Spoken Language Interaction

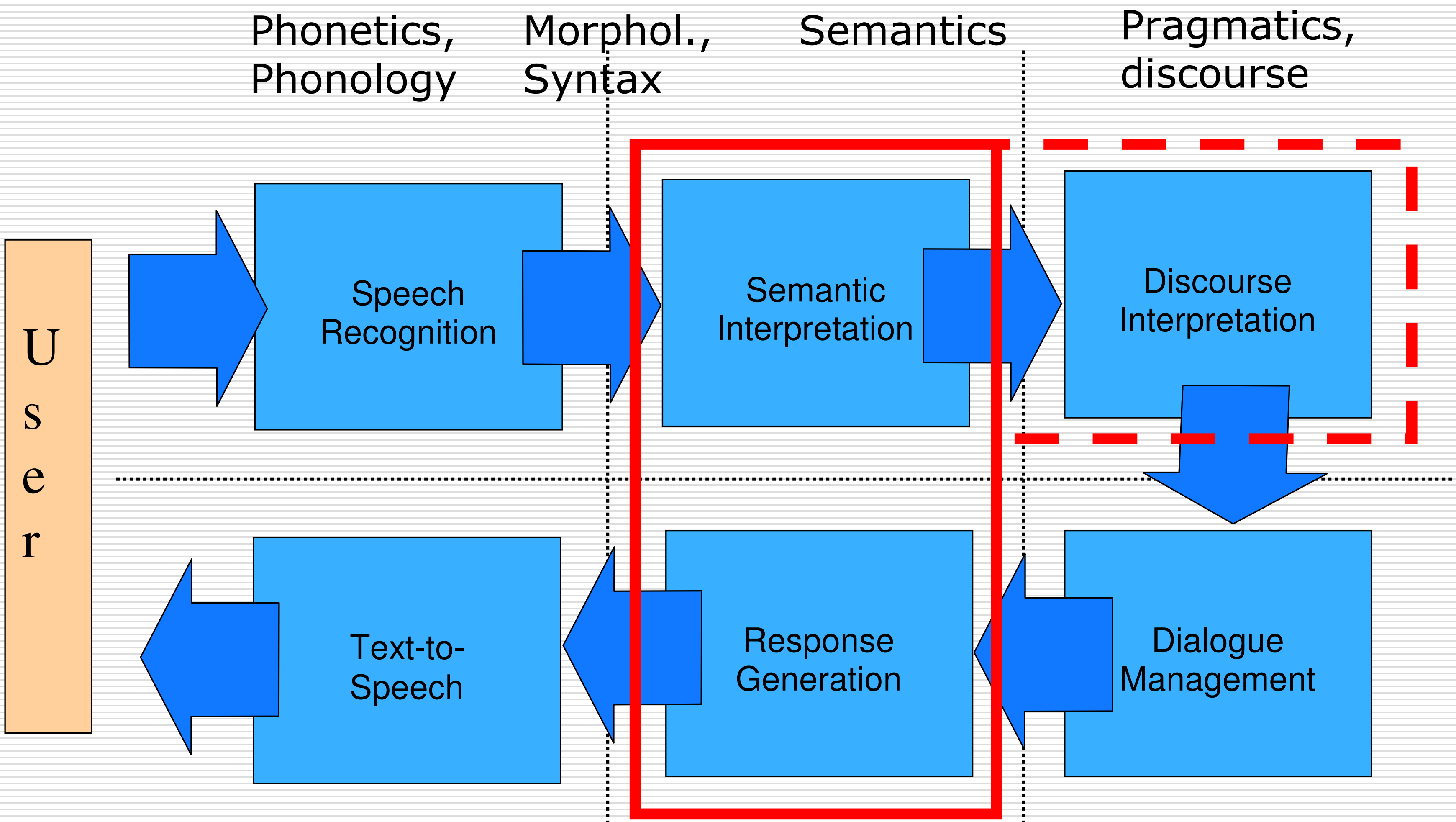
Natural language
…understanding (NLU)
…generation (NLG)

Jurafsky & Martin (2000) *Speech and Language Processing*. Prentice Hall.

Spoken Dialogue Systems

- A system that allows a user to *speak* his queries in natural language and receive useful spoken *responses* from it
- Provides an interface between the user and a computer-based application that permits *spoken interaction* with the application in a "relatively natural manner"

