THE SEMANTIC STRUCTURE IN COMPARISON WITH OTHER SEMANTIC REPRESENTATIONS

Johannes Müller (1), *Holger Stahl* (2)

(1) Institute for Human-Machine-Communication, Munich University of Technology, D-80290 Munich, Germany
mue@mmk.e-technik.tu-muenchen.de
(2) Rohde & Schwarz GmbH & Co. KG, Muehldorfstrasse 15, D-81671 Munich, Germany
holger.stahl@rsd.de

ABSTRACT

This paper describes some common semantic representations and compares them with the *semantic structure*, which is introduced as a semantic representation level within a speech processing system. The semantic structure is especially designed to facilitate semantic decoding using exclusively stochastic knowledge within a maximum-a-posteriori decoding algorithm.
Keywords: semantic structure, semantic representation, speech understanding, semantic decoding.

1 INTRODUCTION

The central and most important criterion of speech and language lies in its significance, which enables the transmission of information by speaking or writing. Otherwise, speech and language would be completely useless, since nobody would normally speak purely insignificant sounds or write an insignificant combination of letters.

The human perception of speech and language normally uses the auditory and the visual sensory channel. In exceptions, the tactile channel is used, e.g. for braille. The received acoustic, optical or haptic information (quantitative information aspect [7][12]) is only relevant for transmission or storage, but does not satisfy the interest of the human information recipient. For the latter, the meaning or the semantics is the only worth knowing information (qualitative information aspect).

1.1 What Means 'Semantics'?

The term 'semantics' generally refers to the significance of signs and combinations of signs[1]. In the following, only linguistic signs (words or word chains as representations of spoken or speakable phonemes) might be exclusively considered. Each sign corresponds to expression and content. In this case, the expression includes the correct word formation (morphology) as well as the correct and understandable word pronunciation (phonology). The linguistic sign does not yet combine a thing and a name with each other, but it combines a conceptual and an acoustic imagination. The conceptual imagination is not an object itself but an abstraction of such conceivable objects. The acoustic imagination is not a spoken phoneme chain, but a psychological imagination of such a sound combination. Within humans, acoustic and conceptual imaginations are tightly connected by association that he/she is able to switch between these imaginations within fractions of seconds[2]. Even if the received speech is overlaid by background noise, grammatical mistakes or spontaneously spoken effects, a human information recipient can easily develop a conceptual imagination. This changeover causes greatest difficulties for a computer application, namely for the following reasons:

- The meaning of one or several words can be quite ambiguous. The required unambiguity occurs by including the context. This ambiguity can be explained by the "context sensitivity of language" since the pragmatic environment of a word is very often required for its correct semantic decoding. Each human being must acquire that enormously great knowledge base within a lifelong learning process by upbringing, education, experience, observation, interest, and instruction.
- Robust speech recognition is enormous computing and storage extensive and is not sufficingly mastered at present. Occurring recognition errors can result in an incorrect semantic decoding. Background noise, spontaneously spoken effects or incomplete vocabulary usually cause further recognition errors.
- The acquisition and representation of linguistic knowledge is yet a scientifically incompletely investigated field. In this context, it is not clear, how the human brain stores received information.

1) An image or a figure can have significance (in the sense of semantics), too.

2) As an example, the word "chair" might be considered: As soon as a human hears the corresponding phoneme chain, he/she can switch to the conceptual imagination. The imagination, how a thing has to look like and what is its nature to be a chair, is very easy for every human. If a human sees or feels any chair, the changeover to the acoustic imagination is easy. He/she is immediately able to imagine the respective phoneme chain, to pronounce it correctly or to find synonyms as "seat" or "bench". Even further associations are formed: How heavy is a chair, for what purpose it can be used, whether it is aesthetic or comfortable, etc.

- Language production only succeeds within limited domains. Mostly, distributed naturally spoken utterances are not really online generated, but retrieved from a database by means of simple inference.

A system, which should understand all imaginable spoken inputs and should correctly react on them, appears impossible at the present state of art. If therefore a system should be able to automatically understand speech, then all spoken user inputs should be within a clearly restricted domain. This indeed means a tremendous simplification, however, it is accepted without any problems since for a graphic application or a train schedule information, utterances are to be expected only within the respective domain.

