מיכל שרגל
שלאวล אואה אלון
they acquire this knowledge in the first place? What do they do when they use this knowledge? How do people know when they know a natural language?

Questions that are traditionally addressed by linguistics: formal, rigorous, computationally based investigation of empirical methods and techniques of computer science. A computational linguistics. An approach to linguistics that

What is this course about?
than computers. Linguistically, with people in ways that suit people rather between different human languages, and otherwise interact that can understand everyday human speech, translate in order to provide such novel products as computers, committed to a true simulation of cognitive processes, languages, emulating cognitive capabilities without being is concerned with computational processing of natural science, and in particular artificial intelligence, that.

What is this course about?
Example of an application: machine translation

Using technology developed by SYSTRAN
From http://babelfish.altavista.com/
and production work.

book describes research about how language comprehension
coordinating our day-to-day behavior with others. This
next. In spoken form it serves as our primary means
serves as a long-term record of knowledge from one generation
and is a crucial component of our lives. In written form it
language is one of the fundamental aspects of human behavior

Translation
Example of an application: machine
Example of an application: machine
production work.
appropriately as the understanding of a language and the
approachability of the behavior with others. This book describes the search
deserves from our primary means of coordination of our everyday
service from one generation to following. In the shape of speech, she
speech serves from record of long duration of acquaintance
behavior and is a crucial member of our crew. In the written
The language is one of the fundamental functions of human

Example of an application: machine

Translation
human behavior and is a crucial component of our language is one of the fundamental aspects of language.

**Comparison**
Comparison

Generation to following
Generation to the next
record of long duration of acquaintance from one
a long-term record of knowledge from one
In the written shape servants from
In written form it serves as

Natural Language Processing
Comparison

This book describes the search approximately as the understanding of a language and the production work.

Natural Language Processing
Example of an application: question answering

and http://www.ask.com

From http://www.ask.com
Many of the algorithms are computationally intractable

Most stages of the process involve ambiguity

The necessary knowledge is enormous

Language understanding is complicated

Why are the results so poor?
Different subfields of linguistics study these areas:

- World knowledge
- Discourse knowledge
- Pragmatic knowledge
- Semantic knowledge
- Syntactic knowledge
- Morphological knowledge
- Phonetic and phonological knowledge

What kind of knowledge is required?
Different subfields of linguistics study these areas.

- World knowledge
- Discourse knowledge
- Pragmatic knowledge
- Semantic knowledge
- Syntactic knowledge
- Morphological knowledge
- Phonetic and phonological knowledge

What kind of knowledge is required?
constraints on their combination and rules of pronunciation, phonetics defines an inventory of basic units (phonemes), that relates to sound, abstracting away from their physical properties. Phonology studies the module of the linguistic capability speech sounds, their perception and their production and used in language, including the physical properties of and phonetics studies the sounds produced by the vocal tract.
(occurrence) of phonemes with respect to one another:

**Phonotactic constraints:** restrictions on the distribution

**Allomorphs:** variants of phonemes that are in complementary

**Meaning:** 7 — 7

**Free variation:** alternation of sounds with no change in

weak — weak; to — too; two: *two*; *to*; *too*

**Homophones (homonyms):** words that are pronounced

Problems in phonological processing
Different subfields of linguistics study these areas:

- World knowledge
- Discourse knowledge
- Pragmatic knowledge
- Semantic knowledge
- Syntactic knowledge
- Morphological knowledge
- Phonetic and phonological knowledge

What kind of knowledge is required?
Infixes are inserted into the root/stem
prefixes precede the root/stem
suffixes follow the root/stem
especially roots or stems.
A suffix: a morpheme which is added to other morphemes,
Bound morphemes: books, on the
Free morphemes: book, 1920
bound (not a word) or free (word).
Morpheme: a minimal sound-meaning unit. Can either be
Morphylogy studies the structure of words.

Morphylogy

Natural Language Processing
Ambiguity: 

from base forms and inflectional affixes:  

Inflectional morphology: Inflected forms are constructed

Internationalize ← International

Internationalize ← International

Derivational morphology: Words are constructed from roots

Problems in morphological processing
Different subfields of linguistics study these areas:

- World knowledge
- Discourse knowledge
- Pragmatic knowledge
- Semantic knowledge
- Syntactic knowledge
- Morphological knowledge
- Phonetic and phonological knowledge

What kind of knowledge is required?
Furiously sleep ideas green colorless
Colorless green ideas sleep furiously
Young green frogs sleep quietly

Natural Language sentences have structure.

