DOCUMENT RESUME

ED 339 227 FL 019 905

AUTHOR Wittrock, Merlin C.
TITLE A Classification of Sentences Used in Natural Language Processing in the Military Services.
INSTITUTION California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY Advanced Research Projects Agency (DOD), Washington, D.C.
REPORT NO CSE-TR-294
PUB DATE Jun 89
CONTRACT N-00014-86-K-0395
NOTE 24p.
PUB TYPE Reference Materials - General (130)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Classification; Cognitive Psychology; *Discourse Analysis; Inferences; *Interpersonal Communication; *Language Processing; *Language Role; *Language Styles; *Military Science; Semantics; Sentence Structure

ABSTRACT

Concepts in cognitive psychology are applied to the language used in military situations, and a sentence classification system for use in analyzing military language is outlined. The system is designed to be used, in part, in conjunction with a natural language query system that allows a user to access a database. The discussion of military language and military decision-making from the perspective of cognitive psychology looks at the characteristics of this language style, the intention of the speaker or questioner, use of top-down language processing, the role of background knowledge in military communication, the effects of context on meaning, the relevance of syntactic and semantic analysis for military language use, and the usefulness of inferential and domain-specific processing. The taxonomy is an outline of sentence characteristics incorporating all of these concepts, and is designed to be applied to both single utterances and connected discourse. (MSE)

C / T A

Center for Technology Assessment

A CLASSIFICATION OF SENTENCES USED IN NATURAL LANGUAGE PROCESSING IN THE MILITARY SERVICES

CSE Technical Report 294

Merlin C. Wittrock

Center for Technology Assessment
UCLA Center for the Study of Evaluation

- UCLA Center for the Study of Evaluation

  in collaboration with:

- UCLA Computer Science Department

  Machine Perception Lab

  Artificial Intelligence Lab

A CLASSIFICATION OF SENTENCES USED IN NATURAL LANGUAGE PROCESSING IN THE MILITARY SERVICES

CSE Technical Report 294

Merlin C. Wittrock

Center for Technology Assessment
UCLA Center for the Study of Evaluation

June, 1989

The research reported herein was conducted with partial support from the Office of Naval Research, Defense Advanced Research Projects Agency, pursuant to grant number N00014-86-K-0395. However, the opinions expressed do not necessarily reflect the position or policy of this agency and no official endorsement by this agency should be inferred.

Please address inquiries to: CSE Dissemination Office, UCLA Graduate School of Education, 405 Hilgard Avenue, Los Angeles, California, 90024-1521

Abstract

In this report, concepts from cognitive psychology are applied to the problems of developing a taxonomic system for classifying sentences used in natural language processing in the military services. The report first presents a conception of the characteristics of cognitive psychology that are related to the goal-oriented, top-down, technical, pragmatic, idiomatic, and sometimes non-grammatical nature of language processing in the military services. The paper then focuses on the central importance of the role of pragmatics and inferential analyses when processing language in stressful, goal-oriented military situations. The paper concludes with a taxonomy of sentences that has been derived from this cognitive conception of natural language processing in the military services.

An Application of Cognitive Psychology[1]

Cognitive psychologists interested in comprehension view natural language processing as a form of communication. From cognitive psychological perspectives, the central purpose of natural language processing is to communicate meaning in everyday contexts. When the context becomes that of making military decisions during warfare, natural language processors face difficult conditions for understanding questions and for constructing informed and useful answers.

Under these difficult and often stressful conditions the construc on of informed and useful answers to imperfectly coherent, sometimes ungrammatical questions requires appropriately designed natural language processors. These processors incorporate appropriate grammars, knowledge representation systems, language analyzers, conceptual dependencies, semantic networks, and representations of realistic contextual and pragmatic information, including the intention of the questioners and their restricted use of language under stress and fatigue.

For example, under these realistic conditions, what are the appropriate uses of a transformational generative grammar or a case grammar? Should knowledge be presented in declarative form or as a set of procedures? Should language analyzers use primitives (conceptual dependencies) or patterns of semantic relations (semantic networks)? How should pragmatic or contextual information be incorporated into the natural language processing system? How can the questioner's intention or purpose best be represented in the system? Does the inclusion of the questioner's intention imply a top-down (goal-driven) parsing or a combination of a top-down and a bottom-up (text-driven) processing?