Spoken Dialogue System - overview



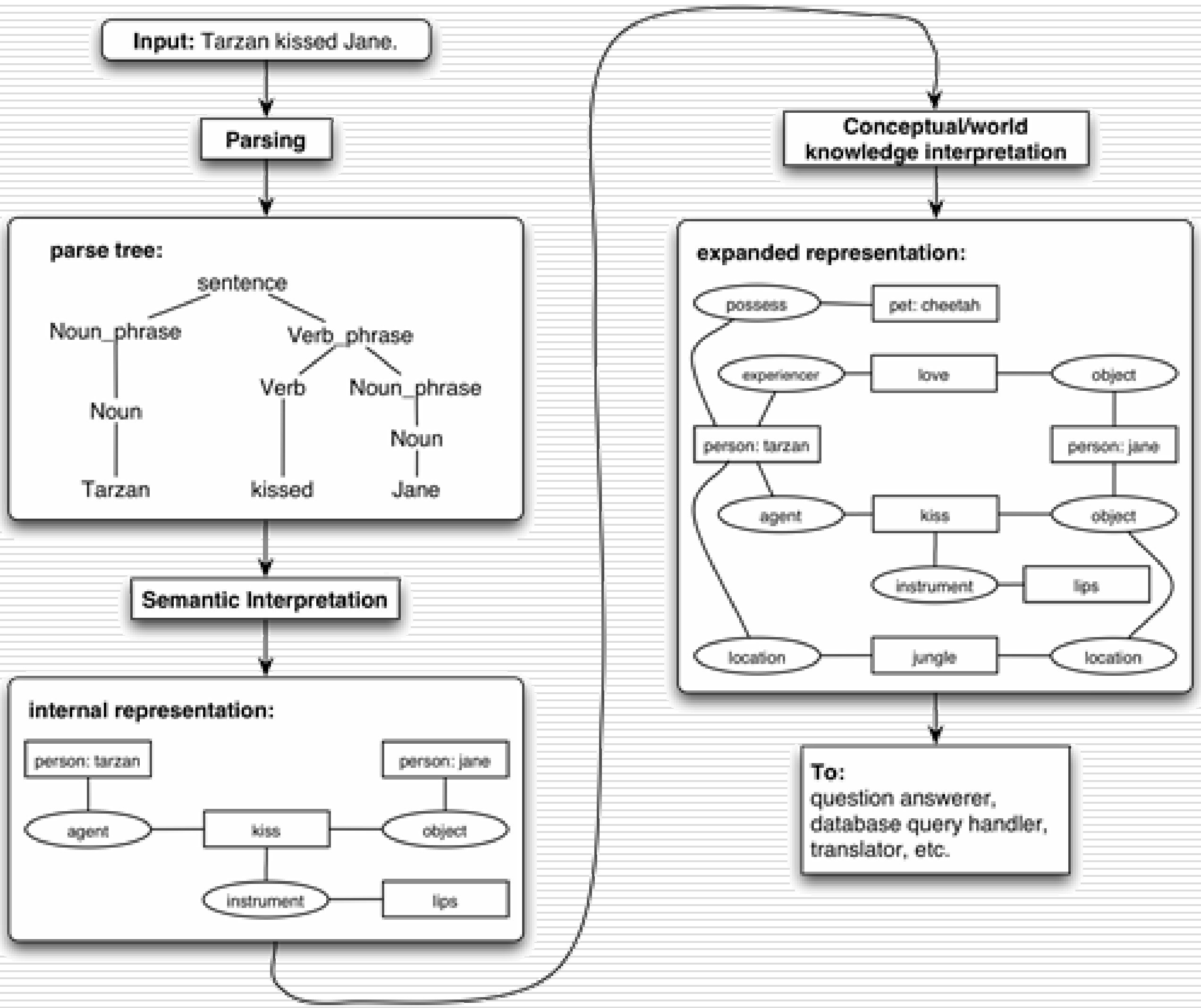
Natural language understanding

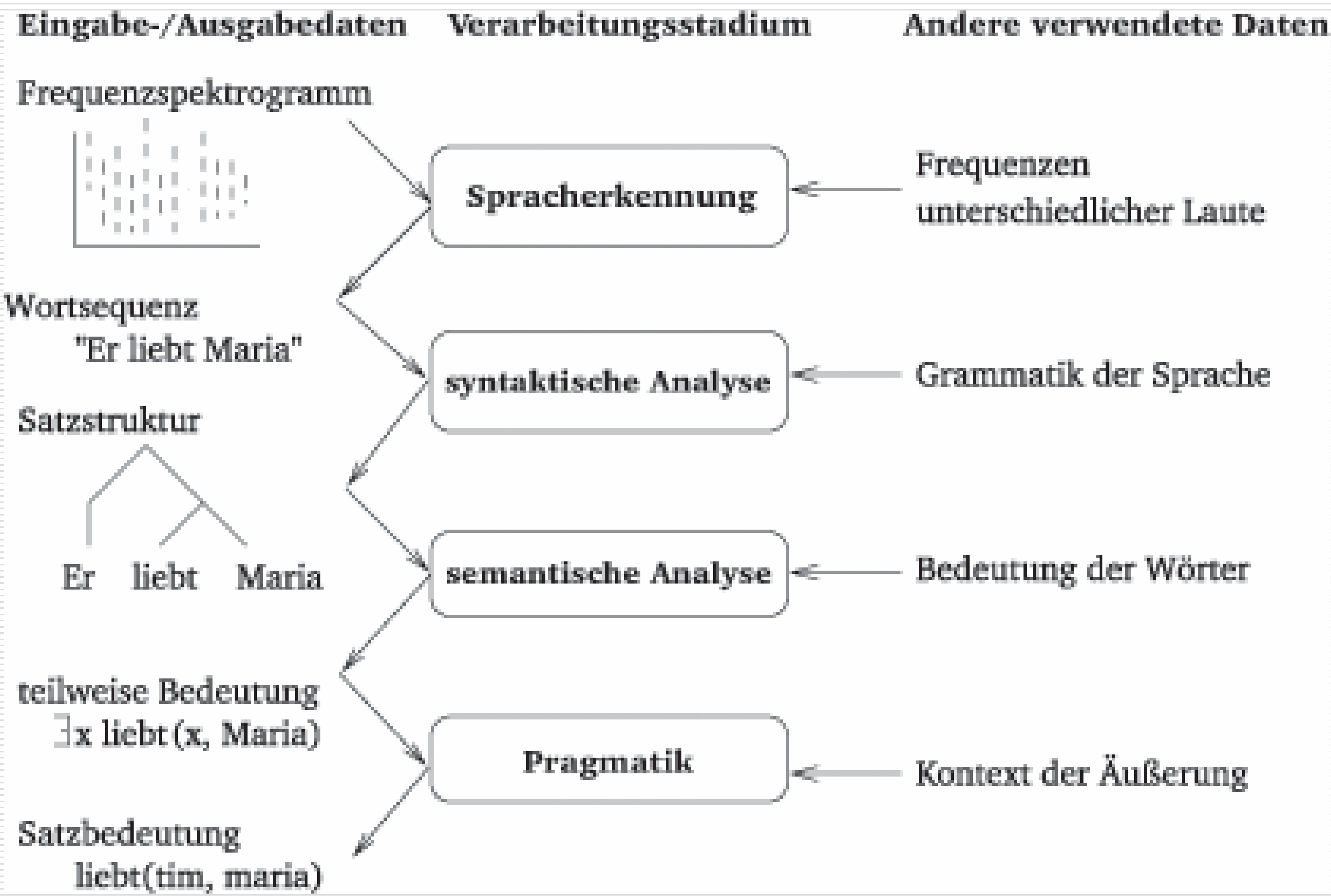
Syntax analysis

Semantic & discourse interpretation

Syntax

Semantics




Eingabe-/Ausgabedaten
Verarbeitungsstadium
Andere verwendete Daten
Frequenzspektrogramm
Spracherkennung
Frequenzen unterschiedlicher Laute
Wortsequenz "Er liebt Maria"
syntaktische Analyse
Grammatik der Sprache
Satzstruktur
Er liebt Maria
semantische Analyse
Bedeutung der Wörter
teilweise Bedeutung
∃x liebt (x, Maria)
Pragmatik
Kontext der Äußerung
Satzbedeutung
liebt(tim, maria)

Natural language understanding

- Syntax analysis:
  - Determine sentence structure from sequence of words
- Semantic interpretation/analysis:
  - Determine word meanings and the overall meaning of their composition in the sentence
- Discourse interpretation/pragmatic analysis:
  - Use context information to complete and disambiguate sentence meaning
  - Determine intention behind the sentence