Syntax
Problems of syntactic processing

• chart-based vs. backtracking
• right-to-left vs. left-to-right
• top-down vs. bottom-up

ungrammatical ones.

Parsing: assigning structure to grammatical strings, reflecting

natural language?

Expressiveness: what formalism is required for describing
This book describes research about how language comprehension and production work.

Coordination:
Kim promised Sandy to call the plumber.
Kim asked Sandy to call the plumber.

Control:

Ambiguity:

Problems of syntactic processing.
Different subfields of linguistics study these areas:

- World Knowledge
- Discourse Knowledge
- Pragmatic Knowledge
- Semantic Knowledge
- Syntactic Knowledge
- Morphological Knowledge
- Phonetic and Phonological Knowledge

What kind of knowledge is required?
Semantics is obtained from the meanings of its subphrases. A good semantics is compositional: the meaning of a phrase A semantic representation must be precise and unambiguous. Semantics assigns meanings to natural language utterances.
every student doesn't like math
every student hates at least two courses

Scope ambiguity:

Word sense ambiguity: book; round; about; 102\%
VP anaphora: Kim loves his wife and so does Sandy.

record of knowledge.
behavior. In written form it serves as a long-term
language is one of the fundamental aspects of human
discussed. Constraints are suggested and problems with them are
in the next few paragraphs, some preliminary
After she robbed the bank, Kim went home
Kim went home after she robbed the bank

Co-reference and anaphora:

Problems of semantic processing
Different subfields of linguistics study these areas:

- World knowledge
- Discourse knowledge
- Pragmatic knowledge
- Semantic knowledge
- Syntactic knowledge
- Morphological knowledge
- Phonetic and phonological knowledge

What kind of knowledge is required?
Sandy’s sister is a ballet dancer.

Kim regrets that he voted for Gore.

The current King of France is bald.

Febriculously uttered:

the class of contexts in which the sentence can be

**Presupposition:** the presuppositions of a sentence determine

is said.

Pragmatics is the study of how more gets communicated than

Pragmatics
Could you pass me the salt?

I have two children.

— How old are you? — Closer to 30 than to 20.

explicitly uttered:

implícit: what is conveyed by an utterance that was not

Pragmatics
I'm really sorry!

I swear that I didn't do it

sentence you to six months in prison

Warning: I warn you that I'll see you later

Promise: I promise that I'll see you later

Prediction: I predict that I'll see you later

I'll see you later

With it:

force of an utterance, resulting from the function associated

speech acts: the illocutionary force, the communicative

Pragmatics
Pragmatics

Non-literary use of language: metaphor, irony etc.
Different subfields of Linguistics study these areas.

- World Knowledge
- Discourse Knowledge
- Pragmatic Knowledge
- Semantic Knowledge
- Syntactic Knowledge
- Morphological Knowledge
- Phonetic and Phonological Knowledge

What kind of knowledge is required?
Give me two tickets for the earlier one, please.

There is one at 2:00 and one at 2:30.

When does the train to Hita leave?

An example dialogue:

extremely important for dialogue systems. Much like sentences do, understanding discourse structure is a discourse is a sequence of sentences. Discourse has structure.

Discourse
It was summer then.
It really surprised Sandy.
She does it every year.
It changed her life.

Reference to non-NPs: Kim visited the University of Haiti.

Cross-sentential anaphora

Non-sentential utterances: she; to Haiti; the last one

Problems of discourse processing
Different subfields of linguistics study these areas.

- World knowledge
- Discourse knowledge
- Pragmatic knowledge
- Semantic knowledge
- Syntactic knowledge
- Morphological knowledge
- Phonetic and phonological knowledge

What kind of knowledge is required?
Bill Clinton left for Vietnam today. This is the last foreign
visit of the American President.

- Is the train to Haifa late? - It left Tel Aviv at 8:30.

World Knowledge
Structure of the course

- Finite-state transducers
- Finite-state automata
- Inflections and derivations
- Introduction to morphology: word structure

Morphology
Applications

Semantics
Montague Grammars

- Unification Grammars: feature structures and unification
- The limitations of CFGs
- Bottom-up chart parsing
- Parse trees, language

Syntax
- Context-free grammars: forms, derivations
- Introduction to syntax: the structure of natural languages

Structure of the course
Attendance: Optional but recommended.

Final Grade: Final exam (50%); final exam (30%); mid-term exam (20%)

Grading: 4–6 home assignments (approximately 20% of the second edition, Benjamini/Cumminigs 1995.

Textbook: James Allen, Natural Language Understanding.

Practicalities