

C SC 620
Advanced Topics in Natural
Language Processing

Lecture 20

4/8

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.*

Translating is EU's new boom industry



BBC NEWS WORLD EDITION

Last Updated: Thursday, 8 April, 2004, 01:55 GMT 02:55 UK

[E-mail this to a friend](#) [Printable version](#)

Translating is EU's new boom industry



By **Angus Roxburgh**
BBC News Online, Brussels

When 10 new countries join the European Union on 1 May, they bring with them an extra nine languages to add to the EU's existing 11.

There could even be 10 new tongues, for if Greek and Turkish Cypriots vote for reunification before then, Turkish will become the EU's 21st language.



How will it cope? Even with 20, Europe's tower of Babel is creaking.

Translators, builders and electronics suppliers are busy ahead of 1 May

Twenty languages gives a total of 190 possible combinations (English-German, French-Czech, Finnish-Portuguese, etc), and finding any human being who speaks, for example, both Greek and Estonian or Slovene and Lithuanian is well-nigh impossible.

BBC SPORT
BBC WEATHER
BBC ON THIS DAY

LANGUAGES

РУССКИЙ

POLSKI

SHQIP

ΕΛΛΗΝΙΚΑ

SRPSKI

TÜRKÇE

MORE >

To get round this problem, the parliament will use much more "relay translation", where a speech is interpreted first into one language and then into another - and perhaps into a fourth or fifth.

Clearly the scope for mistakes in this game of Chinese whispers is huge.

"If I'm first in the chain, and make a mistake, then everyone else down the relay makes the same mistake - or worse," Jana Jalvi, one of the new Estonian recruits says.

The need for translation already takes away the cut and thrust of a normal parliamentary debate.

When the Italian Prime Minister, Silvio Berlusconi, last year likened a German MEP to a Nazi camp guard, it took several seconds before the German realised he was being insulted and pulled off his headphones in disgust.

But the rule is that every language must be provided.

"The European Parliament is the one place you can't expect people to speak a foreign language," Patrick Twidle, who is in charge of recruiting new interpreters, says.

“ Translation costs less than 2 euros per citizen, so it is less than a cup of coffee ”

Juhani Lonnroth
Head of translation service

"The European Parliament is the one place you can't expect people to speak a foreign language," Patrick Twidle, who is in charge of recruiting new interpreters, says.

"People are elected not because of their language skills but to represent their political constituency."

The European Commission already has 1,300 translators, who process 1.5 million pages a year in the EU's 11 languages.

In two years that is expected to rise to almost 2.5 million pages - and the staff, based in two enormous buildings in Brussels and Luxembourg, will almost double in size to cope with the output.

The cost will rise from roughly 550 million euros today to over 800 million euros after enlargement.

Is it worth it?

Juhani Lonroth, the Finn who runs the translation service, has done his sums.

"Translation costs less than 2 euros per citizen, so it is less than a cup of coffee or a ticket to the cinema," he says.

"I think it's worth it because it is part of democracy."

Before and after

- ◆ European Commission has 1,300 translators
- ◆ They process 1.5 million pages a year
- ◆ They cost the EU 550 million euros
- ◆ After 1 May, staff will almost double in size
- ◆ They will translate 2.5 million pages a year
- ◆ Their budget will be over 800 million euros

Paper 19. Montague Grammar and Machine Translation. Landsbergen, J.

- Year is 1985
- Montague Grammar
 - Meaning as Higher-Order Intentional Logic
 - Compositional
 - Meaning of an expression is a function of the meaning of its parts
 - Close mapping between syntax and semantics
 - Possible-world semantics

Paper 19. Montague Grammar and Machine Translation. Landsbergen, J.