Syntax analysis

Grammars

Parsing

Syntax analysis

Ziel: Baumartige Zerlegung des sprachlichen Ausdrucks in seine Komponenten gemäß einer *Grammatik*

```
PARSE ("the dog is dead", G):
        [S: [NP: [Article: the][Noun: dog]]
        [VP: [Verb: is][Adjective: dead]]]
```

- Ausgangspunkt für Bedeutungsanalyse: Komposition der Teilbedeutungen (*kompositionale Semantik)*
  (1) *Der Hund fraß den Knochen*
  (2) *Der Knochen wurde vom Hund gefressen*
- Struktur aus Syntaxregeln hilft der Bedeutungsfindung
  - Nicht: „es ist immer das 2. Substantiv, das gefressen wird"
  - Sondern: auf Basis von syntaktischen Zerlegungen

  (3) *sp[Der Hase mit den langen Ohren*] *erfreute sich an* sp[*einem großen grünen Salatblatt*]

Grammars

- Prinzipielle, endliche Beschreibung der *Struktur* der Elemente einer (evtl. unendlichen) Sprache

- Zentrale Eigenschaften
  - *Korrekt*: Generiert nur wohlgeformte (wf) Ausdrücke
  - *Vollständig*: Jeder wf. Ausdruck ist generierbar
  - *Adäquat*: korrekt und vollständig für eine Sprache
  - *Entscheidbar*: Wortproblem immer lösbar
  - *Effizienz*: Wortproblem im *worst case* effizient lösbar

- Elementarformen:
  - Phrasenstrukturgrammatiken (PS-Grammatik)
  - Kategorialgrammatiken (C-Grammatik)

Phrasenstrukturgrammatik

- Basiert auf *Konstituentenstruktur*
- Beispiel: NP -> Det Adj N beschreibt einfache Sätze wie
  - „der alte Mann"
  - „ein junger Mann"
  - „dem großen Haus"
- Menge von Regeln bilden PS-Grammatik (PSG)
- PSG sind formale Regelgrammatiken

Grammatik	$G = (\Sigma, \Phi, P, S)$
Terminale	$\Sigma = \{alte, dem, der, ein, gro\beta en, Haus, junger, Mann, Vater, ...\}$
Nonterminale	$\Phi = \{Adj, Det, N, NP, ...\}$
Regeln	$P = \begin{Bmatrix} S \rightarrow NP\,VP \mid ..., NP \rightarrow Det\ N \mid Det\ Adj\ N \mid ... \\ Det \rightarrow dem \mid der \mid ein \mid ..., Adj \rightarrow alte \mid junger \mid ..., \\ N \rightarrow Haus \mid Mann \mid Vater \mid ... \end{Bmatrix}$
Startsymbol	S

PS-Grammatiken

- Weitere, abgeleitete Formalismen
  - Syntactic structures
  - Generative semantics
  - (Revised) (Extended) Standard theory (R)(E)ST)
  - Government and binding (GB)
  - Generalized phrase structure grammar (GPSG)
  - Lexical functional grammar (LFG)
  - Head-driven phrase structure grammar (HPSG)

Probleme mit PS-Grammatiken

- NP -> Det Adj N beschreibt auch Sätze wie
  - „Der alter Vater"
  - „Ein junges Mann"
  - „Dem großen Häusern"

  Regel generiert über → Grammatik inkorrekt
- Deutsch erfordert Übereinstimmung von Kasus, Genus, Numerus innerhalb einer NP, Adjektiva flektieren abhängig vom Determiner
- NPs kontextfrei beschreibbar:
  - Aufspaltung der Nichtterminale nach *agreement-Merkmalen* und Ersetzung durch Menge neuer Regeln
  - ➢ Große Menge an Regeln
  - ➢ Generalisierung geht verloren

Mächtigkeit von PS-Grammatiken

- Reguläre Sprachen (z.B. a^k)
  - Entscheidbar in O(n)
- Kontextfreie Sprachen (z.B. $a^k b^k$)
  - Entscheidbar in $O(n^k)$
- Kontextsensitive Sprachen (z.B. $a^k b^k c^k$)
  - Entscheidbar in $O(k^n)$
- Rekursiv aufzählbare Sprachen/allg. Regelsprachen
  - Unentscheidbar

- Natürliche Sprachen sind nicht kontextfrei beschreibbar
- *Mild-kontextsensitive* Varianten (Joshi, 1985)
  - Parsing-Problem lösbar in polynomialer Zeit
  - Längenmonotonie
  - Endliche obere Schranke für die Anzahl der kreuzweisen Abhängigkeiten (z.B. Wörter der Form: $x_1 \ldots x_n \ldots y_1 \ldots y_n$)

Kategorialgrammatik

- Basiert auf Funktor-Argument-Strukturen
- Grundidee: Nicht nur Kategorien wie Noun & Verb, sondern auch *komplexe* Kategorien
- Komplexe Kategorien sind Phrasentypen, denen noch Teile fehlen können
  - S/VP : Funktor, der eine VP als Argument nimmt und einen Satz (S) liefert
  - V/(NP/S): Funktor, der ein Verb liefert und einen NP-Funktor nimmt, der einen Satz nimmt
  - Argument kann rechts oder links stehen:

    A/B ... I need a B to my right to become an A

    $A \setminus B$ ... I need a B to my left to become an A

C-Grammatiken

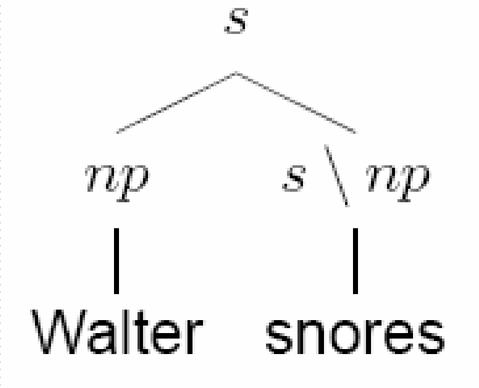
G = <W,C,LX,R,CE>

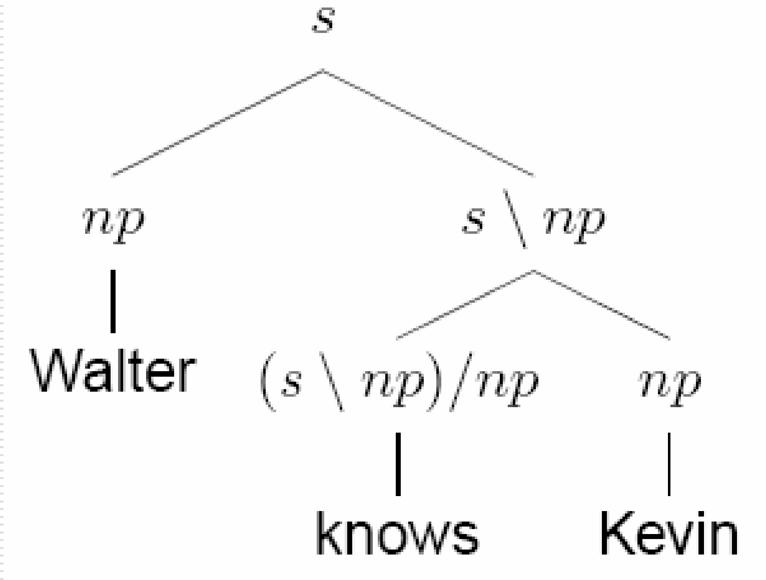
- W: endliche Menge von Wortformoberflächen
- C: Menge von Kategorien, rekursiv definiert nach
  - Vorbedingung: u,v aus C (elementare Kategorien)
  - Induktion: X,Y aus C → (X/Y) und (X\Y) aus C
- LX: endliche Teilmenge von (W x C)
  - definiert welche Wörter welche Kategorien nehmen
  - Elemente geschrieben als A_Y, $F_{(Y/X)}$
- R: Regelmenge mit Regeln der Form
  - $F_{(X/Y)}\ A_Y \rightarrow FA_X$ *(Cancellation)*
  - $A_Y\ F_{(X\backslash Y)} \rightarrow AF_X$ *(Backward Cancellation)*
- CE: Menge der Kategorien (aus C) vollständiger Ausdrücke

C-Grammatik - Beispiel

CE = {s}; LX =

- Walter, Kevin : np
- snores : s\np
- knows : (s\np)/np





C-Grammatiken

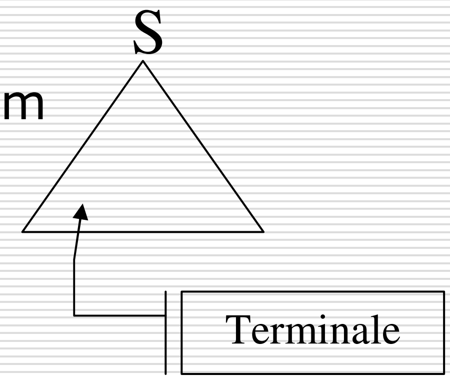
- Probleme
  - Kompliziert, komplex, ineffizient
  - Nicht offensichtlich, wo in einer Eingabe die Ableitung beginnen soll
  - Hoher Grad an lexikalischer Ambiguität nötig, um alternative Wortstellungen in Kategorien zu kodieren
- Abgeleitete Formalismen
  - Montague-Grammatik (MG)
  - Functional Unification Grammar (FUG)
  - Categorial Unification Grammar (CUG)
  - Combinatory Categorial Grammar (CCG)
  - Unification-based Categorial Grammar (UCG)

Baumadjunktionsgrammatiken

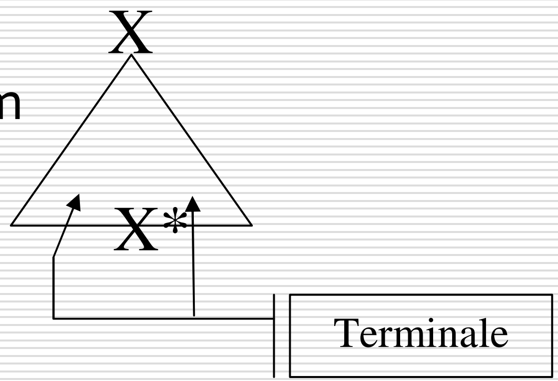
- *Tree Adjoining Grammars* (TAG) (Joshi, 1985)
  - Konkatenation als zweidimensionale Baumstruktur
- Zwei Baumarten:
  - *Initialbaum* (IB): Wurzel=Startsymbol S, Blätter sind ausschließlich Terminalsymbole
  - *Auxiliarbaum* (AB): Blätter sind beliebige Anzahl von Terminalsymbolen und genau ein Nichtterminal, dieses ist das gleiche wie die Wurzel des AB.
- IBs repräsentieren gültige Sätze der Sprache
- *Adjunktion* als Standardoperation:
  - AB kann einen beliebigen Teilbaum ersetzen, wenn er dessen Wurzel „matched“
  - Erzeugt komplexe Sätze (Bäume)

