified as places either on the basis of already mentioned formatting conventions, or by syntactic and semantic features of the context (objects of prepositions that follow phrases denoting actions are classified as places if no other category has been found).

The slot ‘action’ in the Prolog structure above does not contain the exact characteristics of the verb that occurs in the original text. We refrain from placing the verb *stage* (in *staged incursions*) in the sentence representation. We confine ourselves to represent the noun *incursion*, as it is labelled as a ‘military action’ in the lexical database. Similarly, we do not represent the verb *pull* (in *troops pulled out of*), but only the concept of motion and direction – in the Prolog notation: ‘action(motion), direction(out_of(tr,lm))’, which is to be read: there was a motion, and the trajector moved out of the landmark. Such underspecified representations of verb phrases are very useful for idiomatic target text generation. They allow the generator to choose language-specific phrases that express movement of certain objects (or, using Mel’cuk’s (1988) terminology, they allow the realization of language-specific lexical functions). This is shown below by the Swedish and Polish outputs rendering the content of the initial phrases of the text:

Swedish:

Israeliska trupper tågade ut ur Beit Jala
Israeli+pl troops marched out of/left Beit Jala
(*tågade ut ur* instead of default word-for-word-translation **drog ut av.*)

Polish:

Wojska izraelskie wycofaly sie z Beit Jala
Troops Israeli backed out from Beit Jala
(*wycofaly sie* instead of word-for-word-translation **wyciagnely* or **wyciagaly*).

The main advantage of the summarization+multilingual generation concept, vs. the possibility of idiomatic lexical and phrasal choice, is thus preserved, and not negatively affected by the increased degree of robustness of the system as a whole. Even this fact is due to the re-categorisation of WordNet nodes.

3.3. Testing

The first system test was performed on 12 CNN news reports. The average number of sentences per text was 17 (the longest text consisted of 37 sentences, the shortest of 9). As lexical resources we used the noun part of WordNet, the re-categorization table shown in the previous section, and the verbs, adjectives and closed-class items from the prototype Prolog lexicon. In its current shape, the system covers only the domain of military conflicts, terrorist attacks and natural disasters, so this domain functioned as the stable topic of interest in the filtering task. We tried to evaluate the results

along the main lines suggested for DUC (Harman & Over 2002) and TREC (Voorhes and Harman 2001).

The filtering task was, at the first glance, not very challenging, since only one text in the collection dealt with a topic not belonging to the current domain of interest (it was a text about the Pope's journey to Azerbaijan). However, to filter out this text would be problematic, if the mental-space model had not been employed, since the text included fragments containing references to military conflicts (*Azerbaijani President (...) hopes the papal visit will focus attention on the country's conflict with Armenia (...) A cease-fire ended fighting in 1994 after about 30,000 people were killed...*):

The classification of the text was correct due to the fact that no references of this kind were found in the base space, and that the phrases referring to war were embedded either in a hypothetical mental space, or in a non-actual temporal space.

The system produced 11 summaries that were in 98% grammatically correct acceptable cohesion and organization rate (about 75%). Most (70%) of the cohesion and organization shortcomings, and most wrong interpretations occurred in one of the summaries, generated on the basis of a text with an extremely mental-space structure and many syntactic ambiguities. A fragment of this text is shown below:

Sample text 4:

Vice President Dick Cheney lashed out Wednesday at lawmakers who have questioned whether the government missed clues prior to the September 11 terrorist attacks, calling their comments "despicable" and "outrageous" (...) Few lawmakers have actually suggested President Bush knew about the September 11 attacks in advance, but numerous members of the House and Senate (...) have questioned whether the government failed to recognize ...

The remaining 10 summaries were of a quite good quality. Below, we show a fragment of a representative input text and the corresponding part of the summary (fragments of the input texts that have been interpreted as less important or redundant are marked).

Input text:

RAMALLAH, West Bank (CNN) -- Palestinian leader Yasser Arafat said Thursday that elections as part of a reform of the Palestinian Authority will be held this winter, whether or not Israeli forces withdraw from the Palestinian territories.

~~That represented a change of course from Arafat, who said last week that no elections would be held until the Israelis pulled back. Shortly after Arafat's announcement, a committee he had appointed to set up elections resigned, according to Israel Radio, because Arafat would not agree to a specific date for the elections. Other Palestinian leaders said the resignations were a procedural matter.~~ Arafat also condemned Wednesday's suicide bombing in the Israeli town of Rishon Letzion . Two Israelis were killed and at least 37 others wounded when the bomber detonated explosives in the center of a crowded pedestrian district.

