

# Verbmobil

## Speech-to-speech translation system

Rob Malouf  
Susanne Riehemann  
May 23, 1997  
Center for the Study of Language and Information

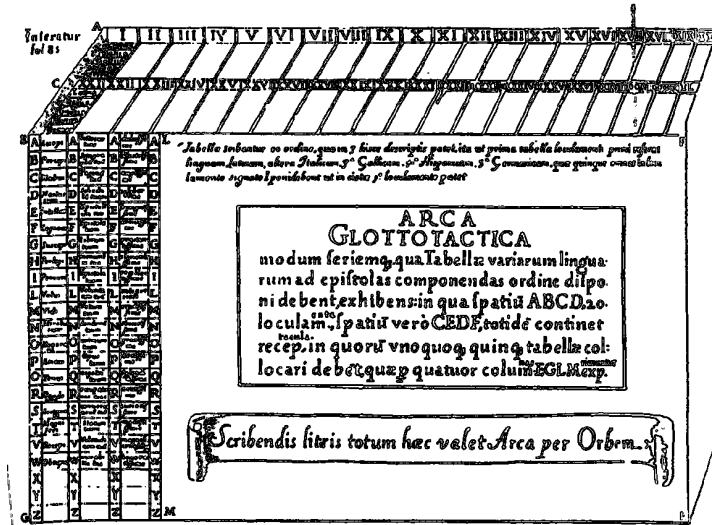
### An old problem



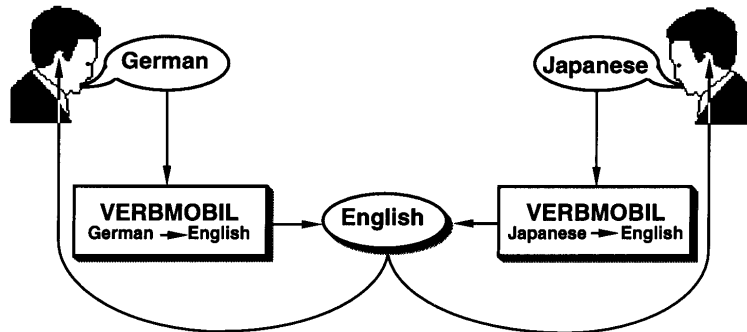
*And the LORD said, "Behold, they are one people, and they have all one language; and this is only the beginning of what they will do; and nothing that they propose to do will now be impossible for them. Come, let us go down, and there confuse their language, that they may not understand one another's speech."*

*Genesis 11 : 6–7*

# An old problem



# Verbmobil



# Verbmobil

- Speech-to-speech translation
- Uses English as pivot language (German to English and Japanese to English)
- Modular architecture
  - Developed by consortium of 22 universities and 7 industrial partners
  - Allows use of existing systems
  - Competition among partners
  - Ability to grow with improving technology

# Participants



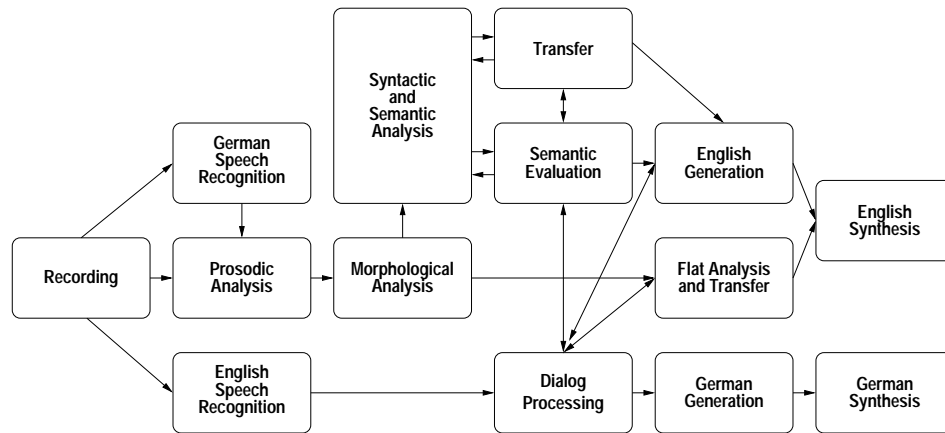
## Flexibility

- Continuous speech input
- Speaker independence with no prior training
- Relatively large vocabulary of 3,000 word forms for spoken input
- Parallel tracks for deep and shallow analysis
- Unification-based grammars for three languages
- "Anytime" processing

