C SC 620
Advanced Topics in Natural Language Processing

Lecture 21

4/13

Reading List

- *Readings in Machine Translation*, Eds. Nirenburg, S. *et al*. MIT Press 2003.
  - 19. Montague Grammar and Machine Translation. Landsbergen, J.
  - **20. Dialogue Translation vs. Text Translation – Interpretation Based Approach. Tsujii, J.-I. And M. Nagao**
  - **21. Translation by Structural Correspondences. Kaplan, R. et al.**
  - 22. Pros and Cons of the Pivot and Transfer Approaches in Multilingual Machine Translation. Boitet, C.
  - 31. A Framework of a Mechanical Translation between Japanese and English by Analogy Principle. Nagao, M.
  - 32. A Statistical Approach to Machine Translation. Brown, P. F. et al.

Tomorrow's Soldier Today

Robot soldiers! Bird-sized airplanes! The Phraselator! At DARPATech, the military shows off its coolest gadgets.

By Phillip Carter

Posted Thursday, March 11, 2004, at 2:23 PM PT

Don't speak Pashto? You do now!

Some of the displays show DARPA success stories—projects conceived by the research agency that have actually made it into production. One example is the Phraselator, a brick-sized one-way translation device designed for use by U.S. soldiers in countries where they don't know the language and don't have time to learn it. Each hand-held unit uses an SD card—the same one used by many digital cameras—that store up to 30,000 common phrases useful for law enforcement, first aid, or war-fighting. To make the device work, a soldier simply says a phrase (such as "Stop at this checkpoint") into the device, and a few seconds later, the Phraselator repeats it in the chosen language—Urdu, Arabic, Pashto, and Korean are available, to name a few. So far, more than 600 of these devices have been shipped to American units in the field—including 15 programmed with Haitian dialects dispatched with U.S. troops to Haiti. (Listen to the Phraselator's Arabic mode here.)

Phraselator

DARPA.asf

PHRASELATOR™

PRODUCTS | PHRASE MODULES | TECHNOLOGY | HISTORY | REAL WORLD APPLICATIONS | CUSTOM SOLUTIONS | SUPPORT | FAQ | PRESS | CONTACT

PHRASELATOR WINS SBIR AWARD AT 2004 **DARPA**Tech! **...read more**

VoxTec heard your plea for a practical, cost-effective means of bridging the cross-cultural communication divide. And we pioneered the solution. The **Phraselator Translation System** is a voice-activated handheld system that translates predetermined phrases into the desired foreign langauge, allowing users to accurately convey critical information in real-time — without a human translator.

How does it work? It's simple. The user speaks or selects a phrase. The **Phraselator** finds the foreign language equivalent. The translated phrase, prerecorded in the desired language, is heard through a built-in high-fidelity speaker.

Combining advanced technology and the highest linguistic standards, the **Phraselator Translation System** boasts proven reliability, efficiency, functionality and durability.

INTRODUCING THE P2

Recent Press

- **MSN Slate Gizmos: *Tomorrow's Soldier Today***
- **Day to Day from NPR News**
- **TechLink: Maryland's Tech, Biotech & Telecom Mag**
- **Military Information Technology**
- **Special Operations Technology**

Paper 20. Dialogue Translation vs. Text Translation – Interpretation Based Approach. Tsujii, J.-I. And M. Nagao

- Time: Late-80s
- Dialogue translation differs from document translation
- Argues that the goal oriented nature of dialogues makes translation more feasible than textual translation

Paper 20. Dialogue Translation vs. Text Translation – Interpretation Based Approach. Tsujii, J.-I. And M. Nagao

- Differences of Environments
  - Example of dialogues:
    - hotel reservation, conference registration, doctor-patient
  - Clear definition of information
  - Active participation of speakers and hearers
    - Writers and readers unavailable during translation
- What Should be Translated?
  - Dialogues (usually) have a purpose
  - Can define what is important and what is not

Paper 20. Dialogue Translation vs. Text Translation – Interpretation Based Approach. Tsujii, J.-I. And M. Nagao