To study these problems of language processing it is productive to use standard sentences or questions as input (e.g., John gave Mary a ring. Why did John give Mary a ring?). Is it more productive to use questions, paragraphs, and longer texts representative of the processing problems that are actually encountered in realistic military contexts?

In a chapter on natural language processors and realistic language comprehension, Riesbeck (1982) answers that question. He writes:

> It is becoming increasingly obvious in the field that we have to deal with real texts, texts that were originally generated to communicate, not to test parsers. The days when we can compare progress by how well they handle "Max went to the store" have passed (p. 37).

He continues by discussing the importance of studying realistic multi-sentential connected texts. In these texts he includes idioms, cliches, and run on statements whose structures make standard control structures of language analyzers "if not wrong, certainly irrelevant to the problems it had" (p. 39).

I too maintain that an appropriate and more productive study of these issues of realistic language comprehension requires language used in military situations, under the difficult conditions faced by the people working with the natural language processors. These situations and conditions include the following characteristics that influence the construction of a taxonomy of questions and the evaluation of a natural language processing system designed to understand them and to respond informatively to them.

---

[1]I want to acknowledge and to thank all of the authors listed in the references, especially Marianne Celce-Murcia, Mary Dee Harris, and Wendy G. Lehnert. Their writings and ideas were invaluable in the preparation of this report and in the design of this classification system.

Characteristics of Natural Language Processing in the Military Sciences

1. The intention of the military questioner and the goal-oriented nature of the communication.
2. The importance of performing top-down (goal-, frame-, or script-driven processing) as well as bottom-up (text-driven) processing.
3. The technical and real-world background knowledge of the questioner and the stored knowledge of the language processor.
4. The importance of pragmatics (information about realistic contexts).
5. The ability to use previous sentences and references to them to understand and respond to questions.
6. The capability to analyze the syntax and semantics of the question or text.
7. The capability to understand non-literal questions and to give non-literal, inferential replies.
8. The capability to understand metaphors, similes, and analogies.
9. The importance of domain-specific processing, including the understanding of non-standard English that contains idioms and technical phrases and even false sentences.
10. The capability to understand and to respond informatively to non-grammatical questions and statements used in the military services.
11. The capability to understand and respond to sentences of less than perfect coherence composed under stress and fatigue.
12. The ability to suggest alternative informative answers when the requested answer either is impossible or is impractical to construct within the time or other constraints of the situation.
13. The goal not only of understanding the question, but of informing the questioner as well.

The following sections elaborate some of the contributions of cognitive psychology to natural language processing in these conditions and situations.

The Intention of the Speaker or Questioner

One of the most significant contributions of recent cognitive research on comprehension involves the importance of knowing the intention or the purpose of the speaker or the questioner. To understand a statement or a question usually requires knowledge about the speaker. What problems is he trying to solve? Why is he asking these questions? What is his purpose? This important theme recurs in a variety of contexts.

In instructional psychology, intention is the purpose of the learners; it is their objective, the focus of their attention.

In linguistics, two schools of thought predominate. Formalists view meaning as the definition of the conditions under which sentences are true. Speakers' or writers' intentions do not pertain to the formalists' conception of meaning, which encounters difficulty in interpreting ambiguous sentences in the real world. In contrast, intentionalists (Grice, 1975; Austin, 1961) maintain that one can understand a sentence only by knowing the intention of the speaker or the writer to communicate meaning.

The intentionalist approach to language processing applies well to the conditions and situations of natural language processing in the military services.

In computer science, expectation-based parsing includes this element of purpose or intention. Dictionary definitions of common words do not always convey meaning adequately. "You're the one for the job" may or may not be a compliment. The meaning of the sentence depends upon the intention of the speaker, just as the answer to the question "Do you know what time it is?" depends upon understanding the intention of the questioner. Some implications of these concepts upon language processing are developed in the next section.

Top-Down Processing

Bottom-up or script-driven parsing uses keywords of the text to construct summaries of the input. Script-driven parsers are notorious for their lack of understanding (Riesbeck, 1982) because keywords can have multiple meanings. When a keyword's meaning differs from the sense interpreted by the bottom-up processor, it constructs erroneous summaries of text.