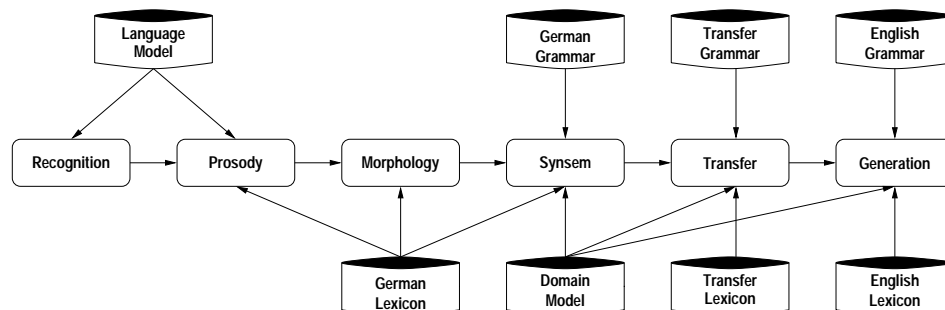
## Performance

- No more than 7 times real time for total processing time of an utterance
- Better than 70% of test inputs independently judged to result in "approximately correct" output end-to-end

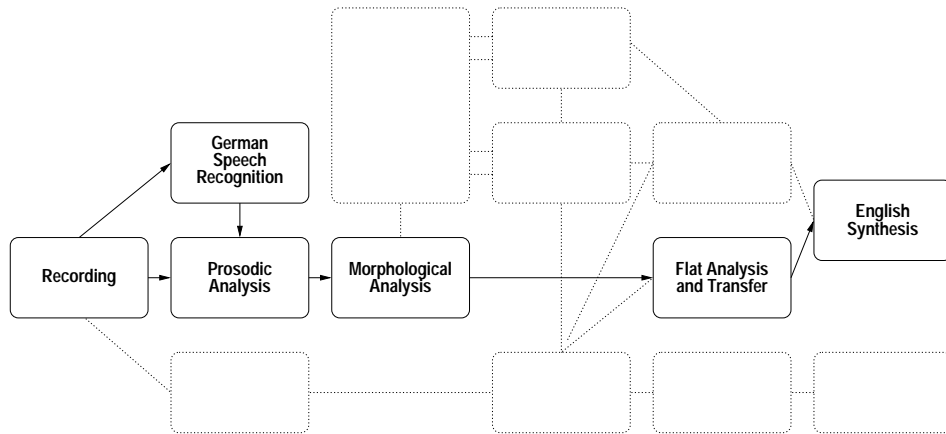
# Overview



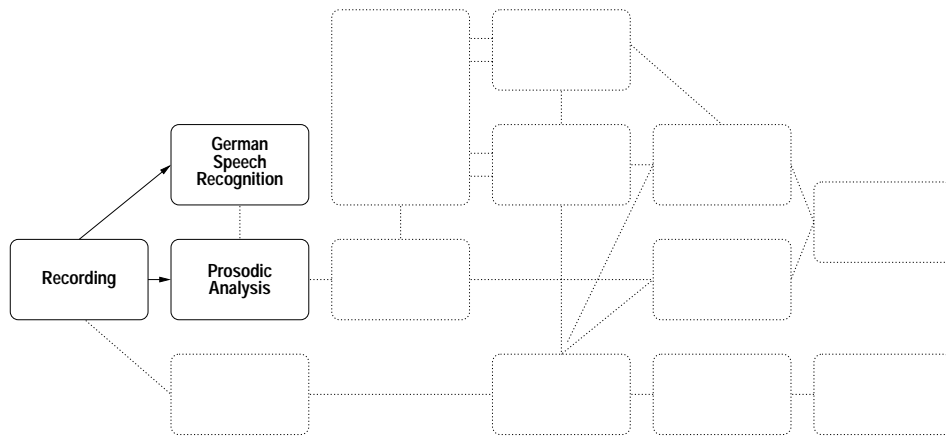
# Pipeline architecture



# Shallow processing

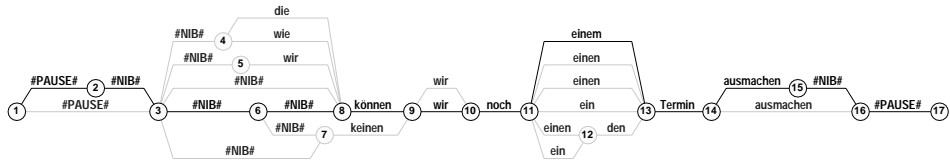


# Waveform

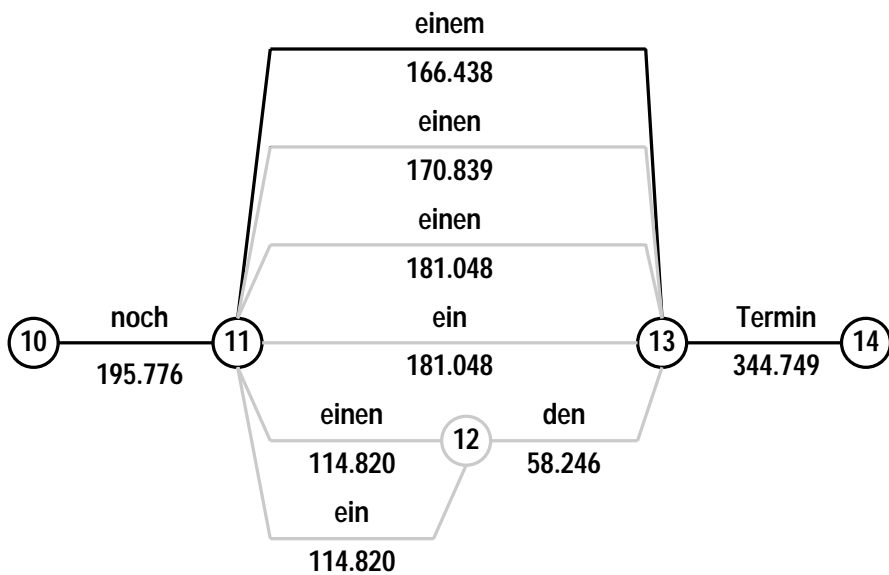




