Multilingual Conversational Interfaces: n NTT-MIT Collaboration

Stephanie Seneff

Spoken Language Systems Group

MIT Laboratory for Computer Science

January 13, 2000

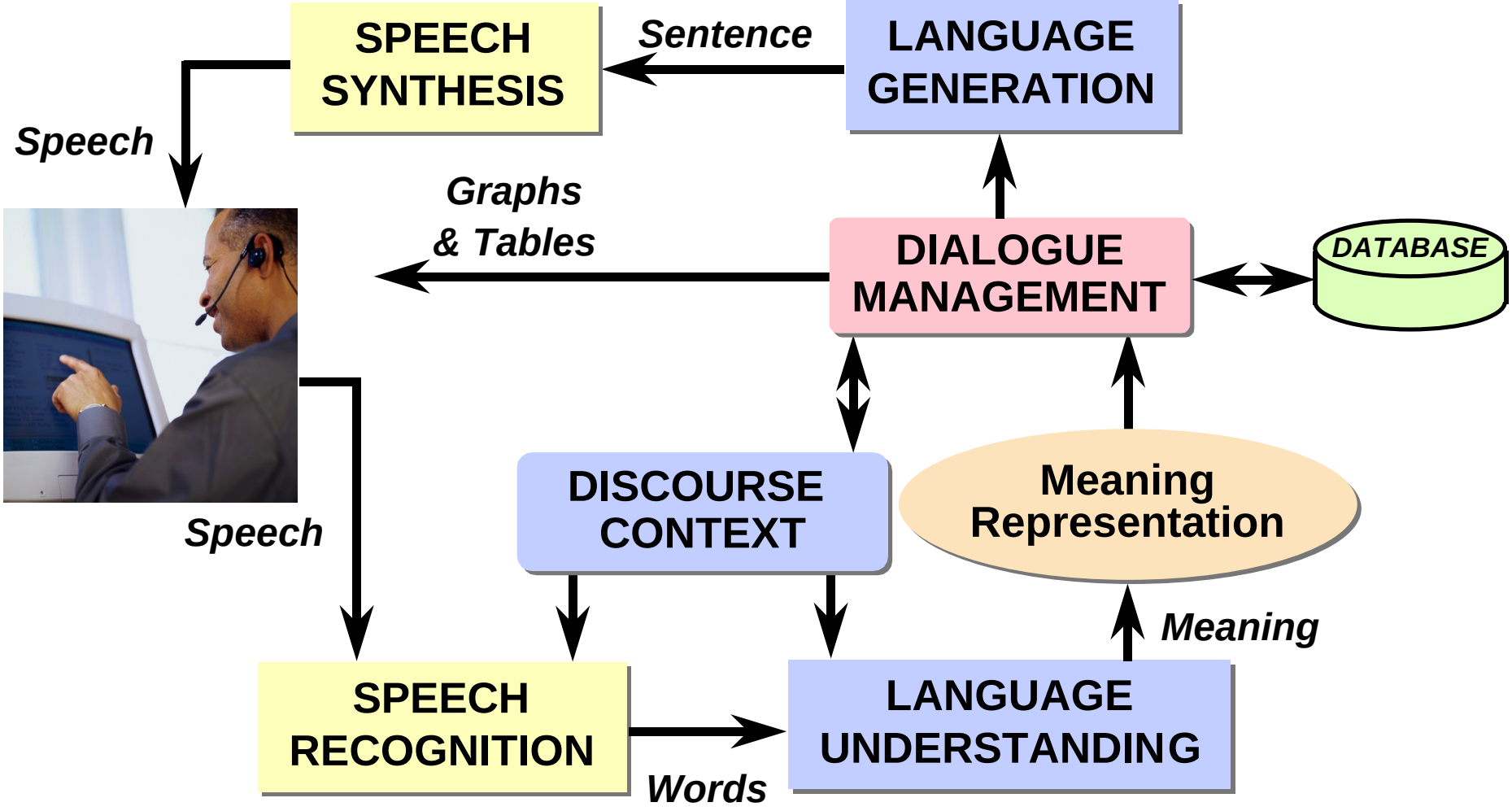
Collaborators

- James Glass (Co-PI, MIT-LCS)
- T.J. Hazen (MIT-LCS)
- Yasuhiro Minami (NTT Cyberspace Labs)
- Joseph Polifroni (MIT-LCS)
- Victor Zue (MIT-LCS)