~~The terror attack marked the second time in two weeks a suicide bombing directed at civilians has rocked Rishon Letzion, a town about 15 miles southeast of Tel Aviv. On May 8, a suicide attack at a pool hall killed 15 people and wounded dozens of others.~~

~~"Suddenly there was an explosion," 16 year old Shmuel Voller told The Associated Press on Wednesday.~~

The bombing occurred on Rothschild Street in the heart of the town around 9:15 p.m. (2:15 p.m. ET).

Summary:

RAMALLAH, West Bank (CNN) -- Palestinian leader Yasser Arafat said Thursday that elections as part of a reform of the Palestinian Authority will take place this winter, whether or not Israeli forces withdraw from the Palestinian territories.

On Wednesday, a suicide bombing took place in the Israeli town of Rishon Letzion, on Rothschild Street in the center of a crowded pedestrian district, around 9:15 p.m. (2:15 p.m. ET). Two Israelis were killed and at least 37 others wounded. Arafat condemned the attack.

3.4. Some remarks on the Polish lexicon: a strategy for developing a module for a target language typologically different from the source language

The main target language of the prototype system was Swedish, the second best developed – Danish. In the new version of the system – the one connected to WordNet – we decided to focus on a language that is typologically more different from English. We chose Polish partially for practical reasons (the language competence of the staff), but also for linguistic reasons – the rich morphology of the Slavic languages, especially the case system and the verbal aspect and Aktionsart system, constituted an interesting challenge.

The work on the Polish module started by developing a closed-class lexicon, where special attention has been paid to the semantics of prepositions. Since most of the Slavic prepositions also function as verbal prefixes, modifying the values of aspect and *or*/Aktionsart (Janda 1986), encoding the selectional restrictions of prepositions and their ‘image-schemata’ was considered to be of crucial importance. The hypotheses about the semantics of prepositions were continuously tested by means of a domain-restricted English-to-Polish machine translation system (translation of weather forecasts, fragments of news reports, and sample sentences from bilingual dictionaries). The analysis of prepositions has shown that they are not only of importance for the verbal paradigms (which is a well-known fact), but that they also provide valuable cues for the semantic classification of Polish nouns. For example, the Polish preposition *na* (translated as ‘on’, ‘at’, ‘in’, ‘onto’, ‘into’, ‘in’, ‘for’) turned out to select for nouns that are conceptualised as either one- or two-dimensional objects. In prototypical cases, this selectional restriction coincides with the one of English ‘on’ (*na dachu* – ‘**on** the roof’, *na talerzu* – ‘**on** the plate’), but the conceptual patterns differ in most conceptual metaphors (*na ulicy* – ‘**in** the street’, *na granicy* – ‘**at** the border’, *na Litwie* – ‘**in** Lithuania’ etc.). The analysis of other prepositions and verbal prefixes confirmed that the distinction between one- or two-dimensional objects on the one hand, and three-dimensional ‘containers’ on the other hand is at least as important for the grammatical structure of Polish as the mass/count distinction is for the Germanic languages. This distinction between ‘flat objects’ and ‘containers’ cannot be established on an ontological basis alone, as the conceptualisation patterns are clearly language specific. For example, a street is a ‘container’ in English, German and French (*in the street*, *in der Strasse*, *dans la rue*), but a flat object in Polish and Russian (*na ulicy*, *na ulice*). Another semantic distinction (in contrast to the +/-container feature, well described in linguistic literature) that is grammaticalized in Slavic is the difference between animate and inanimate objects.

The Polish animacy system comprises an additional complication: the distinction between inanimate, animate, and ‘superanimate’ (male humans) nouns. All those facts lead to the conclusion that it is not sufficient to restrict the work on a target lexicon – especially when there are considerable typological differences between the source and the target – to connecting the target entries with the source entries and copying the source semantic classifications into the target lexicon. What is a conceptual ‘container’ in English, can be a conceptual two-dimensional object in Polish, and vice versa. Distinctions that may be irrelevant for the source language system (such as the superanimate-animate-inanimate feature) can be crucial for the target language. Other distinctions (e.g. the mass/count feature) may apply to nouns in one language (English), and to verbs in another language (the Slavic aspectual system seems to be based on conceptual features related to the mass/count difference). Our ongoing work with the Polish lexicon and generator indicates that the telic hyper/hyponymy hierarchies can quite safely be copied from WordNet to the Polish lexicon, but that language specific classification with respect to conceptual categories like dimensionality, plexity (i.e. the mass/count and uniplex/multiplex distinctions), and animacy is necessary.

4. Conclusions

Automatic understanding of text genres in which different information sources are cited or employed in some other way, or where different senders interact, requires an identification of speech acts and ‘mental spaces’. Our work on mental space identification considers news reports, but can be applied even to other domains, like dialogs, e-mail correspondence, and even scientific texts, where different sources are cited and different points of view are taken into account.