Top-down or interest-driven and goal-driven parsers, such as Fast Reading Understanding and Memory Program (FRUMP) and Integrated Partial Parsing (IPP), use internally stored scripts or Memory Organization Packets (MOPs)[2] as templates of basic elements to understand text such as stories. Actually, these processors combine bottom-up and top-down processing. FRUMP begins by searching for words or phrases of the first sentence of the text that refer to one of its scripts. From this sketchy survey, it selects one of its scripts. After that, the processing is all top-down. FRUMP no longer asks the meanings of words. Instead it seems to ask "Can I make this word mean what the sketchy script requires it to mean?"

Integrated Partial Parsing (IPP) also uses scripts, but it includes frames several times in each sentence. These frames add meanings to words that do not fit well into the script. MOPs—which are like scripts, but smaller, more modular, more domain specific, and more closely tied to episodic memory (Riesbeck, 1982; Schank, 1982)—offer greater utility in goal-driven parsing and in selecting the most relevant or useful interpretation of ambiguous words and sentences.

The domain-specific problems that characterize natural language processing in military settings require a goal-driven parser, but one less procrustean than FRUMP. The goal-driven parser needs to be able to process non-grammatical, incomplete sentences with missing punctuation and even with missing verbs. The parser also needs inference rules to check the reasonableness of its interpretation of a question, a colloquial phrase, or a text. Parsers such as McMAP apparently can perform these difficult tasks (Riesbeck, 1982).

The language processor should be capable of interpreting the ill-formed structure of the language input that is encountered in realistic military situations. It also should be sensitive and flexible enough to modify its interpretations to accommodate meanings that earlier top-down processors would have ignored.

---

[2]A high-level script that abstracts information common across two or more scripts, which enables the prediction of future events for which no appropriate script is available.

Background Knowledge

Perhaps the most important finding of recent research on comprehension in cognitive psychology centers on the relationship between background knowledge and the text or the utterances. In reading comprehension, schema theory (Rumelhart, 1980) focuses upon this relationship. According to schema theory, to understand a sentence or a text involves categorizing or "slotting" it into a schema, much as one would fit an example into a generic category. In my model of generative reading comprehension (Wittrock, 1974, 1981) this concept of understanding is carried one step further to include the generation of new schemata by the learner or reader.

The learners' models of the events they are trying to comprehend summarize an important part of their background knowledge. In recent research on learning and instruction, the learners' models of addition and subtraction (Brown & Vanlelm, 1982; Carpenter, Moser, & Romberg, 1982), motion, and DC current flow (Osborne & Wittrock, 1983) have been extensively studied. In these studies the learners' models critically determine what information can be understood, and how it can be best presented to the learners to convey meaning.

In linguistics, pragmatics[3] includes and extends beyond the psychological presuppositions of the speaker or questioner (Bates, 1976). Semantic or logically derived meanings of utterances are not broad enough to encompass the impact of pragmatics upon comprehension; in this case, these utterances are not broad enough to encompass the background knowledge of the learner. In computer science, background knowledge sometimes appears as scripts and frames (Schank & Abelson, 1977) or as MOPs (Schank, 1983). However, learner background knowledge is not the same as semantic information or stored dictionary knowledge, or the same as the users' goals, intentions, and expectations, or the same as pragmatics. This variable represents a psychological characteristic useful for understanding speakers' or writers' questions and for designing natural language processors appropriate for domain specific processing.

Context

A real-world setting or situation comprises a complex set of features important to everyday communication. For example, within an entire text, context includes all paragraphs and the title. In a paragraph, it includes all sentences; within a sentence, it includes all words and phrases; in a word, it includes all letters or morphemes. Context, however, is derived from more than the text alone.

In a real-world setting on a Navy ship, the context involves all the characteristics of the setting relevant to communication. These might include: (a) the time constraints for making a decision, (b) the presence of an enemy force, (c) the state of readiness of the ship and its crew, (d) the orders given to the ship's captain or his commanding officer, and even (e) the health and mood of the ship's personnel. (The intention of the questioner is a psychologically distinct variable, although in linguistics it is sometimes included as a contextual or pragmatic variable.)

In cognitive approaches to the study of instruction, the relationships between different contexts of learning, such as a word problem or a number problem, and previous instruction often mean the difference between making a concept comprehensible or incomprehensible. In linguistics and computer science, context represents a significant part of pragmatics, which the intentionalists consider critical to explaining natural language processing.

---

[3]The use of the linguistic form appropriate for the situation and the relationships among the communicating parties (from Celce-Murcia, 1983, p. 525).