TAGs

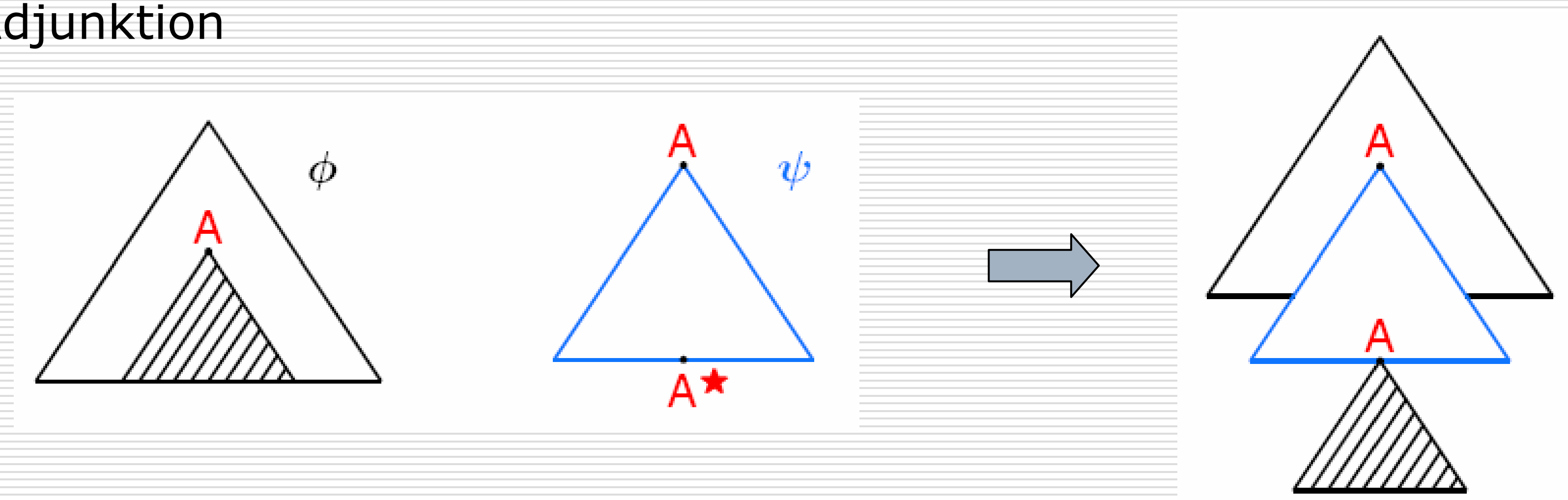
Iinitialbaum



Auxiliarbaum

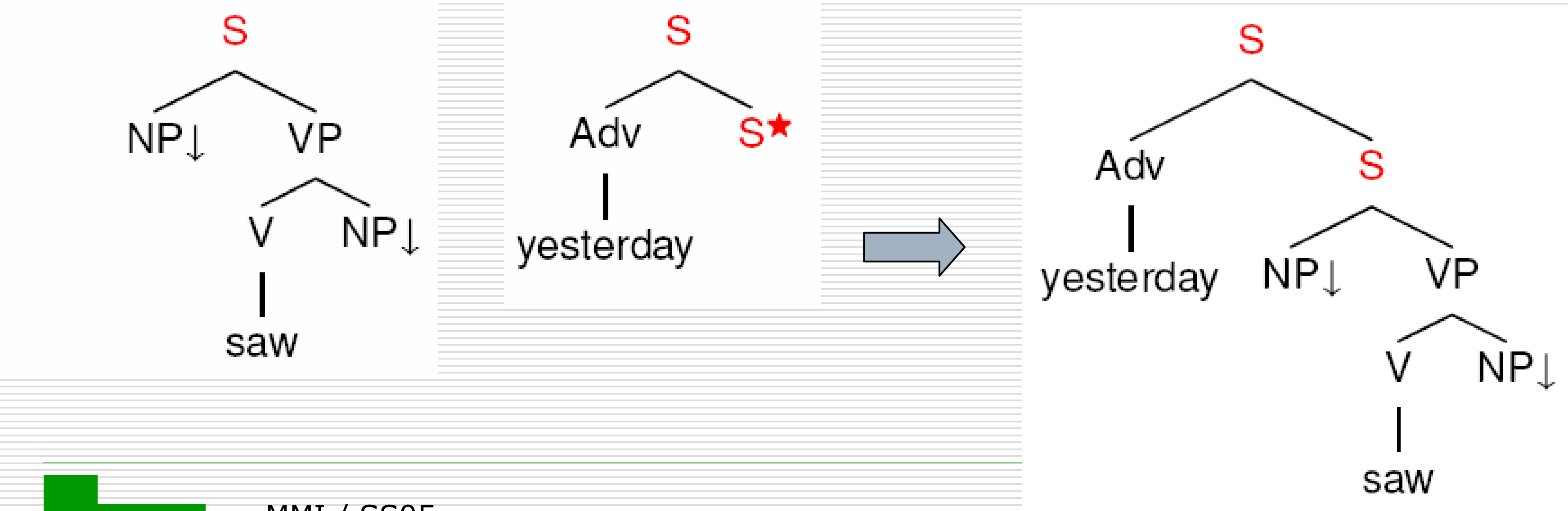


Adjunktion



TAG

- G = <N, T, I, A, S>
  - N: Menge von Nichtterminalsymbolen
  - T: Menge von Terminalsymbolen
  - I: Menge von Initialbäumen
  - A: Menge von Auxiliarbäumen
  - S: Startsymbol aus N



TAGs

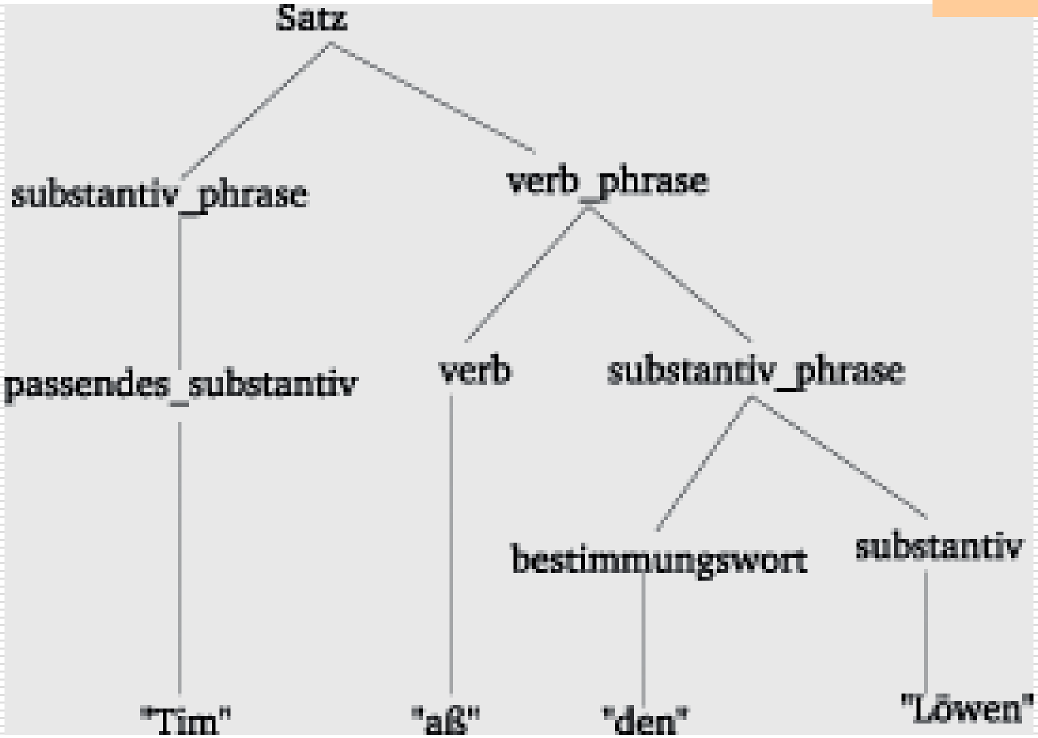
- Oftmals verwendet für natürliche Sprache
  - Große Menge an Elementarbäumen vorhanden
  - z.B. XTAG-Projekt
    http://www.cis.upenn.edu/~xtag/

- Mächtigkeit:
  - können kontextabhängige Sprachen dargestellen
  - können auch Bäume zuweisen, welche durch eine PSG nicht erzeugbar wären

