

C SC 620
Advanced Topics in Natural
Language Processing

3/9

Lecture 14

Machine Translation

- *Readings in Machine Translation*, Eds. Nirenburg, S. *et al.* MIT Press 2003.
- Part 1: Historical Perspective
- Reading list:
 - *Introduction*. Nirenburg, S.
 - 1. *Translation*. Weaver, W.
 - 3. *The Mechanical Determination of Meaning*. Reifer, E.
 - **5. *A Framework for Syntactic Translation*. Yngve, V.**
 - 6. *The Present Status of Automatic Translation of Languages*. Bar-Hillel, Y.

Paper 5: A Framework for Syntactic Translation. V. Yngve

- Introduction: MIT Approach
 - “We are attempting to go beyond simple word-for-word translation: beyond translation using empirical, *ad hoc*, or pragmatic syntactic routines. The concept of full syntactic translation has emerged: translation based on a thorough understanding of linguistic structures, their equivalences, and meanings.”

Paper 5: A Framework for Syntactic Translation. V. Yngve

- Contextual clues:
 1. The field of discourse
 - specialized dictionaries
 2. Recognition of coherent word groups
 - idioms and compound nouns
 3. The syntactic function of each word
 - help solve word order problems as well as multiple-meaning problems
 - (German) *der*
 - article, relative or demonstrative pronoun
 - aominative, genitive or dative

Paper 5: A Framework for Syntactic Translation. V. Yngve

- Contextual clues:
 4. Selectional relations between open class items
 - nouns, verbs, adjectives and adverbs
 5. Antecedents
 - pronouns etc.
 6. All other contextual cues, especially those concerned with an exact knowledge of the subject under discussion
 - ... last to be mechanised

Paper 5: A Framework for Syntactic Translation. V. Yngve

- Two Approaches
 - 95% approach
 - Full understanding
- Decoding verb, noun and adjective inflection
 - Can do a perfect job: one need not be satisfied with anything less
- Verb category translation
 - ? Auxiliary, ergative, unergative
 - Imperfectly understood

Paper 5: A Framework for Syntactic Translation. V. Yngve

- Syntactic Translation
 - Examination of the 6 types of clues above reveals that they are predominantly concerned with the relationships of one word to another in patterns
 - #3 (syntactic function of each word), is basic to the others

Paper 5: A Framework for Syntactic Translation. V. Yngve

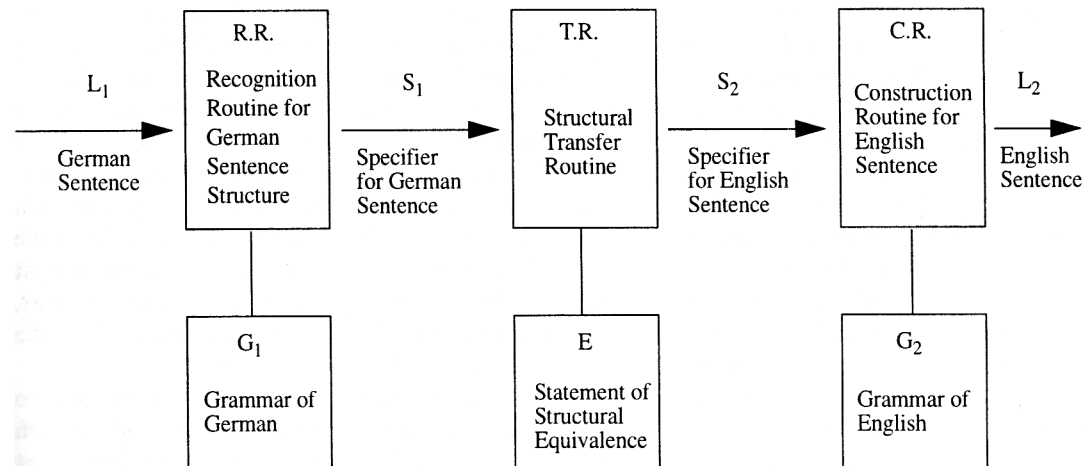
- Syntactic Translation
 - The central role of syntax
 1. Field of discourse
 - must use the relation of words in syntactic patterns as the key for finding which words refer to which field
 2. Recognition of coherent word groups
 - Idioms, noun compounds, and so on, are merely special patterns of words that stand out from more regular patterns
 - Modern Note:
 - Idioms are usually constituents and obey the rules of syntax
 - E.g. V O (VP) much more common than S V
 - Very few exceptions to the rules of syntax: *by and large*, *by the by*