For natural language processing in the military services, context definitely affects the meaning or interpretation given to questions and other sentences. It plays an important role in the evaluation of the usefulness of natural language processors that are designed for military use.

Syntactic and Semantic Analyses

The domain-specific characteristics of language usage in the military, including its technical terms, colloquial phrases, and occasional non-grammatical structures, imply constraints upon syntactical and semantic analyses. Non-grammatical statements or questions must not be rejected in a syntactic analysis because they are not grammatically correct. The deep structure and meaning of the question should be the focus of the syntactical and semantic analyses, not the surface structure or the dictionary meaning of the words in the sentences. The syntactic analyses of the structure should be closely tied to the semantic analyses and should be limited to their implications for understanding the functional relationships of the phrases in the sentence.

A case grammar provides a useful semantic analysis of the sentences used in the military services. It goes beyond the structure of the sentence to emphasize the functional relations among the phrases of the sentences. In contrast, a transformational grammar works closely with the surface structure of the sentence. It works closely with the syntax of the sentence, but not closely enough with the semantics of the sentence. With transformational grammar, we cannot reverse the rules for generating sentences; we cannot use them to analyze sentences. Transitional Networks (TN), Augmented Transitional Networks (ATN), and Recursive Transitional Networks (RTN) solve some of these problems of transformational grammars, but they produce highly complex analyses of sentences.

Fillmore (1968) introduced a paradigm of case grammar that consisted of six cases:

1. agentive
2. dative
3. facilitative
4. locative
5. instrumental
6. objective

These cases have been modified somewhat by Fillmore and others since their introduction. Stockwell, Schacter, and Partee (1973) developed a set of five cases:

1. agent
2. dative
3. instrumental
4. locative
5. neutral

Celce-Murcia's paradigm of case grammar (1972), developed for recognizing sentences such as those encountered in military services, provides a parsimonious categorization of predicate relations:

1. causal actant: the person who causes the action

2. theme: the person or thing about whom a statement is made (Fillmore's objective case)

3. locus: the location

4. source: directional information

5. goal

Case grammars, such as Celce-Murcia's, Stockwell's, and Fillmore's, lead to manageable and semantically informative sentence analyses that are flexible enough to encompass many of the problems of intergrammatical structure and non-standard English encountered in the military services. Celce-Murcia's set of cases and Robert Simmons (1973) set of modes are used in the taxonomy presented in this paper.

Inferential Processing

The problems of developing inferences occur in many contexts of language processing, including those encountered in the military services. The language processor needs to understand non-literal questions within the context of the situation and with the intention of the questioner. "Can you tell me how many ships are battle-ready?" requires an answer other than "yes." The need to go beyond dictionary meanings to include the understanding of metaphors, similes, and analogies occasionally used in military contexts implies a need for a sophisticated language processor capable of difficult inferential analyses based upon stored knowledge about the language usage of military personnel.

Domain-Specific Processing

The domain-specific processing of natural language includes a diverse set of semantic and pragmatic issues, many of which have been discussed (such as the questioner's intention and language usage in the military). In addition, personnel in the military services use a technical vocabulary that implies real-world knowledge and that changes sentence structure.

Many utterances in the military are idiomatic. Some are non-grammatical. Some are stated exactly the opposite of their intention: "It is impossible to underestimate the destruction of an atomic war" states the opposite of its military speaker's intention. A natural language processor needs to be able to use a script to infer the implausibility of the sentence. It needs to reason from the speaker's purpose that another interpretation, the opposite meaning, was intended. With these problems, and with idiomatic expressions, the goal is not to make a literal interpretation, or even to make a definitive interpretation of the meaning of the sentence. Instead, the goal is to infer the meaning within the context of the sentence and within the intention of the speaker.

At this point we come close to FRUMP's procrustean approach to making sentences "mean what we want them to mean." But because the objective of communication in warfare usually is clear—clearer, at least, than the syntactical structures of the sentences—a more useful way to build meaning is to infer it from the script and not to get lost in unwieldy structural ambiguities.

Domain-specific processing in the military services often involves language usage under conditions of emotional stress, severe time constraints, and the necessity to act quickly upon incomplete data. The evaluation of natural language processors for use under these conditions implies that we test them using sentences composed under high stress and with great fatigue. We need to learn how these conditions affect sentence usage and language coherence.