Parsen

- Suchen nach einer möglichen Ableitung eines Satzes in einer Grammatik
- Top-down, bottom-up, chart parser, ATN parser, …
- Ergebnis: Ableitungsbaum
  - Beispiel für „*Tim aß den Löwen*“

Siehe z.B. VL „Spezielle Methoden der KI“



Semantic & discourse interpretation

Lexical & compositional semantics
Discourse interpretation
Reference resolution
Intention recognition, speech act theory

Interpretation

- Aufgabe: *Bedeutungsrekonstruktion*
  - Was ist die *Bedeutung* der folgenden natürlich-sprachlichen Eingabekette: *„Er beginnt um zwei im Raum V2-122."*

- Unterscheide:
  - *Semantisches Potential:* Linguistisch bestimmte Bedeutung, lässt sich allein mit linguistischem Wissen ermitteln

$$Begin(e,t,l) \wedge Event(e) \wedge Time(t) \wedge Location(l)$$

$$\wedge\, Equal(t,2) \wedge Room(l, V2-122, ?b)$$

  - *Aktueller semantischer Wert:* Volle Interpretation unter Anwendung nicht-linguistischens Wissens (Kontext, Domäne, Welt):

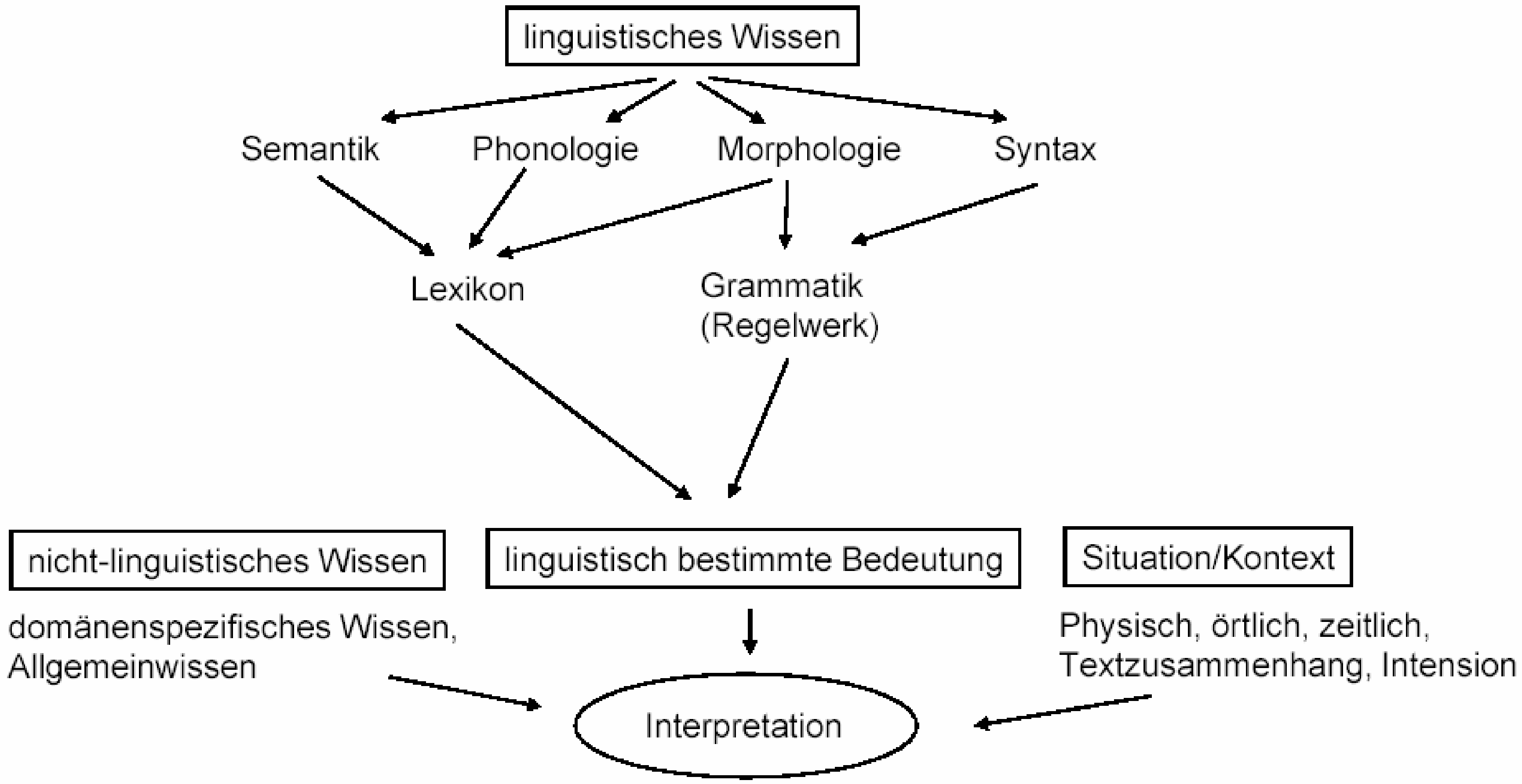
$$Begin(e,t,l) \wedge Event(e) \wedge Time(t) \wedge Location(l)$$

$$\wedge\, Equal(t,2) \wedge Room(l, V2-122, ?b)$$

$$\wedge\, Talk(e,s,l) \wedge Proffessor(s, Cambridge)$$

$$\wedge\, Name(s, Steven-Hawking) \wedge Building(b, Uni-Bielefeld) \wedge \ldots$$

Interpretation



Semantic interpretation/analysis

Ziel: Bestimmung des semantischen Potenzials

- Umformung des *Parse*-Baumes in eine *interne Repräsentation* (z.B. Prädikatenlogik, Frames, …)
- Zwei wesentliche Schritte:
  1. *Lexikalische Semantik:* Bestimmung der Bedeutung einzelner Wörter
     - Homonymie, Polysemie: bank/bank
     - Synonyme: big/large
     - Antonyme: boy/girl, hot/cold
     - Siehe z.B. *WordNet* (http://wordnet.princeton.edu/)
  2. *Satzsemantik:* Konstruktion der Gesamtbedeutung aus den Einzelbedeutungen (*kompositionelle* Semantik)

Satzsemantik

- *Kompositionalitätsproblem*: Wie kann die Satzsemantik auf Basis der syntaktischen Struktur ermittelt werden?
  - *Tom arbeitet.* arbeitet(Tom)
  - *Jemand arbeitet.* $\exists$x arbeitet(x)
  - *Jeder Student arbeitet.* $\forall$x(student(x) $\rightarrow$ arbeitet(x))
  - *Kein Student arbeitet.* $\neg\exists$x(student(x) $\wedge$ arbeitet(x))