- What Should be Translated?
  - Example:
    - [Japanese] hotel-topic, friends-with disco-to want-to-because, Roppongi-gen be-near-nom is-good
    - [SBT] As for hotel, because I would like to go to Disco with friends, to be near to Roppongi is good
    - [English Translation] Because I'd like to go to disco with friends, I prefer to stay at a hotel in Roppongi
    - SBT = Structure Bound Translation
    - *Prefer* and *stay* not in source utterance

Paper 20. Dialogue Translation vs. Text Translation – Interpretation Based Approach. Tsujii, J.-I. And M. Nagao

- Architecture of Dialogue Translation Systems
  - Extract important information from source utterances

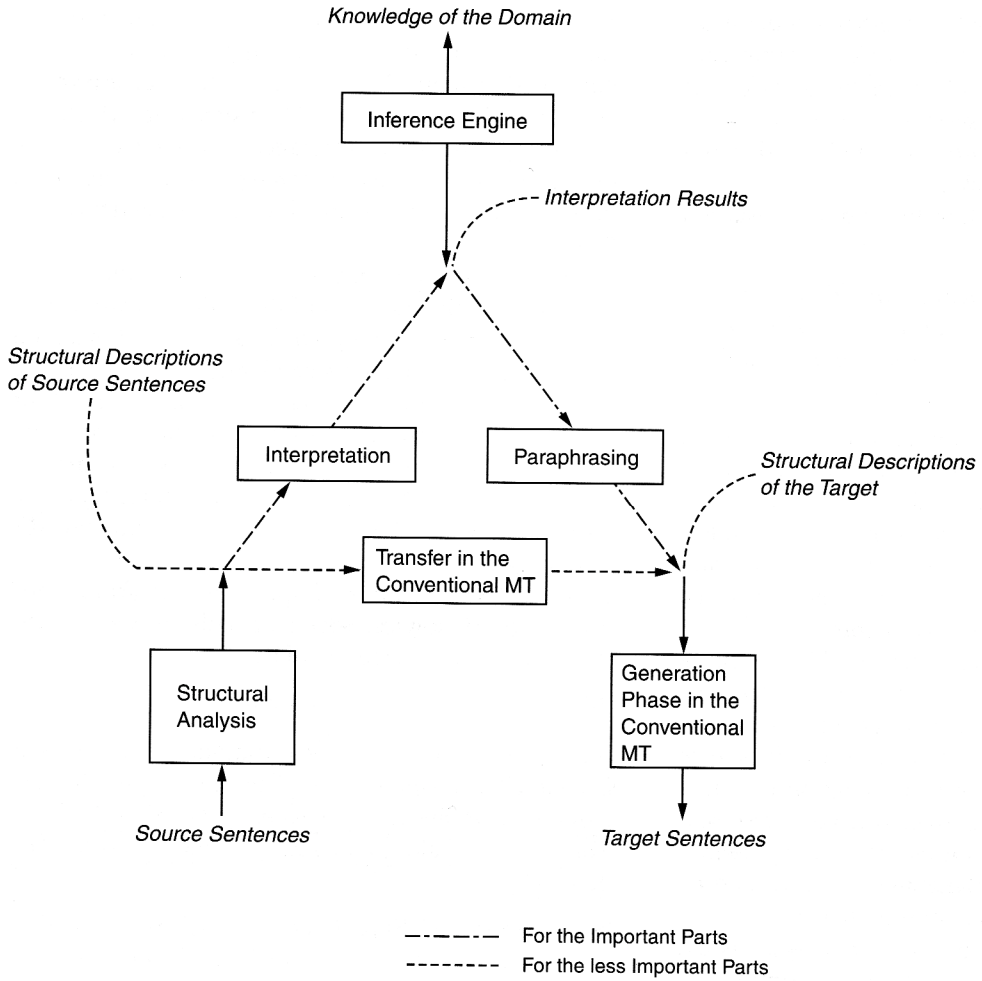

Figure 20.2

Paper 20. Dialogue Translation vs. Text Translation – Interpretation Based Approach. Tsujii, J.-I. And M. Nagao