Goal of the Processing

The goal of the processing of natural language is not only to understand questions, it is to inform the questioner. To inform the questioner means to answer the questions, but it also means to suggest alternative information or alternative ways to get information. If the question cannot be answered within a reasonable time, the natural-language processor should indicate what information it can provide more quickly that would be useful to the questioner. The processor can suggest that the questioner can find out something of value, even if it is not the precise answer requested. An understanding of the goal of the questioner leads to this capability. The intention of the questioner again influences the language processing and gives meaning to the question and to its answers.

A Taxonomy

The taxonomy presented below represents the synthesis of the concepts I have discussed. It has several features designed to process realistic language in military settings. First, it includes the dialect or language use of military personnel, along with their degree of stress and fatigue. Language use differs across dialects and deteriorates under stress and fatigue.

Second, it includes an elaborate section on the speaker's intentions with questions and with declarative statements. Wendy Lehnert (1986a) suggests 12 categories of questions: causal/antecedent, goal orientation, enablement, causal consequent, verification, disjunctive, instrumental/procedural, concept completion, expectational, judgmental, quantification, and feature specification. After considering her useful scheme, I decided to develop an alternative that can serve also to classify declarative statements and that makes additional distinctions that are important in the military service.

Third, because even realistic single utterances require real-world pragmatic and contextual information to give them meaning, I have used the pragmatic classification system with single utterances and with connected discourse. The frame-driven, top-down processing discussed previously seems important in the military context, especially under conditions of stress and fatigue. The use of pragmatic classification means that single utterances, contextual discourse, and dialogue can all be analyzed with a classification system that includes syntax, semantics, and pragmatics because all types of texts involve all three types of linguistic analyses.

Fourth, the system can be used to categorize declarative sentences as well as questions or imperatives such as the IRUS queries.[4]

No classification system can be complete or perfect. There is no intent for this system to classify the English language. Instead, the intent is to develop a system that will classify, in a useful way, the realistic language encountered in natural language processing in the military services.

---

[4]IRUS is a natural language query sytem which allows a user to access a database. The IRUS queries referred to above are 163 questions and imperatives regarding military applications which were taken from an August 1986 CINCPACFLT demonstration.

A Taxonomy of Sentences Used in Natural Language Processing in the Military Services

I. Processing single utterances

A. Processing syntax
1. Type of sentences
a) Declarative
(1) Active
(2) Passive
b) Interrogative
c) Imperative

2. Types of phrases and clauses
a) Nouns vs modifiers
b) Noun groups
c) Prepositions in noun groups
d) Appositives
e) Other modifiers

3. Word order
a) Standard
b) Inverted

4. Referents
a) Structure-based
(1) Of adjectives
(a) Post nominal
i) Participle phrases
ii) Relative clauses
(b) Standard
(2) Of pronouns
(a) Direct
(b) Indirect
(3) Of objects
(4) Of phrases
b) Frame or script-based
(1) Definite references
(a) Referential
(b) Attributive
(c) Set
i) Generic
ii) Individual
iii) Attribute
(d) Number
i) Singular
ii) Plural
(e) Articles
i) Definite
ii) Indefinite
(f) Time
i) Past
ii) Present
iii) Future
(g) Interjection
(h) Previous utterance
i) Sentence
ii) Text

(2) Indefinite reference
(3) Pronoun
(4) Other

5. Non-grammatical sentences
   a) Spelling
   b) Punctuation
   c) Missing words
   d) Missing constituents
   e) Wrong order
   f) Semantic constraint violations
   g) Sentence fragments

B. Processing semantics
   1. Lexicon
      a) Single word
      b) Phrase

   2. Literal meaning
      a) Dictionary meaning
         (1) Single meaning
         (2) Multiple meaning
      b) Technical term or phrase
         (1) Single meaning
         (2) Multiple meaning

   3. Frame-driven meaning
      a) Frame relevant to disambiguating meaning
      b) Frame irrelevant to disambiguating meaning

   4. Inferential meaning
      a) Idiom
      b) Metaphor
      c) Simile
      d) Analogy
      e) Colloquial phrase
      f) Domain specific (novel)
         (1) Indirect speech
         (2) Focus or theme of text
      g) Ellipses
         (1) Sentence based
         (2) Discourse based
      h) Conjunctions
         (1) Context based
         (2) Context irrelevant
      i) Comparatives
      j) Anaphora
         (1) Forward reference
         (2) Standard reference