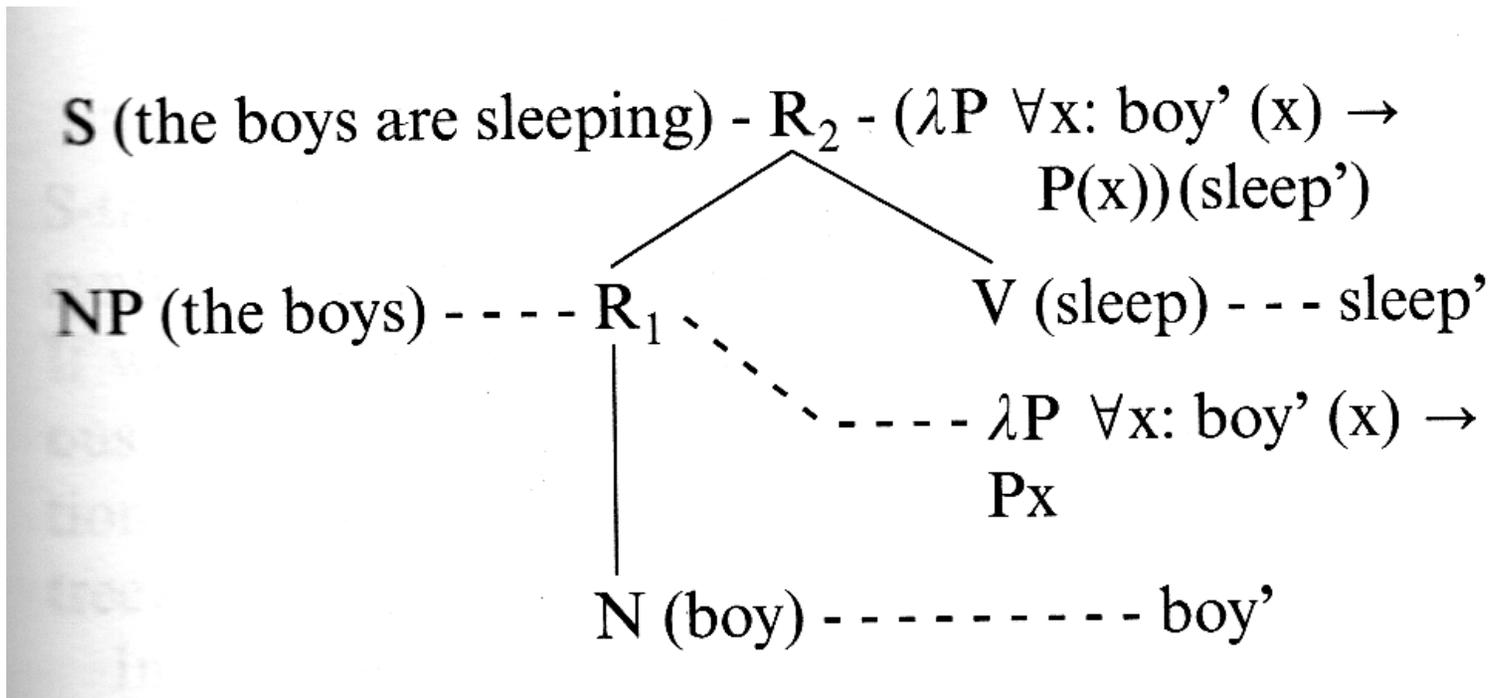


Figure 19.3

The boys are sleeping $\rightarrow \forall x (\text{boy}'(x) \rightarrow \text{sleep}'(x))$

Paper 19. Montague Grammar and Machine Translation. Landsbergen, J.

- Montague Grammar and Computer Applications
 - Strong and weak points?
 - Attention given to semantics
 - Sound semantic base is needed for determining what a correct answer or a correct translation is...
 - NLP Q&A
 - Machine Translation
 - Advantage over some other linguistic theories
 - Exactness and constructiveness
 - Syntax and semantics defined locally over phrase composition rules
 - cf. Grammar with several syntactic levels, where the semantics is defined at the deepest level

Paper 19. Montague Grammar and Machine Translation. Landsbergen, J.

- Montague Grammar and Computer Applications
 - Strong and weak points? (contd.)
 - Weak syntax
 - Incidental property of Montague's examples
 - Intentional logic and possible-world semantics too complex for practical use
 - Purely generative framework
 - Syntax and semantics in parallel

Paper 19. Montague Grammar and Machine Translation. Landsbergen, J.

- M-grammars
 - Transformational power
- Consists of:
 - Syntactic component
 - Morphological component
 - Semantic component

Paper 19. Montague Grammar and Machine Translation. Landsbergen, J.