2 CLASSIC SEMANTIC REPRESENTATIONS

Showing the meaning of language or speech is not only a formal-theoretic discipline, but also necessary as intermediate level within a speech understanding [15] or even speech translating [10] application. Linguistic techniques on one hand, mathematic-logic techniques on the other hand are well established yet. For the representation of the semantics, Winston defined some instructive distinguishing criterions [18]:

- **Institutional based semantics:** An informative description of the facts is available which was not made among formal-logic conditions and does not contain any firmly defined semantics.
- **Equivalence semantics:** The descriptions of the facts are correlated with descriptions of another representation, which has a firmly defined semantics, for example the transcription by means of predicate logic. (The transcription by a formal-logic, syntactic-semantic representation within a speech understanding application would be in the sense of such an equivalence semantics.)
- **Procedural semantics:** If there is an amount of programs, which work on account of the descriptions within the representation language, then the semantics is defined by the process or the result of these programs. (The reaction of a speech understanding application would be exactly in the sense of a procedural semantics: The result, which is shown on the screen, is the semantics of the utterance.)
- **Descriptive semantics:** The semantics is represented by descriptions in natural language.

In the sense of the equivalence semantics, a formal-logic representation level has to be found, in order to reflect the significance of spoken inputs. Semantically equivalent word chains (e.g. *"move the bullet to the left"* or *"please move the sphere to the left side"*) should correspond to identical semantic representations. For simplifying purposes, this representation level might be completely independent from the current application status, i.e. from the pragmatic influence, and independent from preceding or succeeding utterances. Whether the sphere, which should be moved to the left, really exists within the pragmatic context, is absolutely irrelevant for the semantic representation – only the utterance itself is crucial.

Classic notations of the semantics or the structure of speech and language are predicate logic, semantic networks and valence/dependence grammar, which are shortly described in the following chapters for the introduction of the semantic structure.

2.1 Predicate Logic (PL)

PL can excellently be used to represent natural spoken information. In this paper, only first order PL with the following inventory is described:

- Logic constants
  – junctors: $\neg$ (not), $\wedge$ (and), $\vee$ (or), $\rightarrow$ (results in)
  – quantors: $\forall$ (for all), $\exists$ (for an existing)
  – identity signs
- individual variables a, b, c, d, ...
- non-logic constants
  – individual constants
  – n-place relational constants for $n \geq 0$

During the simplest form of semantic analysis of linguistic patterns, each word is assigned to a non-logic constant of a certain type [4]. The word category (e.g. noun, verb, adjective) is completely independent of the resulting logic type. For example, the object *"red sphere"* is assigned to the individual variable a, which is specified by the connected one-place relational constants:

$$(\text{sphere}(a) \wedge \text{red}(a)).$$

The PL representation of the utterance *"move all red spheres five centimetres to the right"* is:

$$\text{move}(a,b) \rightarrow \forall a(\text{sphere}(a) \wedge \text{red}(a)) \wedge \text{toRight}(b) \wedge 5\text{cm}(b)$$

With these examples, the transformation of a linguistic statement into the corresponding predicate logic notation seems to be relatively easy. By the help of the PL-notation, some specific tools for the further formal-logic evaluation (induction, deduction, composition, substitution) can be set, if this should be required. Since there is not any direct relation to the respective word chain, it is not possible to estimate a probability (see chap. 3) for the correlation between any PL-notation and any word chain.

2.2 Semantic Networks (SNs)

SNs are well established in Artificial Intelligence for representing spoken or written knowledge [18]. The syntax of a SN is quite simple. There are

- objects, which are shown by named rectangles, being the nodes of the SN,
- relations between these objects, which are shown by named arrows, being the edges of the SN, and
- generalized statements (e.g. *"all"*, *"each"*), which take a SN as formula, referring to all imaginable cases.

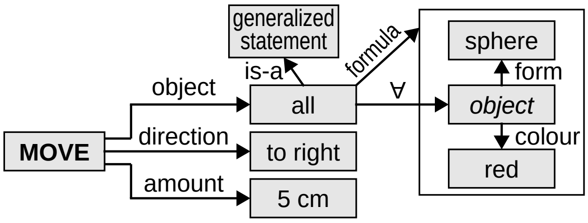

Fig. 1: SN representation of the utterance
"move all red spheres five centimetres to the right"

Fig. 1 gives an idea about the hierarchy, starting form the central MOVE-object to the direction of the edges. The kind and the number of object's specifications are flexible and optional.