An appropriate semantic classification of the lexical entries is a prerequisite for successful text understanding. Hypo/hypernymy relations are very useful, but only if the distinction between telic, constitutive and formal hyponymy is encoded. In the domain of news reports, telic hyponymy is especially useful for identification of semantic roles and anaphora resolution. Telic hyponymy seems also to be the lexical relation that can most safely be copied from a source lexicon into target lexicons, while constitutive hyponymy relations display considerable cross-linguistic variation and must be treated with caution, especially when typologically different languages are concerned.

Extending the classification of lexical entries in order to cover more domains is one of the future goals of the project. Up to now, the re-classification and re-grouping of WordNet has not been automatized. A tool for semi-automatic re-grouping WordNet entries with respect to different topical domains is under development.

The noun part of WordNet has proved to be useful for summarization, if domain-related modifications are added to the original structure of the database. The verb part of WordNet, however, contains a lot of vague definitions that lead to too many ambiguous interpretations of the input text. For this reason, the verb part of WordNet has not been used in the text understanding procedure, but an extension of the topical domain would require a larger verb lexicon than the currently utilized Prolog verb list.

A partial solution to this problem is to identify the most relevant verb ‘synsets’ in WordNet while developing the target (Polish) lexicon: the ‘synset’ linked to Polish entries can be given priority in the text understanding procedure. We also consider introducing some form of supervised learning, i.e. learning from user feedback.

References

- Alexandersson, J., Poller, P. and Kipp, M. (2000) *Generating Multilingual Dialog Summaries and Minutes*. In Wahlster, W. (Ed): *VerbMobil: Foundations of Speech-to-Speech Translation*, Springer-Verlag, Berlin, pp. 507-518.
- Austin, J.L. (1962) *How to do things with words*. Oxford: Clarendon Press.
- Beale, S., Nirenburg, S. and Mahesh, K. (1996). *HUNTER-GATHERER: Three Search Techniques Integrated for Natural Language Semantics*. AAAI/IAAI, vol.2, pp.1056-1061.
- Carnap, R., *Meaning and Necessity*, Chicago: U. Chicago Press, 1947
- Dancygier, B. and Sweetser, E. (1996) *Conditionals, Distancing, and Alternative Spaces*. In A.E. Goldberg (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.
- Farwell, D. and Helmreich, S. (1997) *What floor is this? Beliefs and Translation*. Proceedings of the 5th International Colloquium on Cognitive Science, pp. 95-102.
- Farwell, D. and Helmreich, S. (2001) *Towards Pragmatics-based Machine Translation*. MT Summit VIII: MT 2010 – Towards a Road Map for MT. Proceedings of the Workshop, Santiago de Compostela, Spain, pp. 22-25.
- Grice, H.P. (1975) *Logic and Conversation*. In P.Cole and J.Morgan (Eds): *Syntax and semantics*. Vol.3: *Speech acts*. New York: Academic Press, pp. 43-58.
- Hahn, U., and Reimer, U. (1999) *Knowledge-Based Text Summarization: Saliency and Generalization Operators for Knowledge Base Abstraction*. In Mani, I. and Maybury, M.T.(Eds): *Advances in Automatic Text Summarization*. Cambridge, Massachusetts, London, England: The MIT Press.
- Harman, D. and Over, P. (2002) DUC: an Evaluation on Summarization. Retrieval Group, Information Access Division, National Institute of Standards and Technology, <http://www.darpa.mil/ipto/research/proceedings/tides2001jul/Enabling-summarization-Harman.pdf>
- Hintikka, J. (1962) *Knowledge and belief. An introduction to the logic of the two notions*, Cornell University Press, New York.
- Ide, N. and Véronis, J. (1998) *Introduction to the Special Issue on Word Sense Disambiguation: The State of the Art*. *Computational Linguistics* 24 (1), pp. 1-40.
- Jackendoff, R.S. (1983) *Semantics and Cognition*. Cambridge, MA: MIT Press.
- Janda, L. (1986) *A semantic analysis of the Russian verbal prefixes*. *Slavistische Beiträge*, Verlag Otto Sagner, München.
- Lakoff, G. (1993) *The Contemporary Theory of Metaphor*. In A. Ortony (Ed.): *Metaphor and Thought*, 2d ed. Cambridge: Cambridge University Press.
- Langacker, R. (1991). *Concept, Image, and Symbol*. *The Cognitive Basis of Grammar*. Berlin/New York: Mouton de Gruyter.
- Lee, M. and Wilks, Y. (1996) *An ascription-base approach to Speech Acts*. Proceedings of CoLing’96, pp. 699-704.
- Lehnert, W. (1999). *Plot Units: A Narrative Summarization Strategy*. In I. Mani and M.T. Maybury, (Eds): *Advances in Automatic Text Summarization*. Cambridge, Massachusetts, London, England: The MIT Press, pp. 177-213