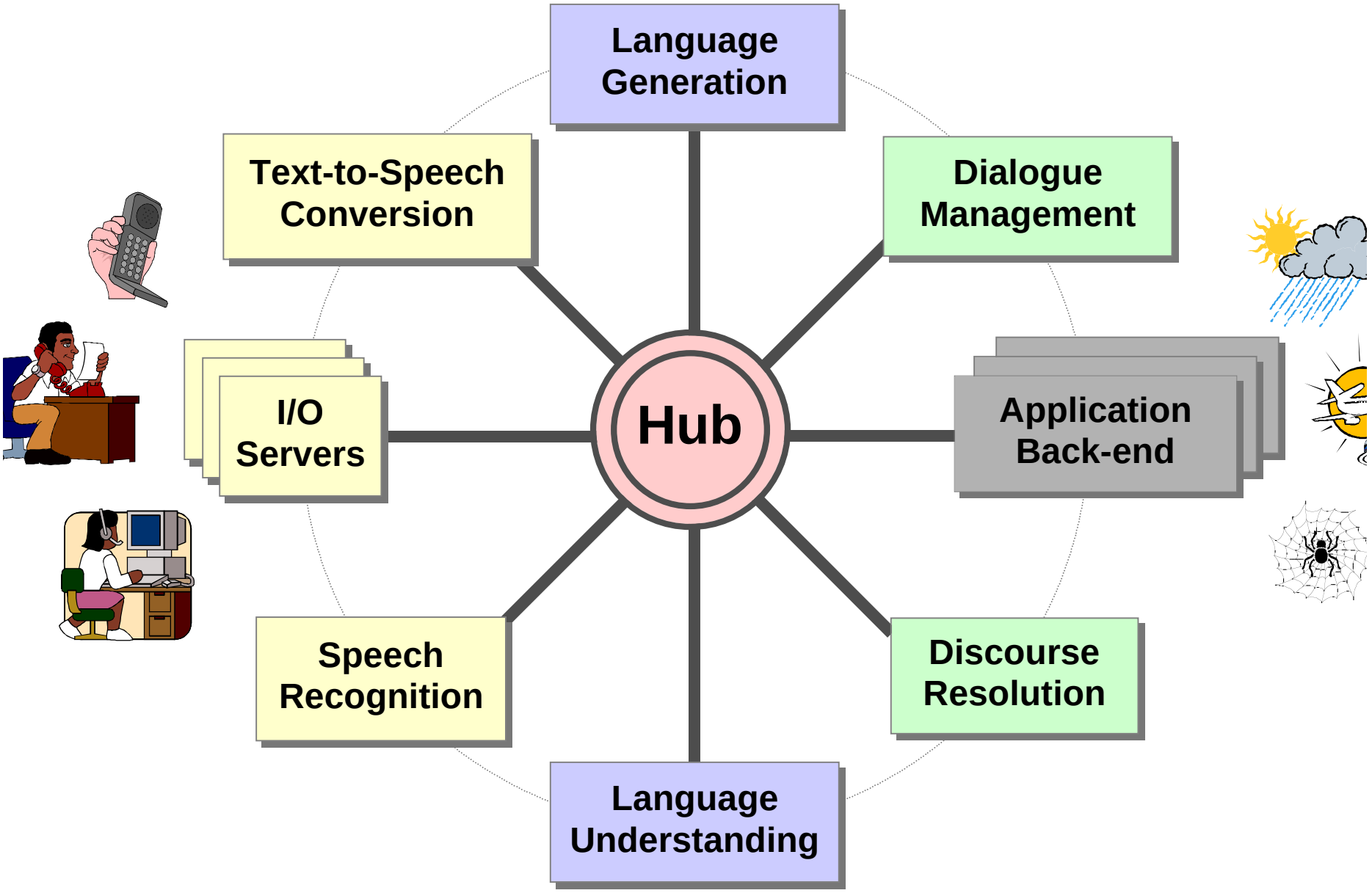
What are Conversational Interfaces

- Can communicate with users through *conversation*
- Can *understand* verbal input
  - Speech recognition
  - Language understanding (in context)
- Can retrieve information from on-line sources
- Can *verbalize* response
  - Language generation
  - Speech synthesis
- Can engage in *dialogue* with a user during the interaction