- Syntactic Component

- S-tree
- Nodes: category + attr/val pairs
- Edges: syntactic relations

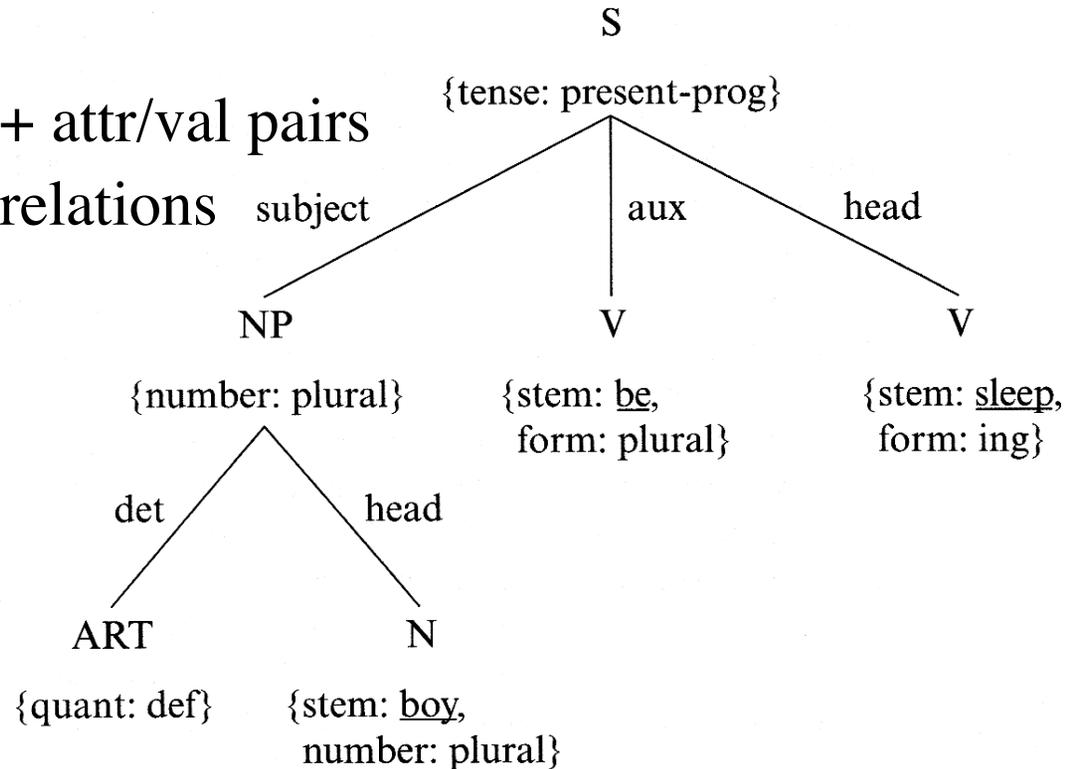


Figure 19.4

Paper 19. Montague Grammar and Machine Translation. Landsbergen, J.

- Rules must be bidirectional to serve as input to
 - M-Parser
 - M-Generator
- Termination of transformational rules guaranteed by measure condition
 - E.g. number of nodes in a tree must be decreasing
- Surface syntax condition
 - Covering grammar?
 - S-PARSER