2.3 Valence / Dependence Grammar (VDG)

In contrast to constituent grammars or phrase structure grammars [3][6], the VDG goes further than the criterion of the syntactic form by showing that linguistic elements are dependent on each other within a sentence. It states that sentences (in our case word chains) have a hierarchic and non-linear structure fixed by dependence relations [2]. This hierarchic sentence structure is not only built by syntactic constraints, but also by semantic characteristics of the words. The central assumption is that the verb expresses the whole process, so it is treated as the centre of the sentence. Certain successors can depend upon the central verb. This capability can be compared to the valence of chemical elements. Verbs can be classified into groups with one, two, three, or more valences [17].

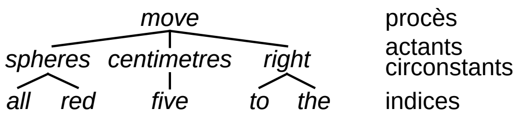

Fig. 2: VDG representation of the utterance
"move all red spheres five centimetres to the right"

3 THE SEMANTIC STRUCTURE

According to Winston's equivalence semantics [18], we look for a formal-logic semantic representation, which is independent of the context and the current dialogue status.

Furthermore, an entirely stochastic approach for semantic decoding was aimed, which integrates semantic, syntactic and acoustic knowledge that was generated within an automatic training process. For this, the maximum-a-posteriori classification formula, which has been well established in speech recognition, was adopted. Hence, solving the semantic decoding problem results in finding the most likely semantic representation S_E [11]:

$$S_E = \underset{S}{\operatorname{argmax}} \max_{W} \max_{Ph} \left[P(O|Ph)P(Ph|W)P(W|S)P(S) \right] \quad (1)$$

In addition to the acoustic-phonetic models delivering the conditional probabilities $P(O|Ph)$ and $P(Ph|W)$, a stochastic grammar is needed for the calculation of the probabilities $P(W|S)$ and $P(S)$. An important precondition for the success of our semantic decoder[3] implementing eq. (1) was the definition of the semantic representation, which should be

1. **close to the word level** to enable the calculation of $P(W|S)$ by a finite number of parameters,
2. **hierarchical** to enable the calculation of $P(S)$ by a finite number of (recursive) stochastic rules,
3. **logically correct** to consistently and comprehensibly represent semantic concepts,
4. **generalizing** to identically represent semantically equivalent, but different utterances and
5. **close to the machine level** for a simple transformation of the classified semantic representation into machine comprehensible commands.

3.1 Definition of the Semantic Structure

According to these requirements, we introduce the *semantic structure S* as semantic representation of an utterance within a restricted domain [8]. It is hierarchic like a tree, which consists of *n semantic units* (abbreviated *semuns*) s_n :

$$S = \{s_1, s_2, \ldots, s_n, \ldots, s_N\} \quad (2)$$

Each semun s_n can be described by $(X+2)$ components, its type $t[s_n]$, its value $v[s_n]$ and $X \geq 1$ references to its successors $q_1[s_n], \ldots, q_X[s_n] \in \{s_{n+1}, \ldots, s_N, \text{blk}\}$:

$$s_n = \left(t[s_n], v[s_n], q_1[s_n], \ldots, q_X[s_n] \right) \quad (3)$$

- The **type** $t[s_n]$ lays down the number X of successors[4] and restricts the set of possible successor-types $t[q_1[s_n]], \ldots, t[q_X[s_n]]$. Furthermore, it makes a selection of the corresponding values $v[s_n]$.
- The **value** $v[s_n]$ shows the exact meaning of s_n.
- Each **successor** $q_x[s_n]$ specifies a certain fact of the semun s_n. If the utterance contains that certain specification, the successor $q_x[s_n]$ is identical with another semun within *S*. In that case, the successor is denoted as *successor semun*

$$q_x[s_n] \in \{s_{n+1}, \ldots, s_N\}. \quad (4)$$

  If the utterance does not contain that certain specification, then it is a *blank successor*

$$q_x[s_n] = \text{blk}. \quad (5)$$

  For the consistent description of our stochastic approach, it is necessary to allow a type for the blank successor:

$$t[\text{blk}] = \text{blk} \quad (6)$$

The types of all successors $q_1[s_n], \ldots, q_x[s_n], \ldots, q_X[s_n]$ of the semun s_n can be denoted as *successor group*

$$t_q(s_n) = \left(q_1[s_n], \ldots, q_x[s_n], \ldots, q_X[s_n] \right) \text{ with } X \geq 1 . \quad (7)$$

3) A detailed description of the implementation of the search algorithm is given in [16], a short overview is in [14].