# Word lattice



# German speech recognition



## Flat analysis

"Können wir noch einem Termin ausmachen?"



"We must still fix a date."

"Ja, ich würde uhm vorschlagen, wir könnten uns im Mai treffen, uhm  
so am siebten"



"Yes. I'd suggest. So we could meet in May. Say on the seventh on."

## Flat analysis

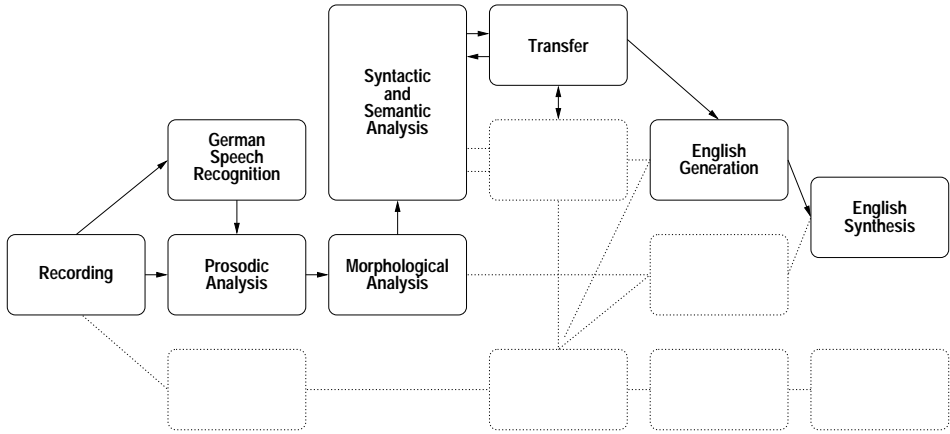
"Wir müssen keinen Termin ausmachen."

"We must arrange no appointment."