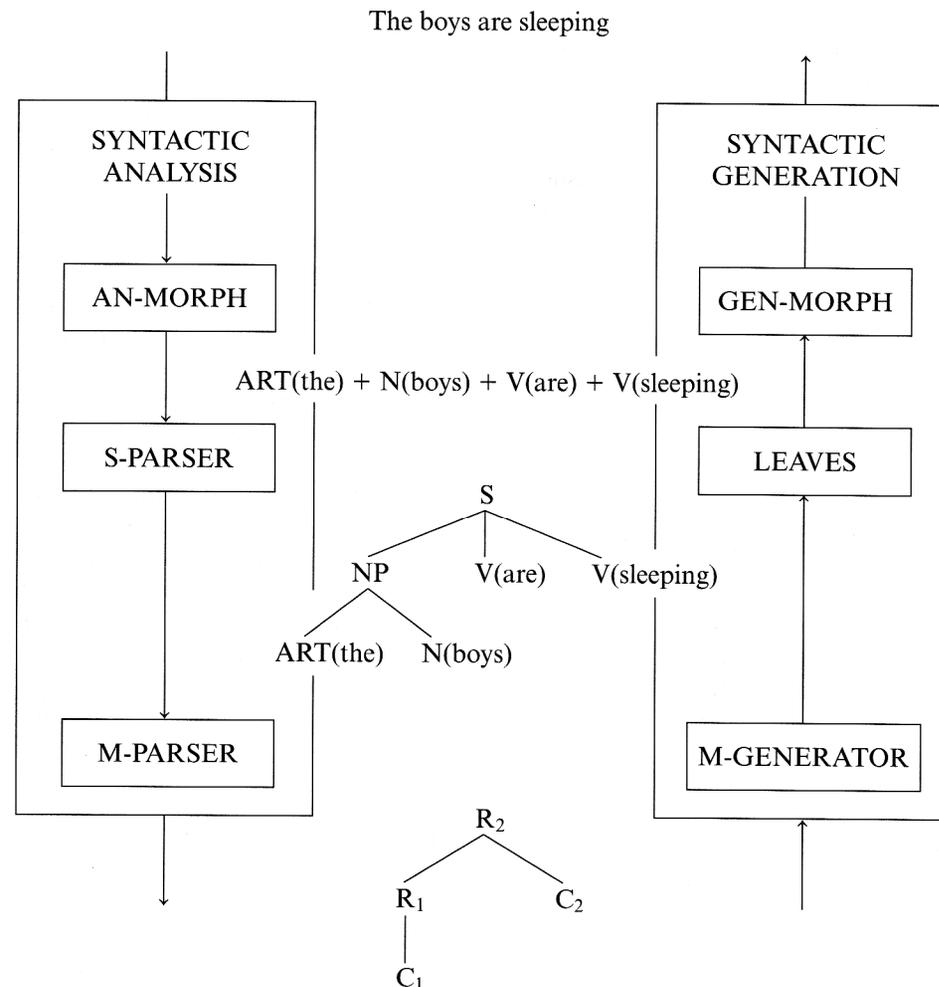


Figure 19.5

Paper 19. Montague Grammar and Machine Translation. Landsbergen, J.

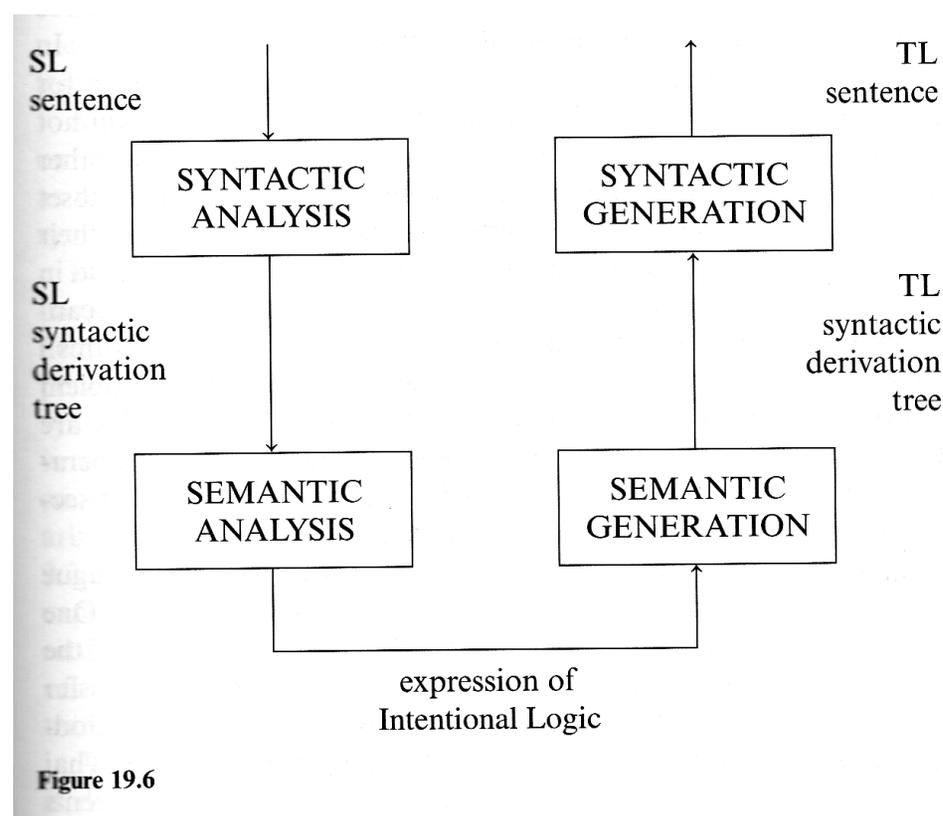
- Morphological Component
 - A-MORPH: words \rightarrow terminal S-trees
 - G-MORPH: terminal S-trees \rightarrow words