- Examples
  - [J] Roppongi-gen be-near-gen hotel-nom good is
  - [SBT] A hotel near to Roppongi is good
  - [J] Roppongi-around-gen hotel-acc please
  - [SBT] A hotel around Roppongi, please
  - [J] Hotel-topic roppingi-no be-near-nom good is
  - [SBT] As for hotel, to be near to Roppongi is good
  - [J] be-convenient-topic roppongi-to near hotel is
  - [SBT] What is convenient is a hotel near to Roppongi

Paper 20. Dialogue Translation vs. Text Translation – Interpretation Based Approach. Tsujii, J.-I. And M. Nagao

- Dialogue translation system need not understand utterances completely, just the important bits
  - Need not translate fluently the unimportant bits
- Real world knowledge
  - Roppongi is a special region in Tokyo where many discos exist
  - In order to go to some place, it is preferable to stay at a hotel near to the place

Paper 20. Dialogue Translation vs. Text Translation – Interpretation Based Approach. Tsujii, J.-I. And M. Nagao

- Active Participation of Speakers and Hearers
  - Translation of dialogues allows for questions from user when translation does not supply necessary information or when translation cannot be understood
  - Also permits system to ask clarification questions
  - Example
    - [E] In which region do you want to stay in Tokyo?
    - [J] Disco-to want-to-go
    - [System] The question is ‘in which region do you want to stay in Tokyo?’ Would you specify the place which you prefer to stay?

Paper 21. Translation by Structural Correspondences. Kaplan, R. et al.

- Lexical-Functional Grammar (LFG)
- Claim: Modularity
  - Modularity of linguistic specifications
    - Not a single level that connects two languages
    - Instead, simultaneous correspondences
  - Permits contrastive transfer rules that depend on but do not duplicate the specifications of independently motivated grammars of source and target languages

Paper 21. Translation by Structural Correspondences. Kaplan, R. et al.

- A General Architecture for Linguistic Descriptions
  - LFG
    - c-structure (constituent)
    - f-structure (grammatical function)

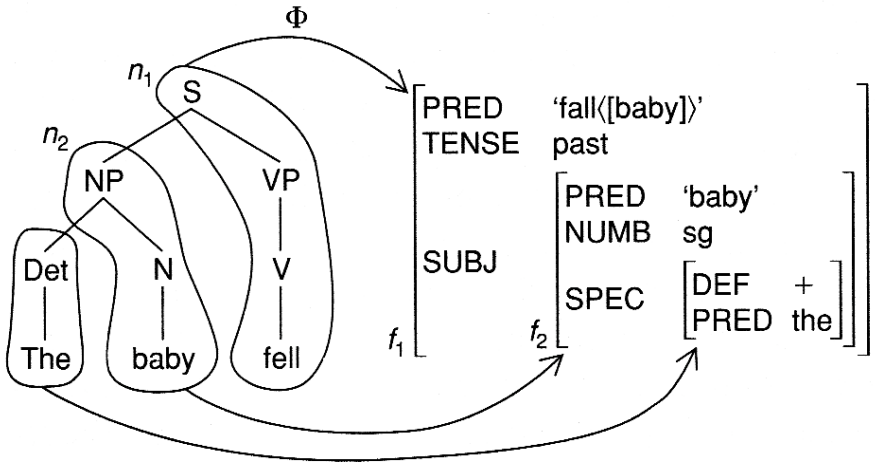

Figure 21.1

Paper 21. Translation by Structural Correspondences. Kaplan, R. et al.

- A General Architecture for Linguistic Descriptions
  - LFG
    - c-structure (constituent)
    - f-structure (grammatical function)
    - Semantic structure

