Naturalness in Speech Communications

B.H Juang

SAP on SSP 2002

Introduction

- Natural
  - Spontaneous dialogue
- Unnatural
  - IVR (Interactive Voice Response) system with one-way command mode.

Introduction

- Questions:
  - What constitutes naturalness in speech communications ?
  - How is naturalness encapsulated in linguistic expressions ?
  - How is naturalness achieved in human communications, which may involve multiple modalities ?
  - How can we design a machine to perform like a human in spoken dialog?

Introduction

- In this paper …
  - We attempt to analyze the human behavioral components that contribute to the naturalness or perceived naturalness in human speech communications.
  - “coherent”
  - “the least effort”
  - “reference” → human-machine dialog.

A MAP of Communication

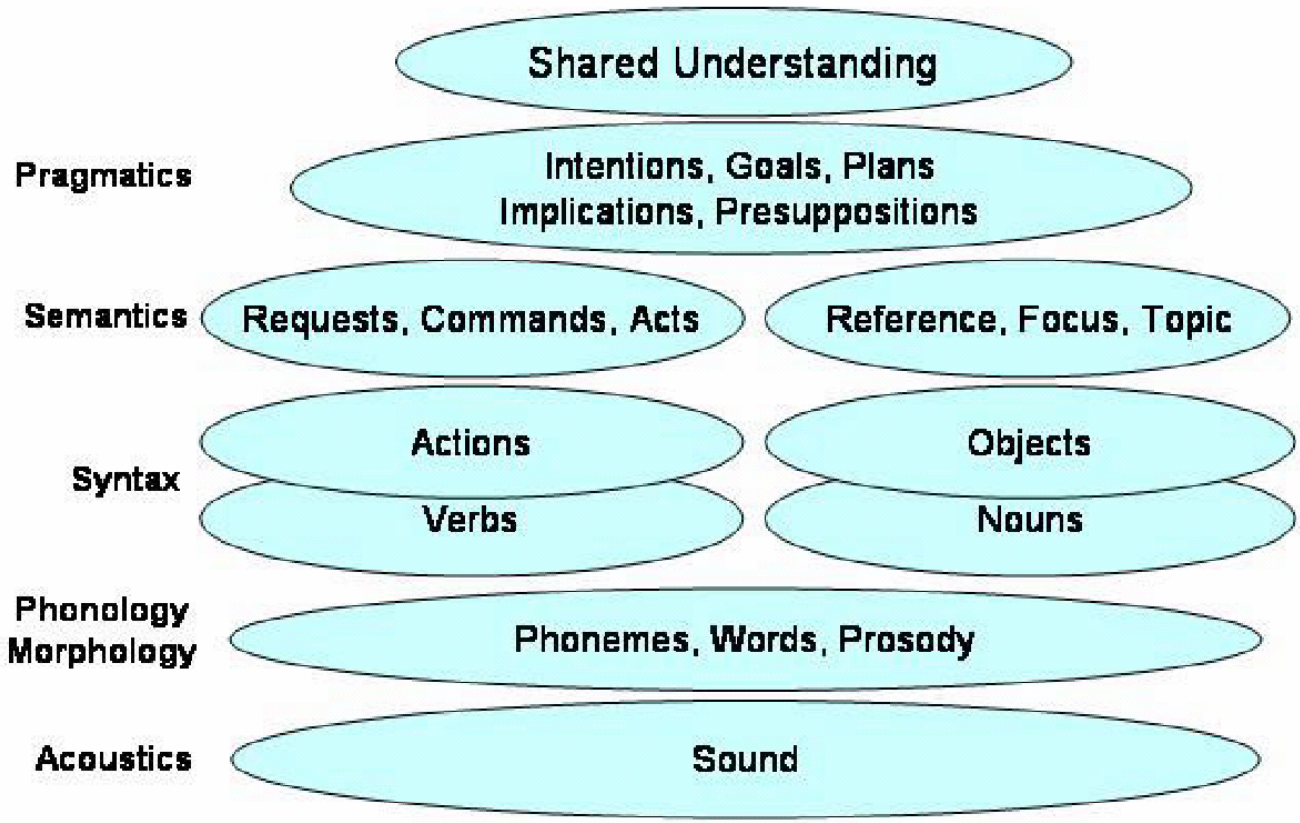

Figure 1: A map of communications

Naturalness in Communications

- “coherent” – Strategy
  - Consistent ←→ hard to comprehend
- “the least effort” – Execution
  - To err is human
  - To rest is human
- “reference” – awareness of the context
  - If every subject matter referred to in the conversation needs a complete definition every time it is unnatural.