omponents of Conversational Interfaces



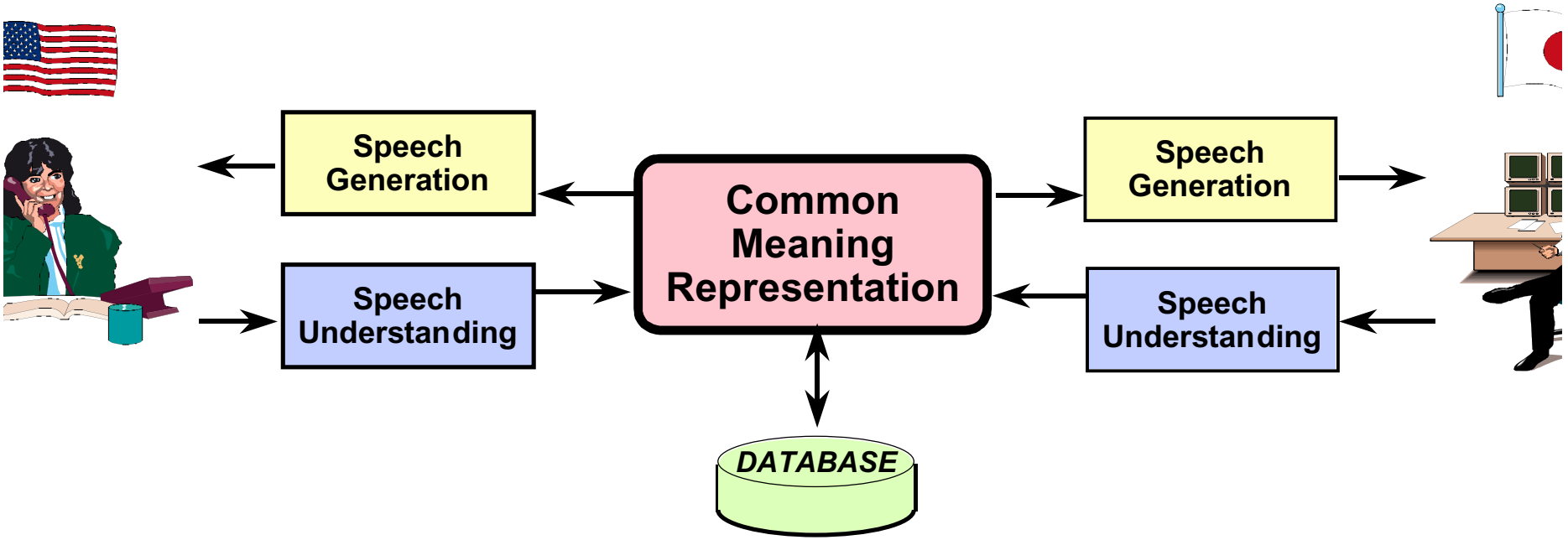
ystem Architecture: Galaxy



MIT pplication Development at MIT

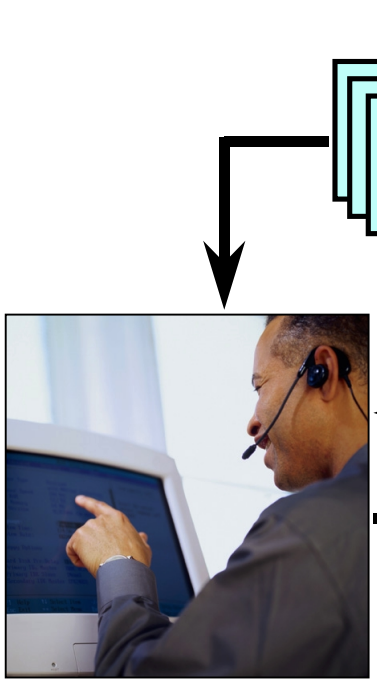
- **Jupiter: Weather reports (1997)**
  - 500 cities worldwide
  - Information updated three times daily from four web sites, plus a satellite feed
- **Pegasus: Flight status (1998)**
  - ~4,000 flights in US airspace for 55 major cities
  - Information updated every three minutes
  - Also uses flight schedule information, updated daily
- **Voyager: (Greater Boston) traffic and navigation (1998)**
  - Traffic information updated every three minutes
  - Also uses maps and navigation information
- **Mercury: Travel planning (1999)**
  - Flight information and reservation for ~250 cities worldwide
  - Flight schedule information and pricing
- **Demonstration: Jupiter in English**

TT-MIT Collaborative Research: Mokusei



- **Explore language-independent approaches to speech understanding and generation**
- **Develop necessary human-language technologies to enable porting of conversational interfaces from English to Japanese**