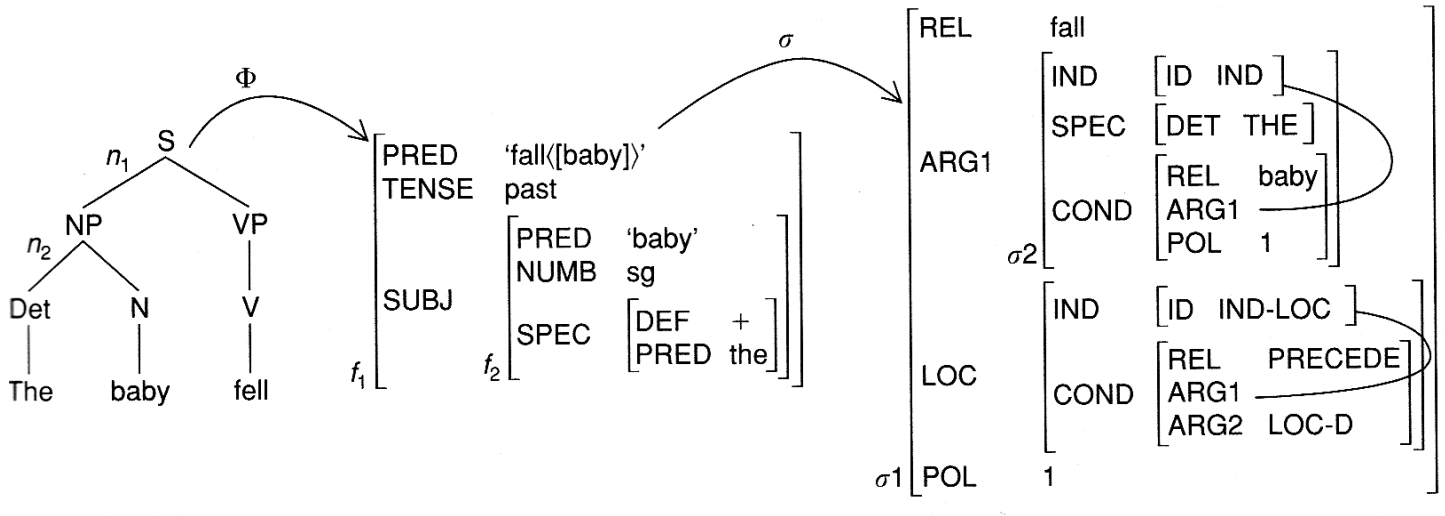

Figure 21.2

Paper 21. Translation by Structural Correspondences. Kaplan, R. et al.

- Examples
  - Change in grammatical function
    - (German) Der Student beantwortet die Frage
    - (French) L'étudiant répond à la question
  - Transitive verb in German, intransitive verb with an oblique complement in French

Paper 21. Translation by Structural Correspondences. Kaplan, R. et al.

(7) *beantworten* V
(↑ PRED) = 'beantworten ⟨(↑ SUBJ)(↑ OBJ)⟩'

while the transfer lexicon for *beantworten* contains the following mapping specifications:

(8) (τ↑ PRED FN) = répondre
(τ↑ SUBJ) = τ(↑ SUBJ)
(τ↑ AOBJ OBJ) = τ(↑ OBJ)

We use the special attribute FN to designate the function-name in semantic forms such as 'beantworten ⟨(↑ SUBJ)(↑ OBJ)⟩'. In this transfer equation it identifies *répondre* as the corresponding French predicate. This specification controls lexical selection in the target, for example, selecting the following French lexical entry to be used in the translation:

(9) *répondre* V
(↑ PRED) = 'répondre ⟨(↑ SUBJ)(↑ AOBJ)⟩'

Paper 21. Translation by Structural Correspondences. Kaplan, R. et al.