   5. Modality
      a) Tense
         (1) Present
         (2) Past
         (3) Future
      b) Aspect
         (1) Present
         (2) Imperfect

c) Form
   (1) Simple
   (2) Emphatic
   (3) Progressive
d) Mood
   (1) Declarative
   (2) Interrogative
   (3) Imperative
e) Essence
   (1) Positive
   (2) Negative
   (3) Indeterminate
f) Modal
   (1) May
   (2) Can
   (3) Must
g) Manner
h) Time

6. Case
   a) Causal actant
   b) Theme
   c) Locus
   d) Source
   e) Goal

7. Metacognition
   a) Direction about inference building
   b) Direction about intention

C. Processing pragmatics
1. Context (setting)
   a) Military
   b) Non-military

2. World knowledge (commen sense)
   a) Script
      (1) Single
      (2) Multiple
   b) MOPs
   c) Experiential
   d) Conceptual
   e) Cultural

3. Domain-specfic knowledge
   a) Technical
   b) Everyday

4. Mode of knowledge
   a) Spatial
   b) Verbal

5. Model used by the speaker

6. Intention (goal, purpose) of the speaker
   a) To obtain or to infer information (question)
      (1) Verification (yes-no)
      (2) Denotation

(a) Quantification (how many)
i) Exact
ii) Estimate
(b) Composition (who, which, what)
(c) Location (where)
(d) Time (when)
(3) Connotation (affect)
(4) Identification
(5) Description
(a) List
(b) Combine
(c) Summarize
(6) Relation
(a) Cause and effect
(b) Instrumental and enablement
(c) Goal
i) Subsumption
ii) Concordance
(d) Correlation
(e) Association
(f) Logical
(g) Among parts
(h) Contradiction
(7) Interpretation
(a) Organization
(b) Explanation
(c) Prediction
(d) Implication
(e) Decision
(f) Conclusion
(g) Action
(h) Planning
(8) Evaluation
(9) Persuasion
(a) Action
(b) Belief
(10) Multiple functions

b) To give or to help someone infer information (statement)
(1) Verification (yes-no)
(2) Denotation
(a) Quantification (how many)
i) Exact
ii) Estimate
(b) Composition (who, which, what)
(c) Location (where)
(d) Time (when)
(3) Connotation (affect)
(4) Identification (name)
(5) Description
(a) List attributes
(b) Combine attributes
(c) Summarize
(6) Instruction
(a) Cause and effect (why)
(7) Relation
(a) Cause and effect (why)
(b) Instrumental and enablement

(c) Goal
i) Relations among goals
ii) Subsumption
iii) Concordance
(d) Correlation
(e) Association
(f) Logical
(g) Among parts
(h) Contradiciton
(8) Interpretation
(a) Organization
(b) Explanation
(c) Prediction
(d) Implication
(e) Decision
(f) Conclusion
(g) Action
(h) Planning
(9) Evaluation
(10) Persuation
(11) Multiple functions
c) To command
d) To make a decision
e) To take action
f) To attain a goal
g) To play a role

7 Expectation
a) Of the speaker
b) Of the listener

8. Presuppositions of the speaker
a) Logical
b) Contextual
c) Psychological

9. Point of view of the speaker

10. Language of the speaker
a) Standard English
b) Non-standard English
(1) Black
(2) Hispanic
(3) ESL
(4) Other

11. Emotion of the speaker
a) Stress
(1) Stressed
(2) Not stressed
b) Fatigue
(1) Fatigued
(2) Not fatigued

II. Processing connected discourse

A. Processing syntax
1. Type of sentences
a) Declarative
(1) Active
(2) Passive
b) Interrogative
c) Imperative

2. Types of phrases and clauses
a) Nouns vs modifiers
b) Noun groups
c) Prepositions in noun groups
d) Appositives
e) Other modifiers

3. Word order
a) Standard
b) Inverted

4. Referents
a) Structure-based
(1) Of adjectives
(a) Post nominal
i) Particpal phrases
ii) Relative clauses
(b) Standard
(2) Of pronouns
(a) Direct
(b) Indirect
(3) Of objects
(4) Of phrases
b) Frame or script-based
(1) Definite references
(a) Referential
(b) Attributive
(c) Set
i) Generic
ii) Individual
iii) Attribute
(d) Number
i) Singular
ii) Plural
(e) Articles
i) Definite
ii) Indefinite
(f) Time
i) Past
ii) Present
iii) Future
(g) Interjection
(h) Previous utterance
i) Sentence
ii) Text
(2) Indefinite reference
(3) Pronoun
(4) Other