- **Use existing Jupiter weather-information domain as test case**
  - **It is the most mature English system**
  - **It allows us to explore language technology for interface and conten**

Multilingual Conversational Systems: Our Approach



SPEECH SYNTHESIS

LANGUAGE GENERATION

Rules

Graphs & Tables

DIALOGUE MANAGER

DATABASE

DISCOURSE CONTEXT

Meaning Representation

Models

SPEECH RECOGNITION

LANGUAGE UNDERSTANDING

Rules

Language Transparent

Language Independent

Language Dependent

okusei: Speech Recognition

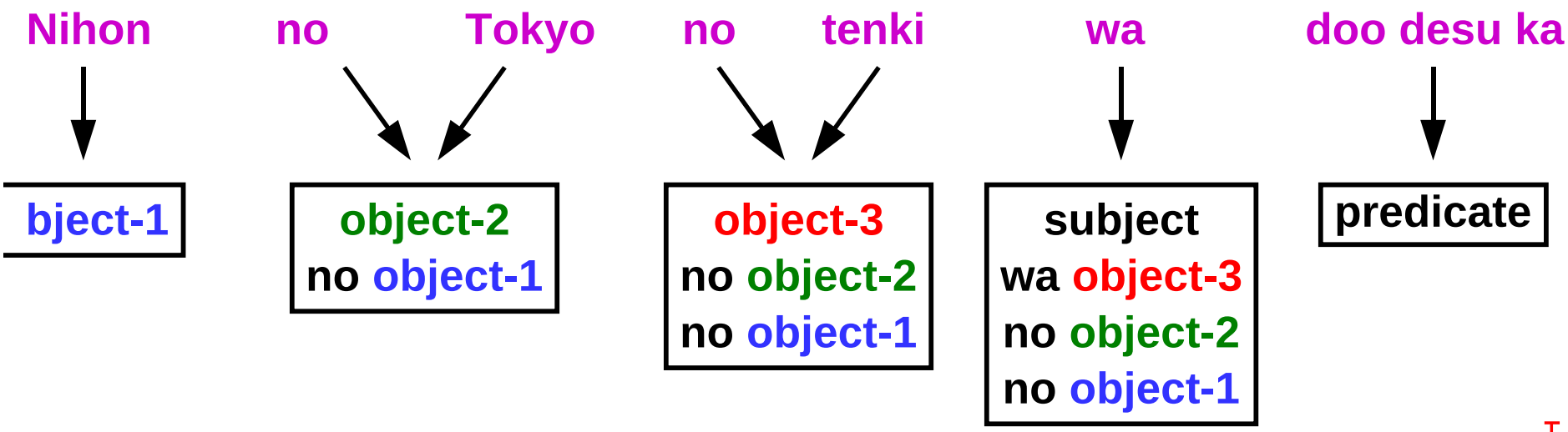
- **Lexicon: >2,000 words**
- **Phonological modeling:**
  - **Japanese specific phonological rules, e.g.,**
    * Deletion of /i/ and /u/: desu ka → /d e s k a/
- **Acoustic modeling:**
  - **Used English models to generate transcriptions for Japanese (read and spontaneous) utterances**
  - **Retrained acoustic models to create hybrid models from a mixture o English and Japanese utterances**
- **Language modeling:**
  - **Class *n*-gram using 60 word classes**
  - **Also exploring a class *n*-gram derived automatically from TINA**