4) This means: $X = X(t[s_n])$. At present, we use $1 \leq X \leq 5$.

A semun s_n together with all references to its X successors $q_1[s_n]$, …, $q_x[s_n]$, …, $q_X[s_n]$ can be graphically depicted as shown in the following figure. The successors are numbered from 1 to X from top to bottom. The reference to a successor semun is marked by the edge "$\longrightarrow$". In contrast to this, the edge "$\longmapsto$" marks the reference to a blank successor:

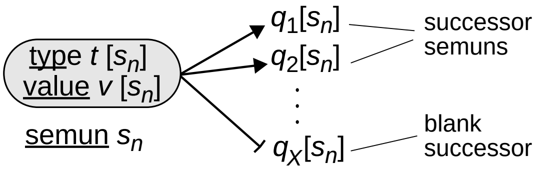

Fig. 3: Illustration of a semun s_n with X successors

The whole semantic structure S forms a 'tree' with the semun s_1 as 'root' and the blank successors as 'leaves'. All semuns $s_2, \ldots, s_N$ belong to exactly one predecessor semun.

A **branch**, denoted as $S(s_n)$, is formed by a semun $s_n \in S$ with (recursively regarded) all its successor semuns down to all terminating blank successors. The semuns within each branch $S(s_n)$ are consecutively indexed starting from n, blank successors are not included.

$$S(s_n) = \{s_n, s_{n+1}, s_{n+2}, \ldots\} \tag{8}$$

Hence, each branch is a subset of the semantic structure:

$$S(s_n) \subset S \quad \text{for} \quad n \neq 1 \tag{9}$$

According to the previous explanations, the branch $S(s_1)$ of the root semun s_1 is identical to the whole semantic structure S.

$$S(s_1) = S \tag{10}$$

Each single semun represents a small semantic part of the whole utterance. The significance of a semun is specified by its successors, which are in a fixed relation to each other. Each branch $S(s_n) \subset S$ represents a connected part of the total meaning.

A **line** is a direct sequence of m+1 semuns $s_n, \ldots, s_{n+m}$ within a branch, each one with respectively X=1 successor. A line starts with the successor semun of a semun with $X \geq 2$ successors and is terminated by a semun, which has one (i.e. X=1) blank successor, or by the predecessor semun of a semun with $X \geq 2$ successors. Within a semantic structure, a line consisting of the m+1 semuns $s_n, \ldots, s_{n+m}$ is denoted by the index vector $\binom{n}{n+m}$. The set of all lines within the semantic structure S is united within the set $L(S)$.

The order of the semuns within a line has no influence on the meaning of the whole semantic structure – clearly spoken:

<u>The order of semuns within a line is optional!</u>

Two semantic structures are *equivalent*, if they contain the same information. An equivalent semantic structure comes into being (apart from the restrictions of the syntactic model) by semun permutations within a line.

As an example, the following figure shows the tree of a semantic structure S_1 with N=11 semuns.

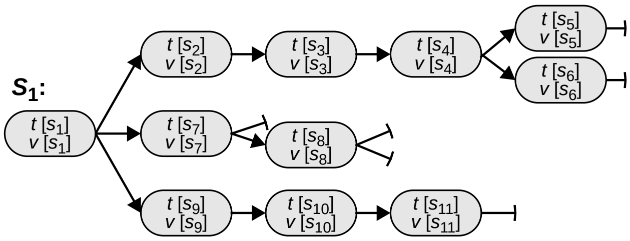

Fig. 4: Combination of several semuns s_n forming the tree of the semantic structure S_1

In the above figure, there is X=3 for s_1, X=2 for s_4, s_7, and s_8 as well as X=1 for all other semuns (the references to the respective blank successors count, too). The branch $S_1(s_2) = \{s_2, s_3, s_4, s_5, s_6\}$ consists of five semuns, the branch $S_1(s_4) = \{s_4, s_5, s_6\}$ of three and the branch $S_1(s_7) = \{s_7, s_8\}$ of two semuns.