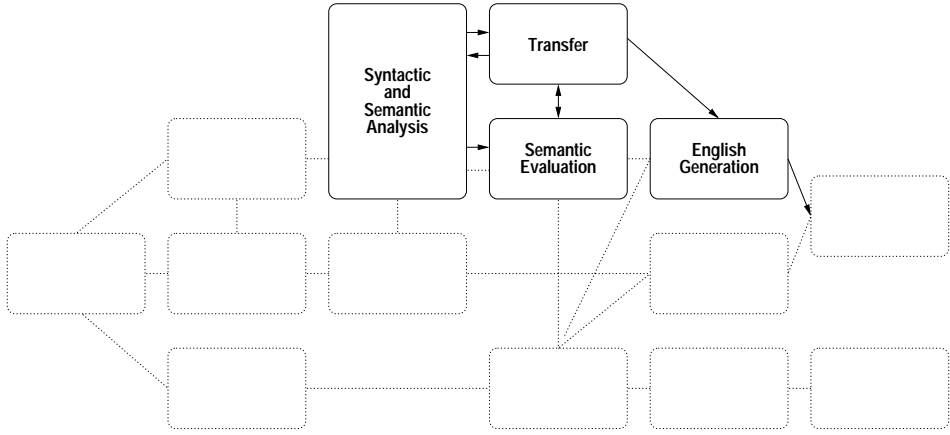


"We must take another date."

# Deep processing



# Vermobil Interface Term





## Transfer

"Das paßt echt schlecht bei mir."



"That really doesn't suit me well."

- [echt(L,I)] ↔ [real(L,I)].
- [passen(L,E), bei(L1,E,X)] ↔ [suit(E), arg2(E,X)].
- [schlecht(L,E), [passen(L1,E)] ↔ [neg(L,A), good(A,E)]
- [schlecht(L,E)] ↔ [bad(L,E)]

## Transfer

"Können wir **noch** einen Termin ausmachen?"

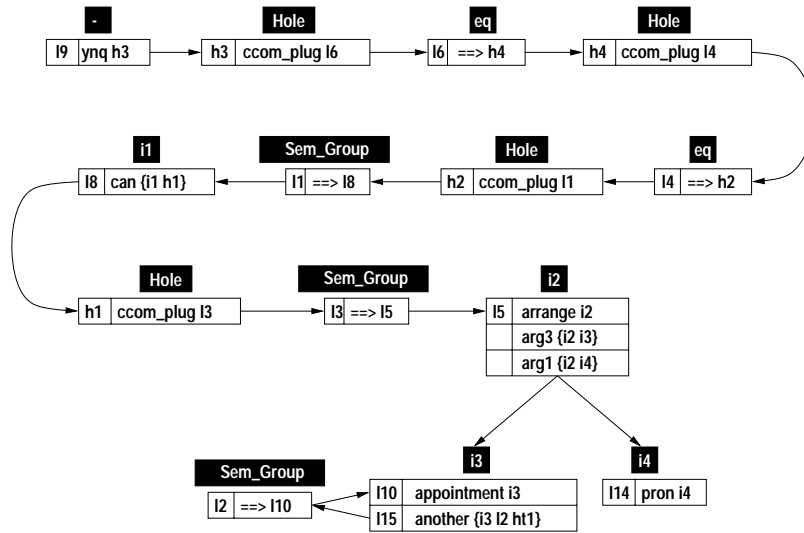


"Can we arrange **another** appointment?"

"Can we **still** arrange an appointment?"

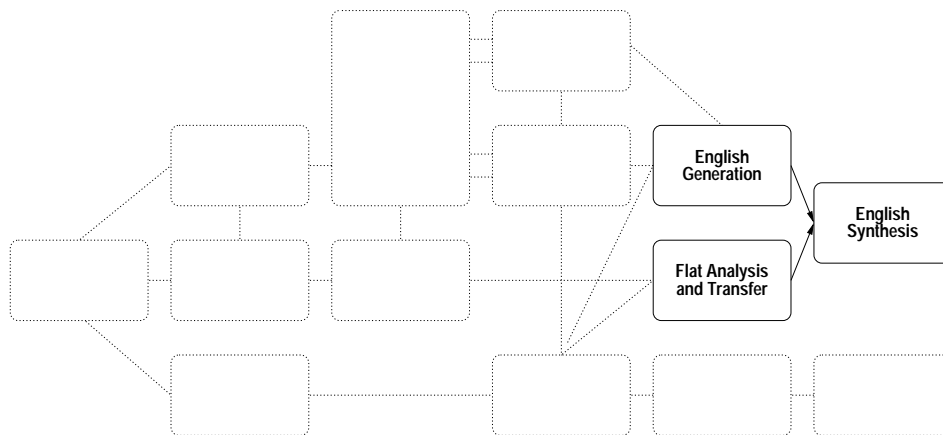
- [noch(L,F,S)] ↔ [still(L,F,S)]
- [noch(L,F,S), ein(L1,R,S1)], [L1 < F] ↔ [another(L1,R,S1), eq(S,S1)]