  Alle syntaktisch gleich (NP + intransitives Verb)

- häufig folgt man direkt der syntaktischer Struktur; *syntaktisch-semantisches Parsen*
  - Zerlegung in *Parse*-Baum enthält nicht nur syntaktische (NP, Verb, Nomen, ...), sondern auch semantische Kategorien (Personenname, Aktionsbeschreibung, etc.), z.B. C-Grammatiken
  - damit schon während des Parsens *interne Repräsentation* der Bedeutung aufgebaut

Discourse interpretation

Ziel: Von Satzsemantik zu Text-/Diskurssemantik

- Nötige Wissensquellen (über ling. Wissen hinaus):
  - Domänenwissen (banking transaction)
  - Diskurswissen (satzübergreifend)
  - Weltwissen (*Common-sense knowledge*, Situationswissen)
- Beispiel:

  U: I would like to open a fixed deposit account.

  S: For what amount?

  U: Make it for 8000 Rupees.

  S: For what duration?

  U: What is the interest rate for 3 months?

  S: Six percent.

  U: Oh good then make it for that duration.

Discourse interpretation

- *Referenzauflösung*: worauf wird Bezug genommen bei…
  - ausgelassenen Wörtern (oder Phrasen)
  - Pronomen: "*John likes that blue car. He buys it*."
    - Kontext nötig, um die Referenten von "*that blue car*", "*he*", und "*it*" zu bestimmen
    - Erst dann: likes(john, car1), buys(john, car1)
    - Auch bezeichnet als: Ko-Referenz, Anapher

Pronominale Referenz

- Auflösung nach *Salienz* der potentiellen Referenten
  - Neuheitswert, Wiederholungen
  - Grammatikalische Rolle (z.B. Subj > Obj)
  - Parallelen: *„Mary went with Sue to… Sally went with her to…"*
  - Verbsemantik: *„John telephoned Bill. He lost…"* vs *„John criticized Bill. He lost…"*
- Beispiel: Diskursrepräsentationstheorie (Kamp & Reyle)
  - Beschränkungen für die Sichtbarkeit von Referenten
  - Hierarchische Diskursrepräsentation: pronominale Referenz nur auf gleicher oder höherer Ebene
- Siehe auch: Salienz-basierte Algorithmen in Jurafsky & Martin (S. 684 ff)

Weitere zentrale Fragen der Pragmatik

- *Intentionserkennung*: Was will der Sprecher?
  - "Do you have the time?" → will die Zeit wissen
  - "When is the last train to London?" → will nach London
- *Informationsstruktur*: Was ist neu?
- *Rhetorische* und *narrative Struktur*: Wie ist der Bezug zum vorher Gesagten?
- *Konversationsanalyse*: Mit welchen sprachlichen Mitteln wurde etwas erreicht?

→ Alles eine Theorie für sich.

Speech Act theory

- Austin 1962, *How to Do Things With Words*
  - some utterances don‘t just ‘say’, but ‘do’ something
  - *performative verbs:* explicate the action they take (promise, swear, …)
- Searle 1969, *Speech Acts*
  - Different forces/acts of an utterance (locution)
    - *Locutionary* act: the act of uttering something
    - *Illocutionary* force: the action performed by the utterance
    - *Perlocutionary* force: the change it is meant to bring about
  - Example: “I’m cold.”
    - Illuctionary act: *inform* the addressee that speaker is cold
    - Perlocutionary force: wish to have the window closed
  - Speech act is the expression of the *illocutionary force*
  - Basic speech acts: *request, inform, ask*

Intention recognition

- Goal:
  to recognize the intent of each user utterance as one of a (usually limited) set of speech acts based on context information
- Sample dialogue actions:
  - Switchboard DAMSL (Allen)
    - Conventional-closing
    - Statement-(non-)opinion
    - Agree/Accept
    - Acknowledgment
    - Yes-No-Question/Yes-Answer
    - Non-verbal
    - Abandoned
  - Verbmobil
    - Greet/Thank/Bye
    - Suggest
    - Accept/Reject
    - Confirm
    - Clarify-Query/Answer
    - Give-Reason
    - Deliberate
- On-going standardization efforts
  - Discourse Resource Initiative
  - Allen's Multiparty discourse group
    http://www.cs.rochester.edu/research/cisd/resources/damsl

Summary

- Von Sätzen, die John zu/über Fred sagt,...
  - "What is the time?"
  - "He has missed the train."
- ...können wir jetzt
  - Parsen → syntaktische Struktur
  - Auf strukturierte Repräsentation abbilden, die Inferenz erlaubt → linguistische bestimmte Semantik
  - Mit Wissen um Kontext und Ziele/Pläne vollständige Interpretation herleiten:
    - wants(john, know(john, time1))
    - believes(john, missed(fred, train2))

Geht's auch ohne Syntax und Semantik?

Ja, z.B. mit "*keyword-spotting*":

- durchsuchen der Benutzereingabe nach bestimmten Schlüsselworten, z.B. "Wetter"
- generieren einer Antwort, die zum Schlüsselwort paßt
- Grundlage vieler *Chatterbots*
  - Eliza (Weizenbaum, 1969), die Mutter aller chatterbots
  - ALICE (http://www.alicebot.org/)
  - Anna (www.ikea.de)
  - …
- bereits bei einfachen syntaktischen Kniffen überfordert

```
Benutzer: "Ich möchte auf keinen Fall über's Wetter reden!"
Bot: "Gern! Hier in Bielefeld regnet es mal wieder."
```

Natural Language generation

E. Reiter & R. Dale (2000) *Building Natural Language Generation Systems*. Cambridge University Press.

What is NLG?

- Goal:
  - produces understandable and appropriate output in natural language, possibly along with prosodic information
- Input:
  - some underlying non-linguistic representation of information
- Output:
  - text to speak, prosodic information
- Knowledge sources required:
  - linguistic knowledge (of language)
  - domain and world knowledge

Language Generation

- Simplest generation method is using *templates*, mapping representation straight to text template (with variables/ slots to fill in).
  - loves(X, Y) → X "loves" Y
  - gives(X, Y, Z) → X "gives the" Y "to" Z
- Templates are very rigid, much more to NLG in general..
  - Consider "John eats the cheese. John eats the apple. John sneezes. John laughs."
  - Better as "John eats the cheese and apple, then sneezes. He then laughs."
- Getting good style involves working out how to map many facts to one sentence, when to use pronouns, when to use "connectives" like "then" etc.