5. Non-grammatical sentences
   a) Spelling
   b) Punctuation
   c) Missing words
   d) Missing constituents
   e) Wrong order
   f) Semantic constraint violations
   g) Sentence fragments

B. Processing semantics
1. Lexicon
   a) Single word
   b) Phrase

2. Literal meaning
   a) Dictionary meaning
      (1) Single meaning
      (2) Multiple meaning
   b) Technical term or phrase
      (1) Single meaning
      (2) Multiple meaning

3. Frame-driven meaning
   a) Frame relevant to disambiguating meaning
   b) Frame irrelevant to disambiguating meaning

4. Inferential meaning
   a) Idiom
   b) Metaphor
   c) Simile
   d) Analogy
   e) Colloquial phrase
   f) Domain specific (novel)
      (1) Indirect speech
      (2) Focus or theme of text
   g) Ellipses
      (1) Sentence based
      (2) Discourse based
   h) Conjunctions
      (1) Context based
      (2) Context irrelevant
   i) Comparatives
   j) Anaphora
      (1) Forward reference
      (2) Standard reference

5. Modality
   a) Tense
      (1) Present
      (2) Past
      (3) Future
   b) Aspect
      (1) Present
      (2) Imperfect
   c) Form
      (1) Simple
      (2) Emphatic
      (3) Progressive

d) Mood
   (1) Declarative
   (2) Interrogative
   (3) Imperative
e) Essence
   (1) Positive
   (2) Negative
   (3) Indeterminate
f) Modal
   (1) May
   (2) Can
   (3) Must
g) Manner
h) Time

6. Case
   a) Causal actant
   b) Theme
   c) Locus
   d) Source
   e) Goal

7. Metacognition
   a) Direction about inference building
   b) Direction about intention

C. Processing pragmatics
1. Context (setting)
   a) Military
   b) Non-military

2. World knowledge (commen sense)
   a) Script
      (1) Single
      (2) Multiple
   b) MOPs
   c) Experiential
   d) Conceptual
   e) Cultural

3. Domain-specfic knowledge
   a) Technical
   b) Everyday

4. Mode of knowledge
   a) Spatial
   b) Verbal

5. Model used by the speaker

6. Intention (goal, purpose) of the speaker
   a) To obtain or to infer information (question)
      (1) Verification (yes-no)
      (2) Denotation
         (a) Quantification (how many)
            i) Exact
            ii) Estimate
         (b) Composition (who, which, what)

(c) Location (where)
(d) Time (when)
(3) Connotation (affect)
(4) Identification
(5) Description
(a) List
(b) Combine
(c) Summarize
(6) Relation
(a) Cause and effect
(b) Instrumental and enablement
(c) Goal
i) Subsumption
ii) Concordance
(d) Correlation
(e) Association
(f) Logical
(g) Among parts
(h) Contradiction
(7) Interpretation
(a) Organization
(b) Explanation
(c) Prediction
(d) Implication
(e) Decision
(f) Conclusion
(g) Action
(h) Planning
(8) Evaluation
(9) Persuasion
(a) Action
(b) Belief
(10) Multiple functions

b) To give or to help someone infer information (statement)
(1) Verification (yes-no)
(2) Denotation
(a) Quantification (how many)
i) Exact
ii) Estimate
(b) Composition (who, which, what)
(c) Location (where)
(d) Time (when)
(3) Connotation (affect)
(4) Identification (name)
(5) Description
(a) List attributes
(b) Combine attributes
(c) Summarize
(6) Instruction
(7) Relation
(a) Cause and effect (why)
(b) Instrumental and enablement
(c) Goal
i) Relations among goals
ii) Subsumption
iii) Concordance
(d) Correlation
(e) Association

(f) Logical
(g) Among parts
(h) Contradiciton

(8) Interpretation
(a) Organization
(b) Explanation
(c) Prediction
(d) Implication
(e) Decision
(f) Conclusion
(g) Action
(h) Planning

(9) Evaluation

(10) Persuation

(11) Multiple functions

c) To command
d) To make a decision
e) To take action
f) To attain a goal
g) To play a role

7. Expectation
a) Of the speaker
b) Of the listener

8. Presuppositions of the speaker
a) Logical
b) Contextual
c) Psychological

9. Point of view of the speaker

10. Language of the speaker
a) Standard English
b) Non-standard English
(1) Black
(2) Hispanic
(3) ESL
(4) Other

11. Emotion of the speaker
a) Stress
(1) Stressed
(2) Not stressed
b) Fatigue
(1) Fatigued
(2) Not fatigued

References

Austin, J. (1962). *How to do things with words.* Oxford, England: Oxford University Press.