Paper 5: A Framework for Syntactic Translation. V. Yngve

- Syntactic Translation
 - The central role of syntax
 - 4. Selectional relations between open class items
 - Selectional relations between words that are syntactic related
 - Modern Note:
 - E.g. subject of *wake up* must be animate
 - 5. Antecedents
 - The relationship of a word to its antecedent is essentially a syntactic relationship
 - Modern Note:
 - Binding Theory? C-command relation
 - *Mary saw herself*
 - **Mary's father saw herself*

Paper 5: A Framework for Syntactic Translation. V. Yngve

- The Framework (Hypothetical)



Paper 5: A Framework for Syntactic Translation. V. Yngve

- Specifier = structural description = “message” or transition language
- Tree-to-tree transfer model
- Stored Knowledge
 - G_1 , E and G_2 are declarative descriptions
 - Not procedural
 - Procedural information in R.R., T.R. and C.R. boxes
- “Earlier estimates that the amount of storage necessary for syntactic information may be of the same order of magnitude as the amount of storage required for a dictionary have not been revised.”

Paper 5: A Framework for Syntactic Translation. V. Yngve

- Construction (C.R.)
 - Presumably does things like re-orders the constituents for English word order SVO (vs. German word order V2)

Paper 5: A Framework for Syntactic Translation. V. Yngve

- Specifiers
 - The specifier of a sentence represents that sentence as a series of choices within the limited range of choices prescribed by the grammar of the language
 - Sentence: affirmative or negative
 - Subject: modified by relative clause or not
 - Finite Verb: person, number, tense
 - Verb: which class
 - Ambiguity
 - Unspecified coordinates

Paper 5: A Framework for Syntactic Translation. V. Yngve

- Specifiers
 - Narrow vs. broad specifiers
 - Examples:
 - Auxiliary verb *can* - present, past (narrow)
 - *will be able to* ... future (broad)
 - *has been able to* ... perfective (broad)
 - Auxiliary verb *must* - present (narrow)
 - *had to* - past
 - Present tense for future time
 - “*He is coming soon*”
 - Two type of specifiers for each language
 - 5 step translation procedure

Paper 5: A Framework for Syntactic Translation. V. Yngve

- Recognition

- The question of how we understand a sentence is a valid one for linguists, and it may have an answer different from the answer of how we produce a sentence.
- But it appears that the description of a language is more easily couched in terms of synthesis of sentences than in terms of analysis of sentences. The reason is clear.
 - A description in terms of synthesis is straightforward and unambiguous. It is a one-to-one mapping of specifiers into sentences.
 - But a description in terms of analysis runs into all of the ambiguities of language that are caused by the chance overlapping of difference patterns: a sentence may be understandable in terms of two or more different specifiers.
- Description in terms of analysis will probably not be available until after we have the more easily obtained descriptions in terms of synthesis.

Paper 5: A Framework for Syntactic Translation. V. Yngve

- Transfer of Structure
 - The real compromises in translation reside in these center boxes. It is here that the difficult and perhaps often impossible matching of sentences in different languages is undertaken.
 - But the problems associated with the center box are not peculiar to mechanical translation.
 - Human translators also face the very same problems.
 - The only difference is that at present the human translators are able to cope satisfactorily with the problem.

Paper 5: A Framework for Syntactic
Translation. V. Yngve