- Mani, I. and Maybury, M.T., eds. (1999) *Advances in Automatic Text Summarization*. Cambridge, Massachusetts, London, England: The MIT Press.
- Mel'cuk (1998) Collocations and Lexical Functions. In A.P. Cowie (ed.): *Phraseology. Theory, Analysis, and Applications*, Oxford: Clarendon Press, 23-53.
- Mendes, S. and Chaves, R.P. (2001) *Enriching WordNet with Qualia Information*. In Workshop on WordNet and Other Lexical Resources: Applications, Extensions and Customizations at NAACL 2001, pp. 107-112.
- Miller, G.A. and Johnson-Laird, P.N (1976) *Language and Perception*, Cambridge University Press, Cambridge-London-Melbourne.
- Miller, G.A. (Ed), (1990). *WordNet: An On-Line Lexical Database*. Volume 3(4) of the International Journal of Lexicography. Oxford University Press.
- Miller, G.A. (1995). *WordNet: An on-line lexical database*. Communications of ACM, 38(11).
- Montoyo, A. and Palomar, M. (2001) *WordNet Enrichment with Classification Systems*. In Workshop on WordNet and Other Lexical Resources: Applications, Extensions and Customizations at NAACL 2001, pp. 101-106.
- Nirenburg, S. (1996) *Supply-side and demand side of lexical semantics*. Introduction to the Workshop on Breadth and Depth of Semantic Lexicons at ACL'96.
- Nirenburg, S. and Mahes, K. (1997) *Knowledge based systems for Natural Language Processing*. The Computer Science and Engineering Handbook, pp. 637-653.
- Nirenburg S. and Raskin V. (1988). *Project Boas: "A linguist in a Box" as a multi-purpose language resource*. In COLING-ACL'98, pp. 975-979.
- Oflazer K., Nirenburg S. and McShane M. (2001) *Bootstrapping Morphological Analyzers by Combining Human Elicitation and Machine Learning*. Computational Linguistics, 27(1), pp. 59-86.
- Pustejovsky, J. (1991a) *The Syntax of Event Structure*. Cognition, 41, 47-81.
- Pustejovsky, J. (1991b) *The Generative Lexicon*. Computational Linguistics, 17, pp. 409-41.
- Pustejovsky, J. (1995) *The Generative Lexicon*. The MIT Press, Cambridge, MA.
- Sanders, J. and Redeker, G. (1996). *Perspective and the Representation of Speech and Thought in Narrative Discourse*. In: Fauconnier, G. and Sweetser, E. (Eds): Spaces, worlds and grammar. The University of Chicago Press, Chicago and London, pp. 290-317.
- Searle, J.R. (1969) *Speech acts*. Cambridge: Cambridge University Press.
- Somers, H., Tsuji, J. and Jones, D. (1990) *Machine translation without a source text*. Proceedings of CoLing'90. Helsinki, vol.3 pp. 271-277.
- Sweetser, E. and Fauconnier, G. (1996) *Cognitive links and domains: basic aspects of mental space theory*. In: G. Fauconnier and E. Sweetser (Eds): Spaces, worlds and grammar. The University of Chicago Press, Chicago and London, pp. 1-28.
- Talmy, L. (1988) *The Relation of Grammar to Cognition*. In B. Rudzka-Ostyn (Ed.): Topics in Cognitive Linguistics, Amsterdam/Philadelphia: John Benjamins, pp.165-205.
- Veronis, J. and Ide, N.M. (1990) *Neural networks generated from dictionaries*. Proceedings of Coling 90. Helsinki, vol.2, pp. 389-394.
- Voorhees, E.M. and Harman, D. (2001) Overview of TREC 2001. http://trec.nist.gov/pubs/trec10/papers/overview_10.pdf
- Wahlster, W. (2000) *Mobile Speech-to-Speech Translation of Spontaneous Dialogs: An Overview of the Final Verbmobil System*. In W. Wahlster (Ed): Verbmobil: Foundations of Speech-to-Speech Translation, Springer-Verlag, Berlin, pp. 3-21.
- Wierzbicka, A. (1994) *Apples are not a 'Kind of Fruit'*. American Ethnologist, 11, pp. 313-328.
- Wilks, Y. (1985) *Relevance, points of view and speech acts: An artificial intelligence view*. Technical Report MCCS-85-25, New Mexico State University.
- Zechner, K. (1995) *Automatic Text Abstracting by Selecting Relevant Passages*. MSc Dissertation in Cognitive Science and Natural Language, University of Edinburgh