Within the semantic structure S_1, there are two lines

$$L(S_1) = \left\{\binom{2}{3}, \binom{9}{11}\right\}, \tag{11}$$

the number of existing lines trivially results in

$$|L(S_1)| = 2 . \tag{12}$$

In the semantic structure S_1, the order of the semuns s_2 and s_3 does not matter, i.e. the order is either $\rightarrow s_2 \rightarrow s_3 \rightarrow$ or $\rightarrow s_3 \rightarrow s_2 \rightarrow$. The same rule refers to the semun order within the other line $\rightarrow s_9 \rightarrow s_{10} \rightarrow s_{11} \mapsto$.

Please note, that within a successor group $t_q(s_n)$, the order of successors $q_1[s_n]$, …, $q_x[s_n]$, …, $q_X[s_n]$ carries information and must not be changed. Hence, changing the branch $S_1(s_2)$ and the branch $S_1(s_7)$ would alter the meaning of the semun s_1 and therefore alter the meaning of the whole semantic structure S_1, too.

Fig. 5 shows an example for a concrete semantic structure S_2 in a graphic depiction:

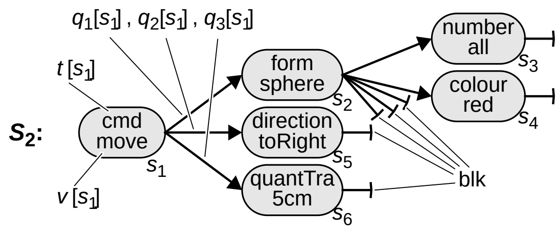

Fig. 5: Semantic structure S_2 corresponding to the utterance *"move all red spheres five centimetres to the right"*

Basically, the structure is quite similar to the notation in VDG (see chap. 2.3). As central part, the respective "ruling" command verb forms the root of the semantic structure. A single semun can be compared to an X-place relational constant, in which a 0-place relational constant is modelled by X=1 blank successor. However, the connection of single semuns to a semantic structure essentially differs from the representation by PL. Although mathematically not as exact, the semantic structure offers the following advantages:

- The semantic structure S is a meaning representation close to the word level and allows an immediate and probabilistic correlation to a word chain W, for which the conditional probability $P(W|S)$ can be calculated by certain knowledge bases. It only considers the semantic content of the utterance without consideration of the pragmatic constellation or the dialogue status. Since the semantic structure has not any knowledge about previous utterances, ellipses[5], anaphoric[6] or kataphoric[7] references cannot be resolved – this must happen in a post-processing stage.
- For the connection of semuns, there is only one single mechanism, namely the reference to other semuns as respective successor semuns. Therefore, only this one kind of the connection has to be considered during the design of the models for the calculation of the a-priori probability $P(S)$.

3.2 Relation to the Word Chain

The semantic structure is a meaning representation close to the word level. Its structure is dependent on the word choice and, although conditionally, on the word order of the corresponding word chain. Nevertheless, it has the capability for generalization: Different, but semantically equivalent word chains correspond to identical semantic structures. For this purpose, the following determinations are established:

- Each word of the word chain W corresponds to exactly one semun s_n of the semantic structure S.
- Each semun s_n of the semantic structure S is assigned to <u>exactly one significant word</u> $w^+[s_n]$ and <u>at most one insignificant word</u> $w^-[s_n]$ of the word chain W. In this manner, two or more words (in morphological sense) can possibly be combined to one contiguous "word". Since only the corresponding semantic structure is important as final result, a morphologically correct word representation is not absolutely necessary. Therefore, the design of the syntactic model (estimation of the probability $P(W|S)$ for the occurrence of the word chain W given the semantic structure S) is possible without any further representation levels.
- Each branch $S(s_n) \subseteq S$ corresponds to an unbroken phrase $W(s_n) \subseteq W$.

5) An ellipsis is a kind of text concatenation, whereat the text reference is created by vacancies, e.g. the utterance *"now a little bit to the right"* after the utterance *"move the sphere to the left"*.