Bates, E. (1976). *Language and context: The acquisition of pragmatics.* New York: Academic Press.

Brown, J.S., & Vanlehn, K. (1982). Towards a generative theory of "bugs." In T.P. Carpenter, J. Moser & T. Romberg (Eds.), *Addition and subtraction: A developmental perspective* (pp. 117-135). Hillsdale, NJ: Lawrence Erlbaum Associates.

Carpenter, T.P., Moser, J., & Romberg, T. (Eds.). (1982). *Addition and subtraction: A developmental perspective.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Celce-Murcia, M. (1972). *Paradigms for sentence recognition* (Technical Training Report AFHRL-TR-72-30). Lowry Air Force Base, CO: Air Force Human Resources Laboratory, Training Research Division.

Celce-Murcia, M., & Larsen-Freeman, D. (1983). *The grammar book: An ESL/EFL teaching course.* Rowley, MA: Newbury House.

Cohen, P.R., Perrault, C.R., & Allen, J.F. (1982). Beyond question answering. In W.G. Lehnert & M.H. Ringle (Eds.), *Strategies for natural language processing* (pp. 245-274). Hillsdale, NJ: Lawrence Erlbaum Associates.

Fillmore, C. (1968) A case for case. In E. Bach & R.T. Harms (Eds.), *Universals in linguistic theory.* New York: Holt, Rinehart, and Winston.

Grice, H. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics. Vol. 3: Speech acts.* New York: Academic Press.

Harris, M.D. (1985). *Introduction to natural language processing.* Reston, VA: Reston Publishing Company.

Lehnert, W.G. (1986a). A conceptual theory of question answering. In B.J. Grosz, K.S. Jones & B.L. Webber (Eds.), *Readings in natural language processing* (pp. 651-657). Los Altos, CA: Morgan Kaufman Publishers.

Lehnert, W.G. (1986b). *The process of question answering: Computer simulation of cognition.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Osborne, R.J., & Wittrock, M.C. (1983). Learning science: A generative process. *Science Education, 67,* pp. 489-508.

Riesbeck, C.K. (1982). Realistic language comprehension. In W.G. Lehnert & M.H. Ringle (Eds.), *Strategies for natural language processing* (pp. 37-54). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rumelhart, D. (1980). Schemata: The building blocks of cognition. In R. Spiro, B. Bruce & W. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33-58). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schank, R., & Colby, K.M. (Eds.). (1973). *Computer models of thought and language.* San Francisco: W. H. Freeman and Co.

Schank, R., & Ableson, R. (1977). *Scripts, plans, goals, and understanding.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Schank, R.C. (1982). Reminding and memory organization: An introduction to MOPs. In W.G. Lehnert & M.H. Ringle (Eds.), *Strategies for natural language processing* (pp.455-493). Hillsdale, N.J: Lawrence Erlbaum Associates.

Simmons, R.F. (1973). Semantic networks: Their computation and use for understanding English sentences. In R. Schank & K.M. Colby (Eds.), *Computer models of thought and language* (pp. 63-113). San Francisco: Freeman and Company.

Stockwell, R.P., Schacter, P., & Partee, B.H. (1973). *The major syntactic structures of English.* New York: Holt, Rinehart, and Winston.

Wittrock, M.C. (1974). Learning as a generative process. *Educational Psychologist,11*, 87-95.

Wittrock, M.C. (1981). Reading comprehension. In F.J. Pirozzolo & M.C. Wittrock (Eds.), *Neuropsychological and cognitive processes in reading* (pp. 229-259). New York: Academic Press.

Wittrock, M.C. (1985, August). *The generative processes of comprehension* (Presidential Address, Division of Educational Psychology). Presented at the Annual Meeting of the American Psychological Association, Los Angeles.