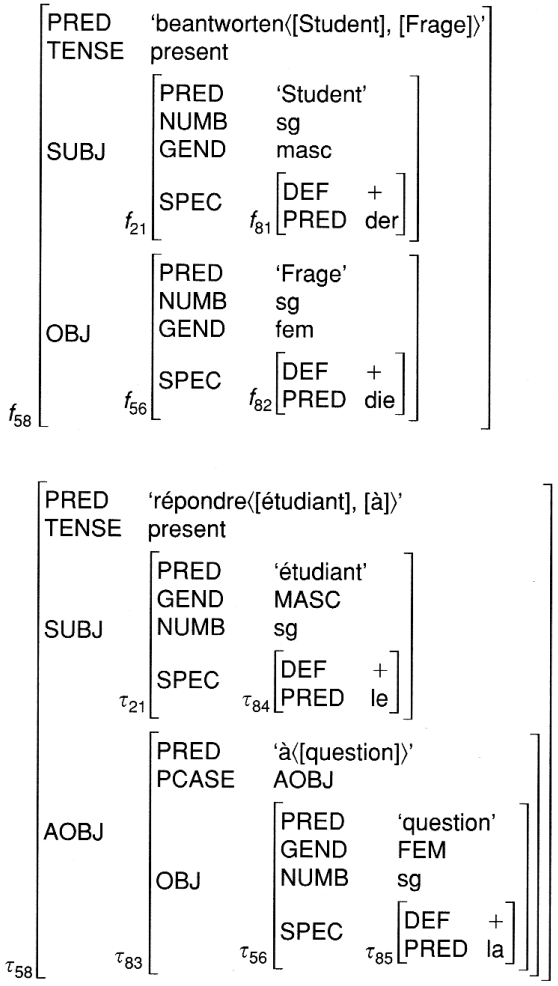

Figure 21.5

Paper 21. Translation by Structural Correspondences. Kaplan, R. et al.

- Example
  - Differences in control
    - The student is likely to work
    - It est probable que l'étudiant travaillera
    - Infinitival complement of a raising verb is translated into a finite clause

likely A

(↑ PRED) = 'likely<(↑ XCOMP)>(↑ SUBJ)'
(↑ SUBJ) = (↑ XCOMP SUBJ)

probable A

(↑ PRED) = 'probable<(↑ COMP)>(↑ SUBJ)'
(↑ SUBJ FORM) = il
(↑ COMP COMPL) = que
(↑ COMP TENSE) = future

(τ↑ PRED FN) = probable
(τ↑ COMP) = τ(↑ XCOMP)

Figure 21.6

Paper 21. Translation by Structural Correspondences. Kaplan, R. et al.

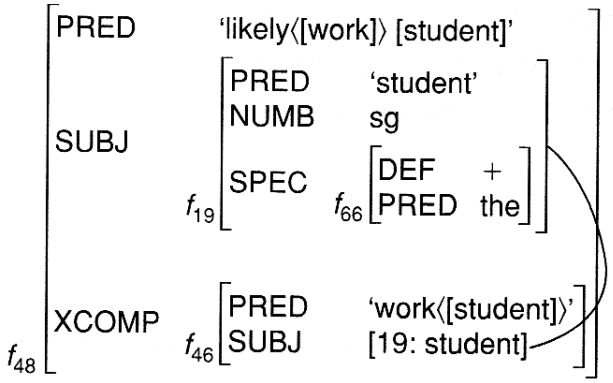

Figure 21.7

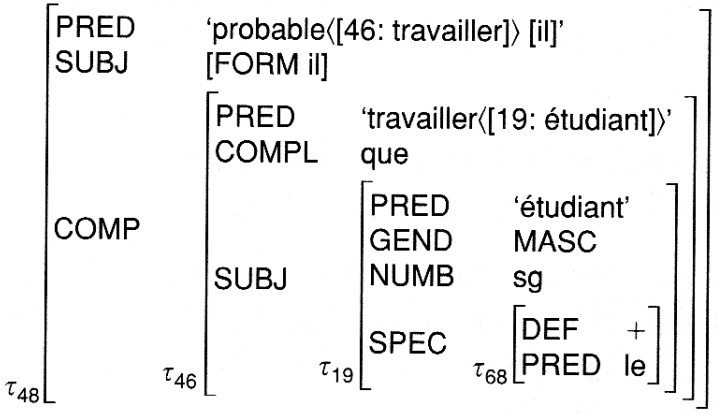

Figure 21.8

Paper 21. Translation by Structural Correspondences. Kaplan, R. et al.

- Differences in Embedding
  - The baby just fell
  - Le bébé vient de tomber

(21) *just* ADV $(\uparrow \text{PRED}) = \text{`just} \langle(\uparrow \text{ARG})\rangle\text{'}$
$(\tau\uparrow \text{PRED FN}) = \text{venir}$
$(\tau\uparrow \text{XCOMP}) = \tau(\uparrow \text{ARG})$

Paper 21. Translation by Structural Correspondences. Kaplan, R. et al.