6) Backward reference, e.g. the utterance *"paint it green"* after the utterance *"create a cone"*.

7) Forward reference; this never occurred during the system tests and the "Wizard of Oz"-simulation [9].

As an example, the relation between the German word chain *"bitte den schönen grünen quader neben der ähm roten kugel löschen"* (*"please delete the nice green cuboid besides the ahm red sphere"*) and its semantic structure S_3 is shown in fig. 6. The significant words $w^+[s_n]$ of the semuns s_n are marked by dark arrows "⟶", the insignificant words $w^-[s_n]$ (*"bitte"*, *"schönen"*, *"ähm"*) by bright arrows "⟶". The phrase $W(s_n)$ belonging to the respective branch $S_3(s_n)$ is shown below. Please note, that the alignment of these phrases are dependent on the hierarchy of the underlying semuns.

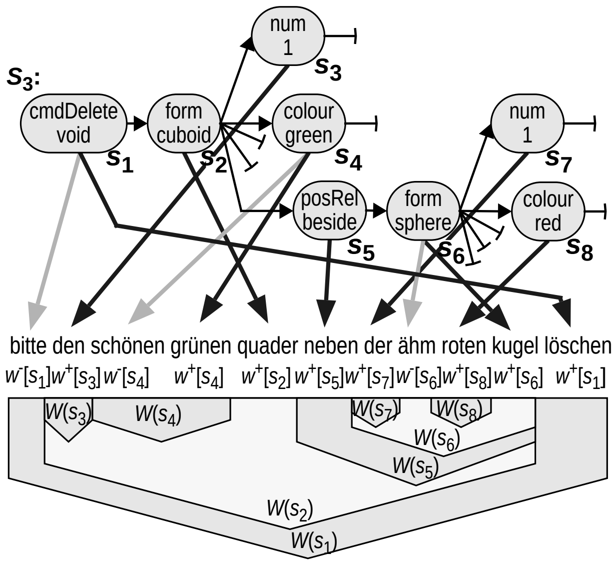

Fig. 6: Semantic structure S_3 and corresponding phrases $W(s_n)$

3.3 Limitation of the Semantics

Since the domain is limited to a certain area of interest, the "semantic portions" representing the user utterances should be within a finite set and be intuitively deduced. The goal is to divide an expected information into small partitions (e.g. certain commands, forms, colours, etc.) in such a way, that all imaginable and correct informations within the current domain can be described.

As one must consider during the development of a new programming language, which commands should be included within the respective (language-) inventory, it is necessary to consider here, which intentions a user can have during the interaction with a domain specific speech understanding application and how he/she could semantically express his/her intention by natural spoken utterances. The inventory, which has to be developed in this case, is the set of existing types and values. By meaningful combinations of these types and values among each other as well as of semuns, which consist of different types, values and successors, an infinite number of semantic structures can theoretically be created with a relatively small type and value inventory.

4 CONCLUSIONS

The semantic structure as a flexible way for showing the meaning of a spoken or written utterance was introduced. With the examples given in chap. 2, it is possible to show the following interesting facts and comparisons:

- An infinite number of semantic structures can be described by a finite set of first order conditional probabilities [13].
- The semantic structure allows a direct and probabilistic correlation with the corresponding word chain *W*.
- The semantic structure treats the verb (as in VDG) as the centre of the utterance and represents it as root semun s_1 in its tree hierarchy.
- The tree hierarchy of the semantic structure is very similar to the hierarchy of SNs or VDG.
- Similar to the principle of SNs or VDG, each semun can be specified by a successor mechanism.
- With the recursive structure, it is simply possible to describe connections as in PL. In contrast to PL, which does not allow a direct correlation to the corresponding word chain *W*, the semantic structure enables the calculation of a probability $P(W|S)$ as in eq. (1).
- In the sense of PL, a semun with *X* successors can be compared to an *X*-place relational constant. In this context, a 0-place relational constant can be realized by a semun s_n with X=1 blank successor $q_1[s_n]$ = blk.

In spite of the domain-specific restrictions of the semantic structure, we achieve 88.4% semantic accuracy (for disjoint training and test set) [15][16] using this representation within a integrated semantic decoder, which utilizes exclusively stochastic and trained knowledge.