okusei: Language Understanding

- **Parse query into meaning representation**
  - **Uses same NL system (TINA) as for English**
  - **Top-down parsing strategy with trace mechanism**
  - **Probability model automatically trained**
  - **Chooses best hypothesis from proposed word graph**
- **Japanese grammar contains**
  - **>900 unique nonterminals**
  - **Nearly 2,500 vocabulary items**
- **Translation file maps Japanese words to English equivalent**
- **Produces same semantic frame (i.e., meaning representation) as for English inputs**

okusei: Language Understanding (cont'd)

- **Problem: Left recursive structure of Japanese requires look-ahead to resolve role of content words**
  - **Nihon wa . . .**
  - **Nihon no tenki wa . . .**
  - **Nihon no Tokyo no tenki wa . . .**
- **Solution: Use trace mechanism**
  - **Parse each content word into structure labeled "object"**
  - **Drop off "object" after next particle, which defines role and position in hierarchy**

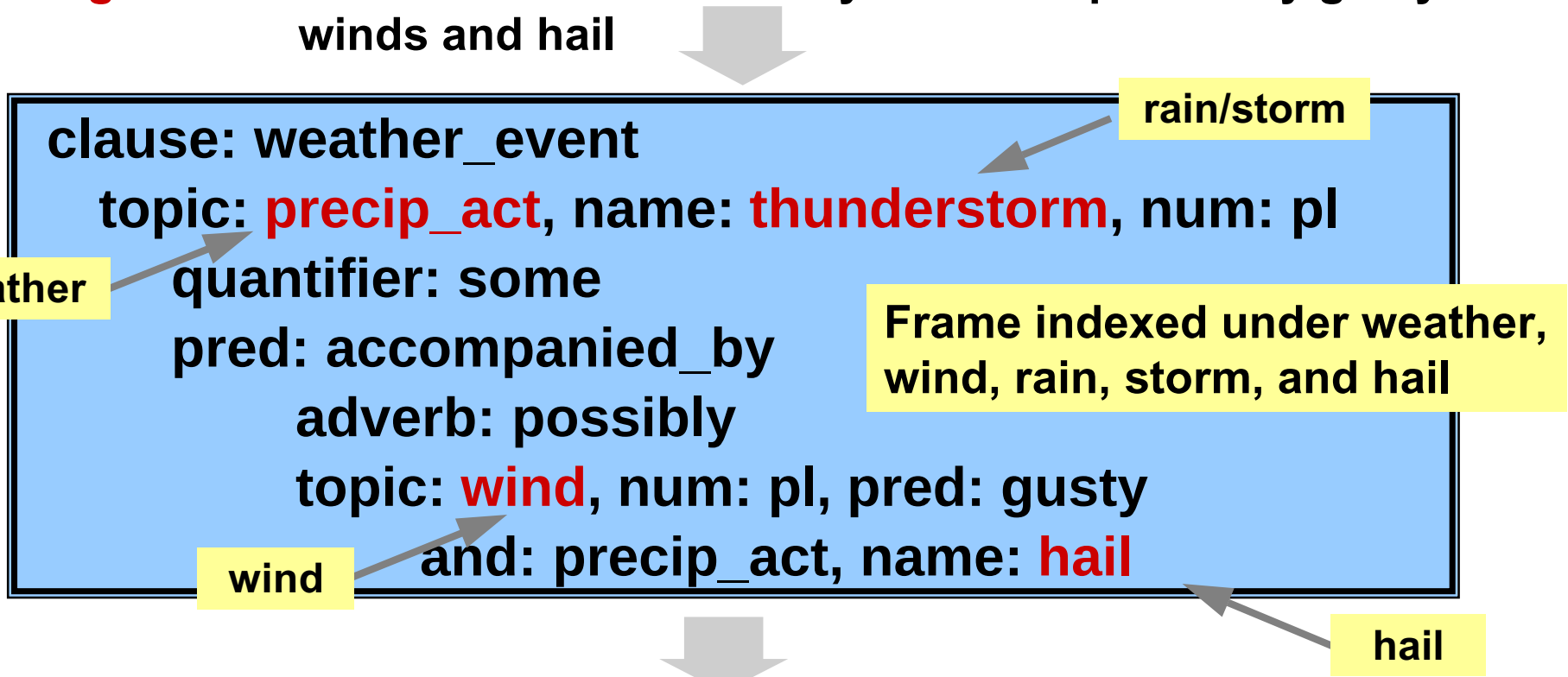


okusei: Content Processing

- **Update sources from Web sites and satellite feeds at frequent intervals**
  - **Now harvesting weather reports for ~50 additional Japanese cities**
- **Use the same representation for English and Japanese**
- **Parse all linguistic data into semantic frames to capture meaning**
- **Scan frames for semantic content and prepare new relational database table entries**

okusei: Example of Content Processing

English: Some thunderstorms may be accompanied by gusty winds and hail



Japanese: ところどころ、強いかぜそしてあられを伴う雷雨

Spanish: Algunas tormentas posiblement acompanadas por vientos racheados y granizo

German: Einige Gewitter moeglichenweise begleitet von boeigem Wind und Hagel

okusei: Language Generation Using Genesis

- **Used English language generation tables as template**
- **Modified ordering of constituents**