Paper 19. Montague Grammar and Machine Translation. Landsbergen, J.

- Montague Grammar and Machine Translation
 - “Possible Translation” System
 - Assumptions
 - Linguistic theory can be clearly separated from the other aspects (extralinguistic information, robustness measures, etc.)
 - Isolated sentences only
 - F-PTR: source language (SL) sentence \rightarrow set of possible translations in the target language (TL)
 - $s' \text{ in F-PTR}(s) \leftrightarrow s \text{ in F-PTR}(s')$
 - Explicit grammars for SL and TL
 - Correctness-preserving property of F-PTR
 - Common information content between source and target sentence

Paper 19. Montague Grammar and Machine Translation. Landsbergen, J.

- Attractive model but there are problems with Intentional Logic as an interlingua
 - Discrepancy between MG literature (detailed semantics for small fragment) vs. what is needed for MT (wide coverage, superficial semantics)
 - Doesn't convey pragmatic and stylistic information
 - Subset problem



Paper 19. Montague Grammar and Machine Translation. Landsbergen, J.

- Subset problem
- Need transfer rules from IL_1 to IL_2

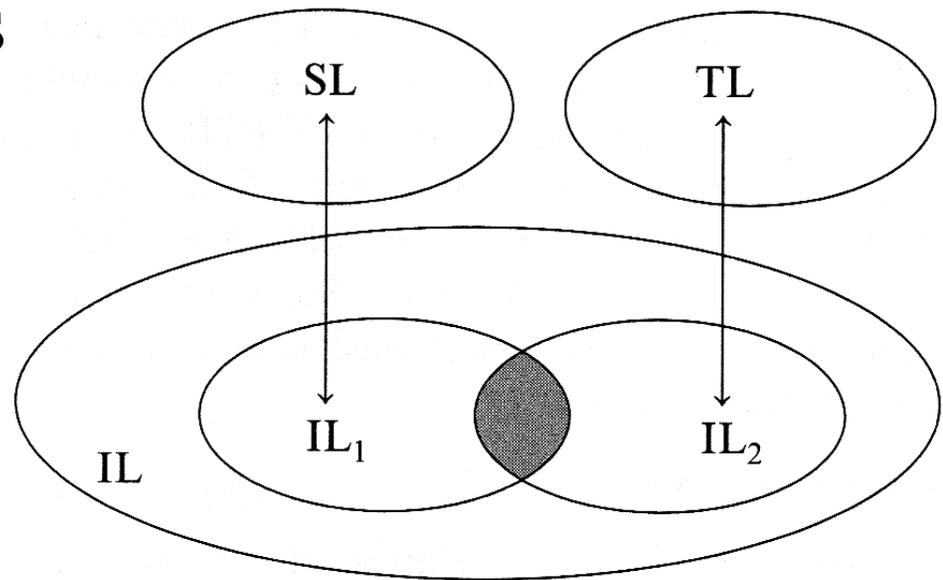


Figure 19.7

Paper 19. Montague Grammar and Machine Translation. Landsbergen, J.