Because of the separation of domain-specific and domain-independent knowledge, we attained an easy portability to various domains. Up to now, we developed semantic structures for two different applications:

- NASGRA (NAtural Speech understanding GRAphic editor) for creating, modifying or deleting three-dimensional objects on the screen (understanding German and Slovenian as well as translation into German, English, French, Slovenian) [10][15][16] and
- ROMAN (ROving MANipulator), speech understanding service robot for carrying out jobs like fetching and bringing things (understanding German) [1].

Currently, we are adopting two further domains:

- Controlling a virtual reality tool for medical image processing by speech (understanding German) [5] and
- spontaneously spoken scheduling dialogues (translating German into English).

REFERENCES

[1] C. Fischer, P. Havel, G. Schmidt, J. Müller, H. Stahl, M. Lang: *Kommandierung eines Serviceroboters mit natürlicher, gesprochener Sprache*, in G. Schmid, F. Freyberger (ed.): Proc. "Autonome Mobile Systeme 1996" (Munich, Germany), Springer „Informatik aktuell“, 1996, pp. 248-261

[2] G. Görz: *Strukturanalyse natürlicher Sprache*, Addison-Wesley Publishing, Bonn, 1988

[3] H. Haugeneder, H. Trost: *Beschreibungsformalismen für sprachliches Wissen*, in G. Görz (ed.): „Einführung in die künstliche Intelligenz“, Addison-Wesley Publishing, Bonn, 1993, pp. 372-424

[4] R.A. Kowalski: *Logic for Problem Solving*, Elsevier North Holland, New York, 1979

[5] C. Krapichler, M. Haubner, A. Lösch, K. Englmeier: *A Human-Machine Interface for Medical Image Analysis and Visualization in Virtual Environments*, Proc. ICASSP 1997 (Munich, Germany), pp. 2613-2616

[6] A. Linke, M. Nussbaumer, P.R. Portmann: *Studienbuch Linguistik*, M. Niemeyer Verl., Tübingen, 1994

[7] H. Marko: *Die Theorie der bidirektionalen Kommunikation und ihre Anwendung auf die Informationsübermittlung zwischen Menschen (subjektive Information)*, Kybernetik 3 (1966), pp. 128-138

[8] J. Müller, H. Stahl: *Stochastic Modelling of Syntax and Semantics*, Proc. KI 1995 Activities: Workshops, Posters, Demos (Bielefeld, Germany), pp. 229-230

[9] J. Müller, H. Stahl: *Collecting and Analyzing Spoken Utterances for a Speech Controlled Application*, Proc. Eurospeech 1995 (Madrid), pp. 1437-1440

[10] J. Müller, H. Stahl, M. Lang: *Automatic Speech Translation Based on the Semantic Structure*, Proc. ICSLP 1996 (Philadelphia, USA), pp. 658-661

[11] R. Pieraccini, E. Levin: *Learning how to Understand Language*, Proc. Eurospeech 1993 (Berlin, Germany), pp. 1407-1412

[12] C.E. Shannon, W. Weaver: *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, 1949

[13] H. Stahl, J. Müller: *A Stochastic Grammar for Isolated Representation of Syntactic and Semantic Knowledge*, Proc. Eurospeech 1995 (Madrid, Spain), pp. 551-554

[14] H. Stahl, J. Müller, M. Lang: *An Efficient Top-Down Parsing Algorithm for Understanding Speech by Using Stochastic Syntactic and Semantic Models*, Proc. ICASSP 1996 (Atlanta, USA), pp. 397-400

[15] H. Stahl, J. Müller, M. Lang: *Controlling Limited-Domain Applications by Probabilistic Semantic Decoding of Natural Speech*, Proc. ICASSP 1997 (Munich, Germany), pp. 1163-1166

[16] H. Stahl: *Konsistente Integration stochastischer Wissensquellen zur semantischen Decodierung gesprochener Äußerungen*, Ph.D. thesis, Department of Electrical Engineering and Information Technology, Munich University of Technology, 1997

[17] L. Tesnière: *Éléments de syntaxe structurale*, Klincksieck, Paris, 1959

[18] P.H. Winston: *Artificial Intelligence*, 3rd edition, Addison-Wesley, 1992