- **Provided translation lexicon for >4,000 words**
- **Challenges:**
  - **Prepositions had to be marked for role: *in_loc, in_time***
  - **Multiple meanings for some other words: e.g., "*well* inland"**
  - **Complex sentences presented difficulties for constituent ordering**
- **A new version of GENESIS is being developed to support finer control of constituent ordering**

okusei: Speech Synthesis

- **Currently use the NTT Fluet text-to-speech system**
- **Fully integrated into the system**
  - **Runs as a server communicating with the Galaxy hub**

okusei Demonstration

- **Entire system running at MIT**
- **Access via international telephone call**
- **Scenario: inquiring about weather conditions in Japan and worldwide**
- **Potential problems:**
  - **The system is VERY new!**
  - **System reliability**
  - **14 hour time difference**
  - **Transmission conditions and environmental noise**

Lessons Learned

- **Our approach to developing multilingual interfaces appears feasible**
  - **Performance is similar to the English system two years ago**
- **A top-down approach to parsing can be made effective for left-recursive languages**
- **Word order divergence between English and Japanese motivated a redesign of our language generation component**
- **Novel technique of generating a class *n*-gram language model using the NL component appears promising**
- **Involvement of Japanese researcher is essential**

Future Work

- **Additional data collection from native Japanese speakers**
  - **Nearly 2,000 sentences were collected in December and January**
- **Improvement of individual components**
  - **Vocabulary coverage, acoustic and language models**
  - **Parse coverage**
  - **Continued development of a more sophisticated language generation component**
- **Expansion of weather content for Japan**