- Take Intentional Logic out
- Or eliminate TL grammar by transfer of terms of the logical expression obtained from Syntactic Analysis

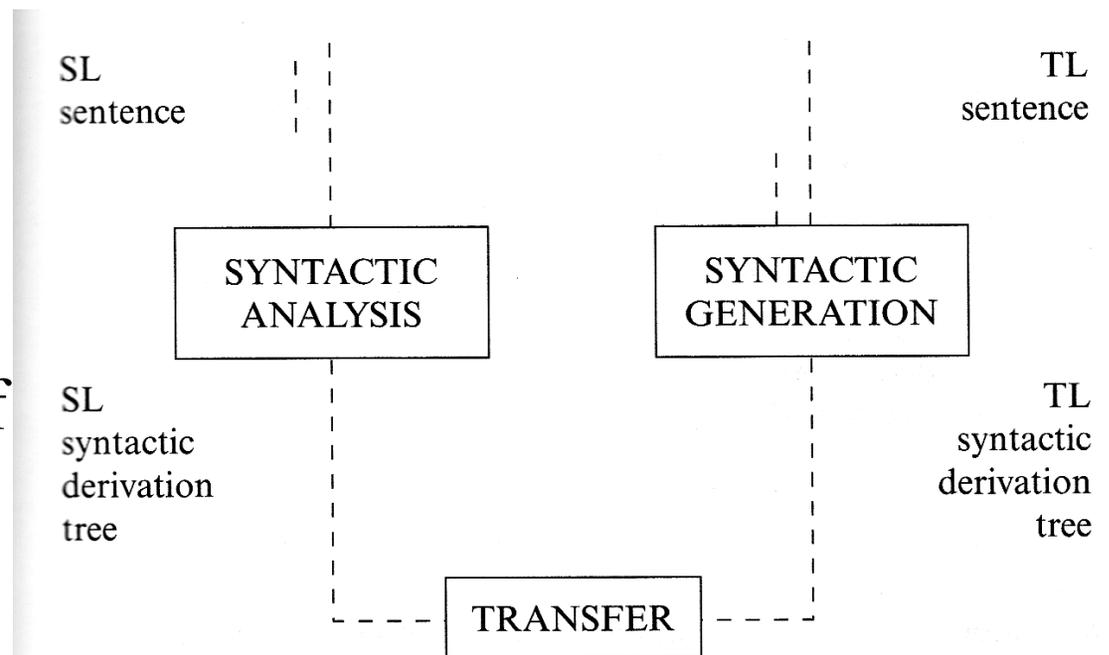


Figure 19.8

Paper 19. Montague Grammar and Machine Translation. Landsbergen, J.

- Isomorphic M-grammars
 - Each expression in one language must have (at least) one corresponding basic expression in the other language with the same meaning
 - Each syntactic rule in one language must have (at least) one corresponding syntactic rule in the other language with the same meaning operation
 - Two sentences are translations of each other if they are derived from corresponding basic expressions by application of corresponding rules

Paper 19. Montague Grammar and Machine Translation. Landsbergen, J.

- Isomorphic M-grammars

DUTCH

basic expressions

N (jongen)

V (slaap)

syntactic rules

NR₁:

N (jongen) → NP (de jongens)

NR₂

NP (de jongens) + V (slaap) → S
(de jongens slapen)

(a) The boys are sleeping.

(b) De jongens slapen.

ENGLISH

basic expressions

N (boy)

V (sleep)

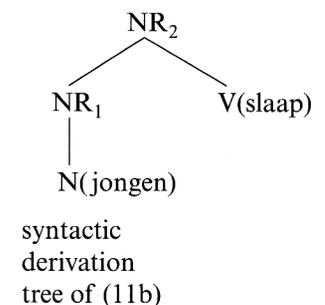
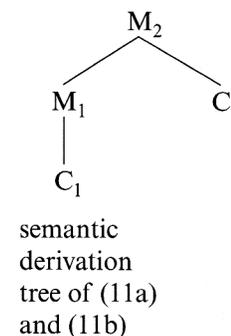
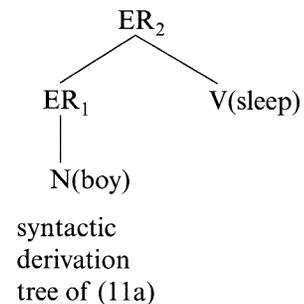
syntactic rules

ER₁:

N (boy) → NP (the boys)

ER₂

NP (the boys) + V (sleep) →
S (the boys are sleeping)



Paper 19. Montague Grammar and Machine Translation. Landsbergen, J.