Questions on Reference

- Questions:
  - Will explicit use of referential semantics make human-machine interactions look and feel more natural?
  - How do we include reference in dialog and grammar design to achieve the perceived naturalness?

Creating Referential Naturalness

- Task:
  - Virtual assistant → “Daisy”

- Three broad headings:
  - Dialog design
  - Dialog processing
  - Utterance comprehension

Dialog Design

- This is to answer the questions of what, how, and when reference can occur in the dialog.
- What – reference fields & objects
  - The questions can be regarded as a process of specifying a set of referential fields
- How – through modes of reference
  - The members of those fields can be referenced as an enumeration of a range of linguistic forms.
- When – in structured dialog
  - Occurs as an aspect of the modal structure of the dialog.

Dialog Processing

- Discourse Referents
  - “co-refer” or “co-referential” e.g. “December 18th”=“this Tuesday”=“tomorrow”=“that day”
- Dialog and Referent States
  - Reference state transition
- Referentially Sensitive Language Generation
  - Providing adequate but unobtrusive feedback about the comprehension of referential acts committed by the human.
  - Formulating appropriate referential expressions as the system commits such acts of its own.

Utterance Comprehension

- The language model is expressed as a finite state grammar.

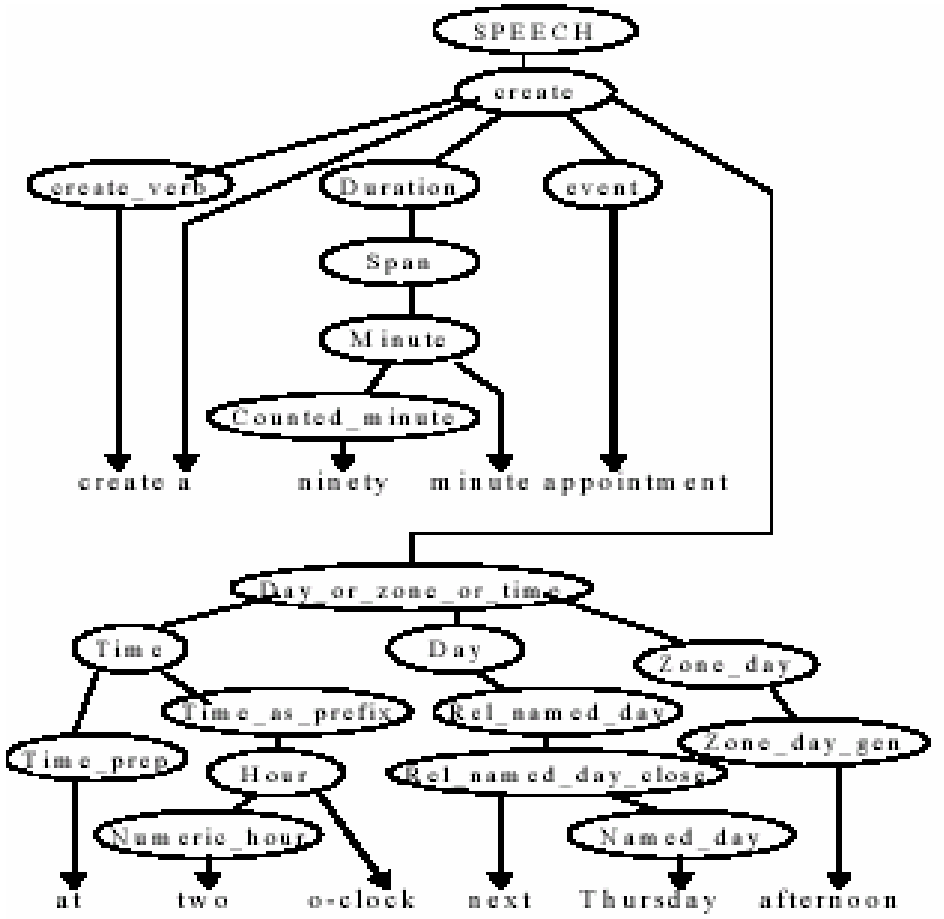

Figure 2: Parse tree returned by ASR