# English meaning

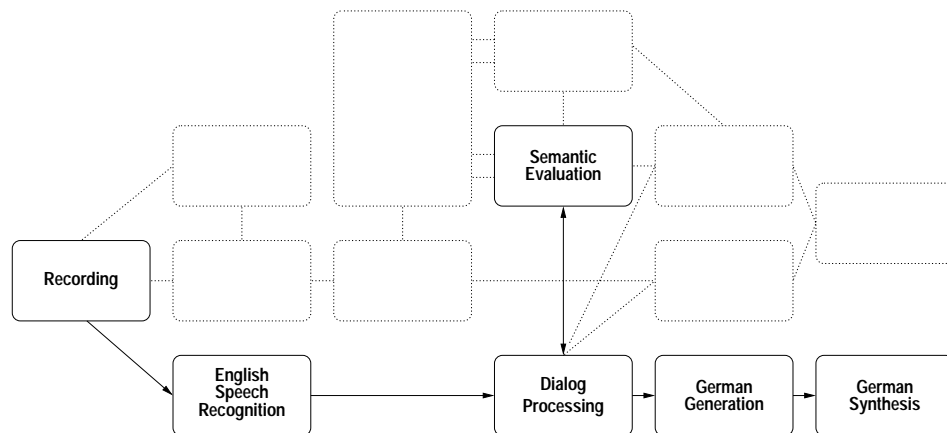


"Can we arrange another appointment?"

# Output



## Dialog



## Dialog

- Supports speech recognition and analysis with top-down predictions
- Provides contextual information
- English keyword spotter follows dialog
- Controls clarification dialogs with user

# Dialog

"Geht es bei Ihnen?"

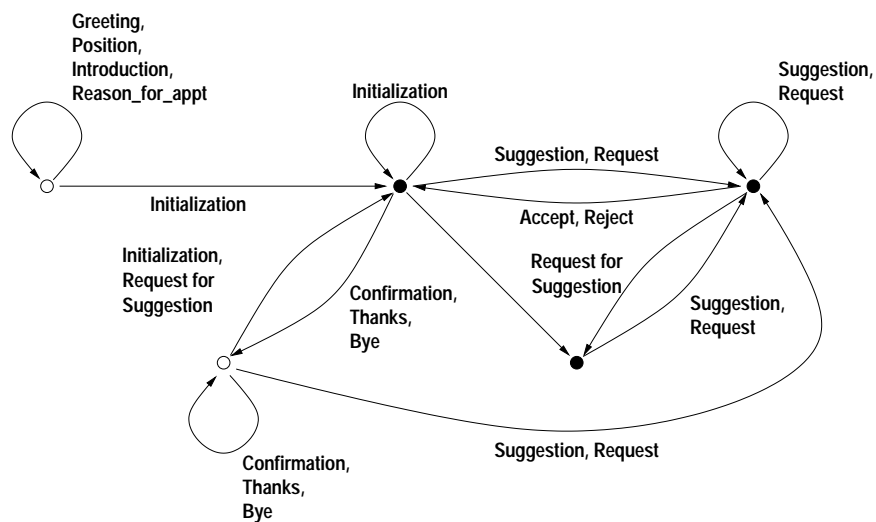


"Does it suit you?"

"How about your place?"

Statistical model + Finite state machine + Discourse Planner

# Dialog



## English Grammar

- Broad coverage
- Extensible vocabulary
- Sound theoretical framework
- Declarative formalism
- Clean semantic representation

## Broad Coverage

- Common syntactic constructions
- Precise analysis to avoid overgeneration
- Large core vocabulary

## Extensible vocabulary

- Core vs. domain-specific vocabulary
- Connections to on-line dictionaries
- Reliable mechanism for adding words
- Morphology for inflection, derivation

## Sound theoretical framework

- Rich lexicon (both generalizations and idiosyncracies)
- Small number of schematic phrase structure rules
- Unification-based constraints
- Language independent principles

## Declarative formalism

- Grammatical constraints independent of processor
- Grammar provides a formal specification for a verifiable implementation
- Allows implementors to draw from relevant computer science research (e.g., declarative specification, object oriented programming)
- Provision for statistical constraints

## Clean semantic representation

- Domain-independent meaning separate from context
- Separation of linguistic meaning from non-linguistic knowledge
- Promotes portability and reusability
- Links to discourse, especially for anaphora
- Lexical access enabled for inference, generation, translation