- Interlingual system
 - But not universal

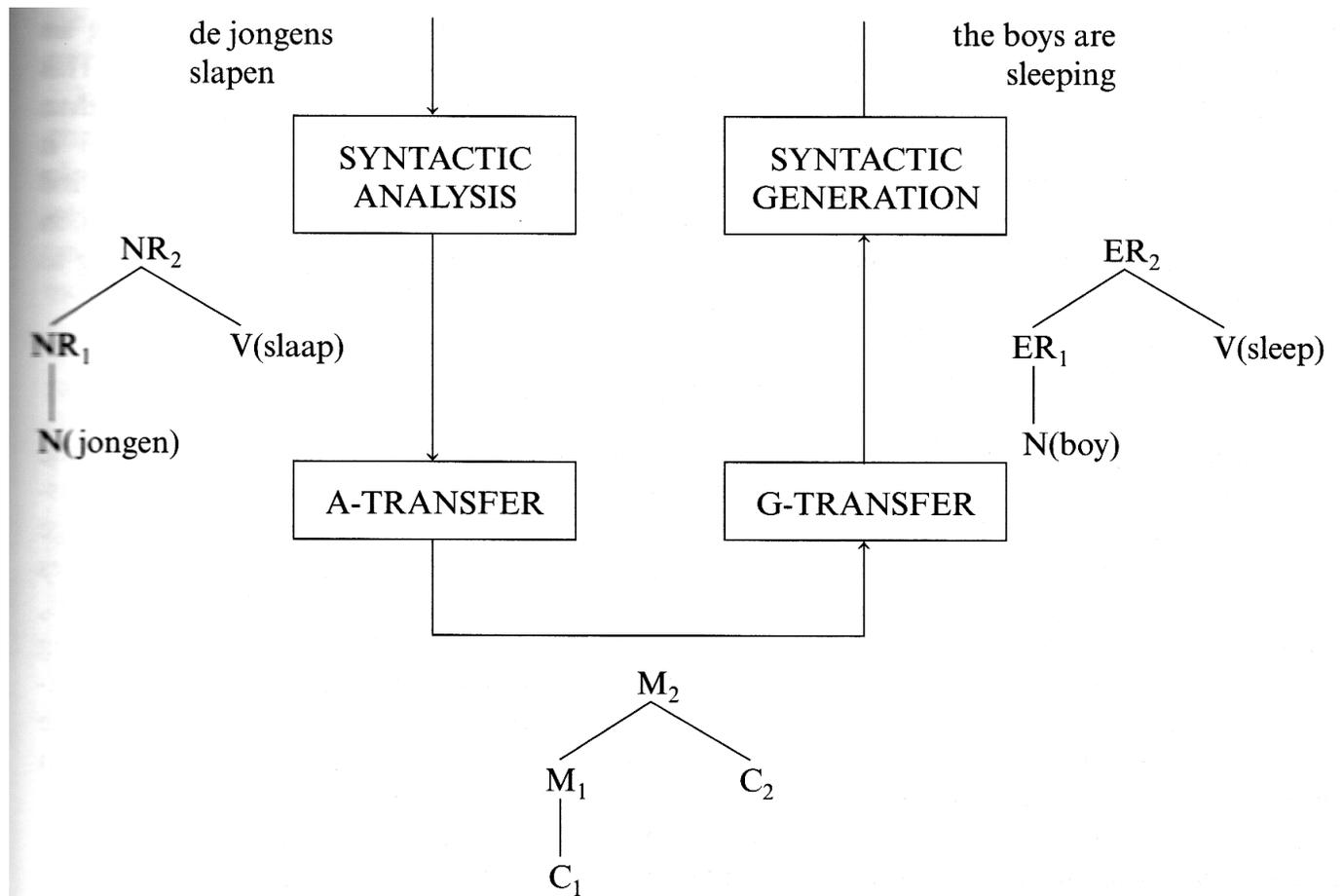


Figure 19.10