f-structure for S

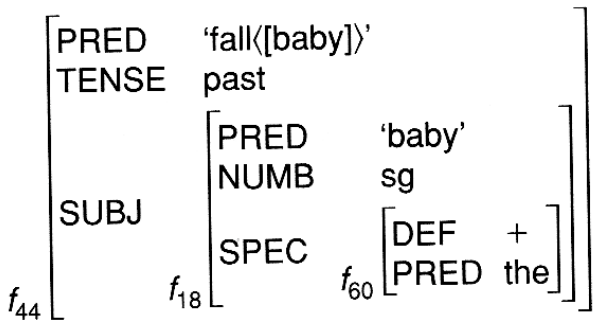

Figure 21.10

f-structure for ADV

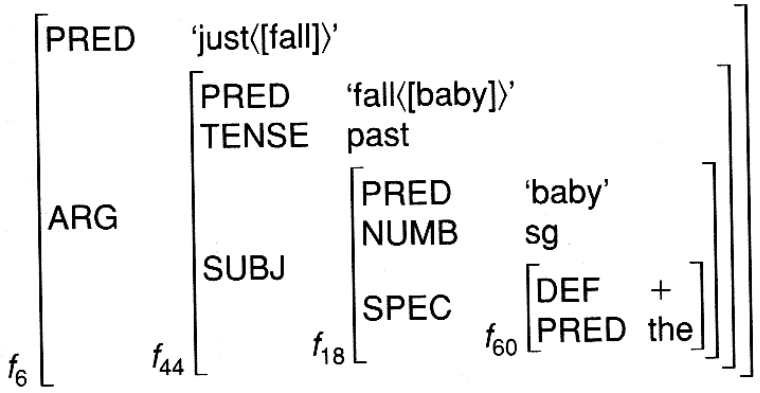

Figure 21.11

Paper 21. Translation by Structural Correspondences. Kaplan, R. et al.

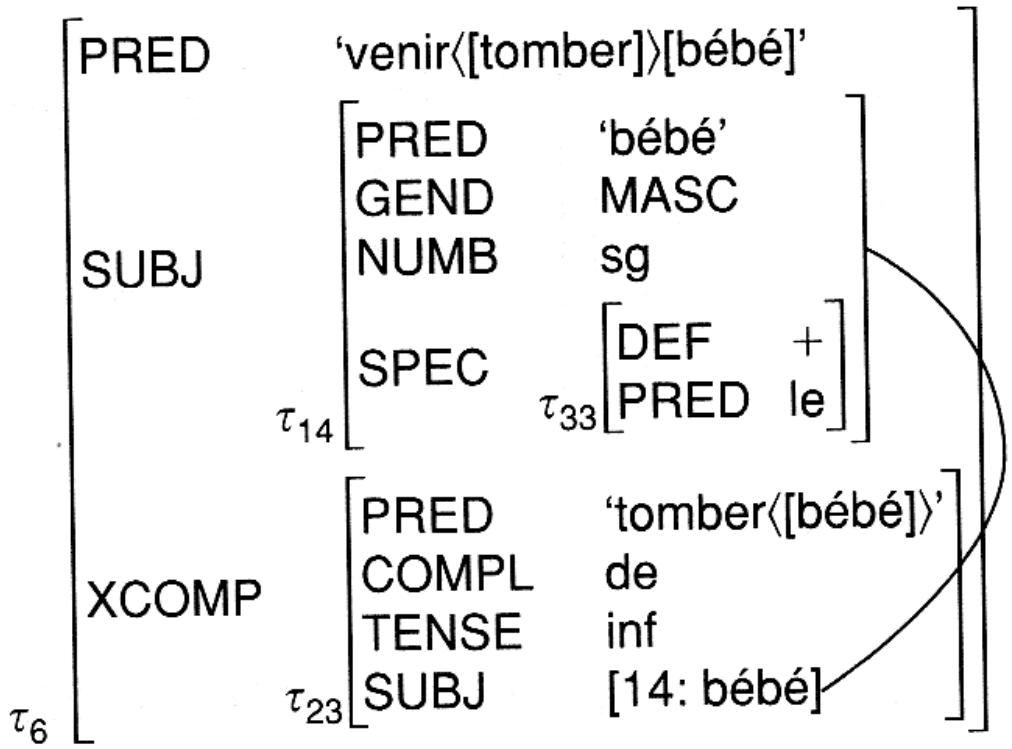

Figure 21.12