Language Generation

- Serious language generation involves deciding
  - what to say
  - how to order and structure it
  - how to break it up into sentences
  - how to refer to objects (using pronouns, and expressions like “the cat” etc)
  - How to express things in terms of grammatically correct sentences
- Often starting point is a *communicative goal*

Why is NLG hard?

- Linguistics does not provide us with a precise theory about how to make such choices
- The choices to be made interact with one another in complex ways
- Many results of choices (e.g. text length) are only visible at the end of the process
- There doesn’t seem to be any simple and reliable way to order the choices

Tasks in NLG

Content Planning	Content Determination Discourse planning (Document Structuring)
Micro-planning	Aggregation Lexicalisation Referring Expression Generation
Surface Realisation	Linguistic Realisation Structure Realisation

1. Content Planning

Goals:

- to determine what information to communicate (content)
- to determine how to structure this information to make a coherent text/discourse

Results: *messages*, predefined data structures that…

- correspond to informational elements (units)
- collect underlying data in ways convenient for ling. expression

- Essentially, a domain-dependent expert-systems-like task
- Common approaches:
  1. based on observations about common utterance structures
  2. based on reasoning about discourse coherence and the purpose of the utterance

1b. Discourse planning via *schemas*

Basic idea (McKeown 1985):

- texts often follow conventionalized patterns
- can be captured by means of ‘text grammars’ that dictate content and structure
- specify how a content plan can be constructed using smaller schemas or atomic messages
  - can specify degrees of variability and optionality
- currently the most popular content planning approach in applied NLG systems

1b. Discourse planning via reasoning

- Observation:
  - discourse coherent by virtue of relationships between their parts — relationships like *narrative sequence, elaboration, justification* ...

- Approach:
  - segment knowledge of what makes a discourse coherent into separate rules
  - use these rules to dynamically compose utterance from constituent elements by reasoning about the role of these elements in the overall discourse

1b. Discourse planning via explicit reasoning

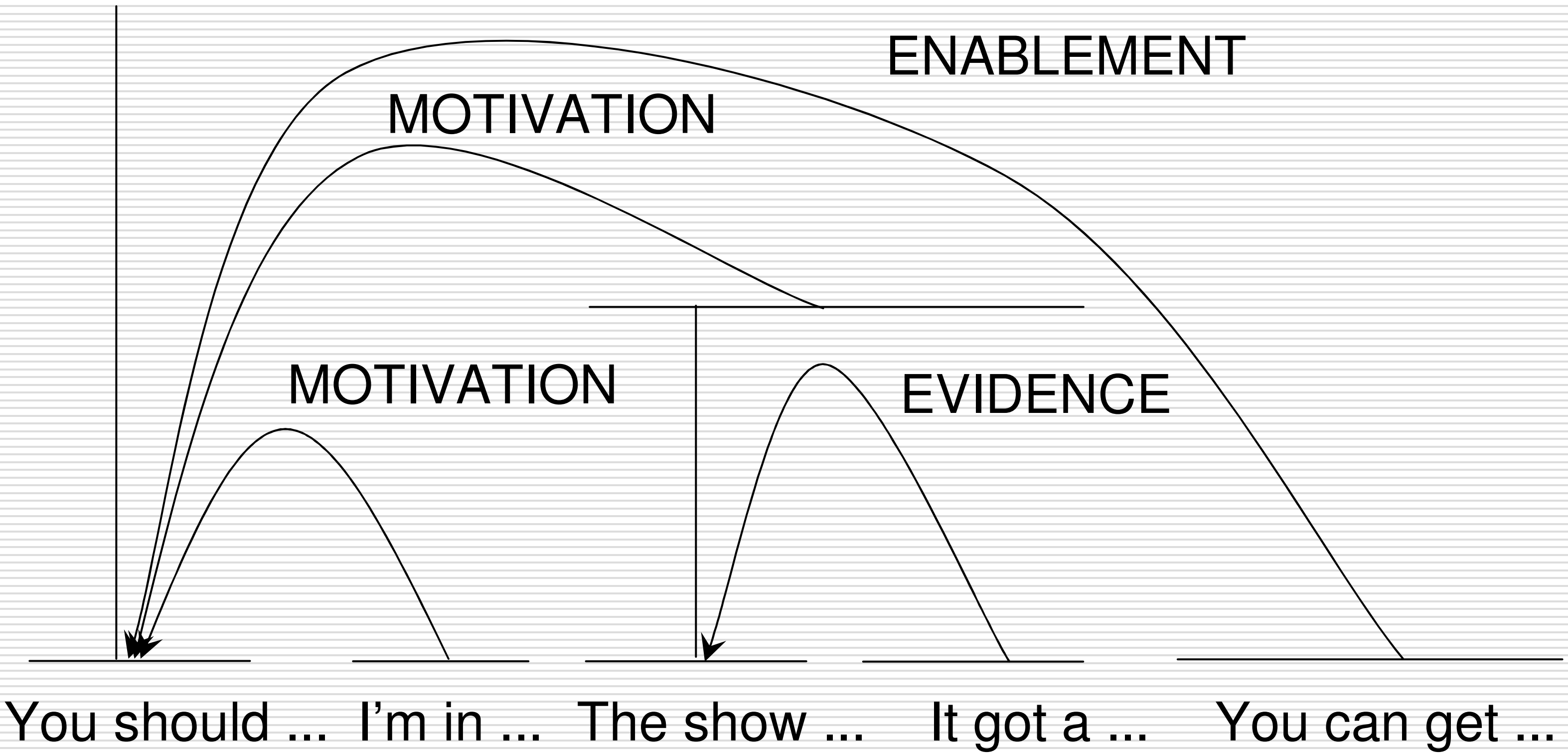
- Adopt AI planning techniques:
  - Goal = desired communicative effect(s)
  - Plan constituents = messages or structures that combine messages (subplans)
- Can involve reasoning about addressee's beliefs, e.g. using modal logics

- Often based on *Rhetorical Structure Theory* (RST) (Mann et al., 1983; http://www.sfu.ca/rst/)
  - every part of a coherent text has some function, some relationship to another part (e.g. *elaborate* a previous part)
  - distinction between *nucleus*, the central segment, and the *satellite*, the less central and more peripheral one
  - many relations asymmetric, defined in terms of constraints on nucleus, satellite, and their combination

See also: Seminal work on discourse structure by Webber (78/83); Grosz (77); Sidner (79/83)

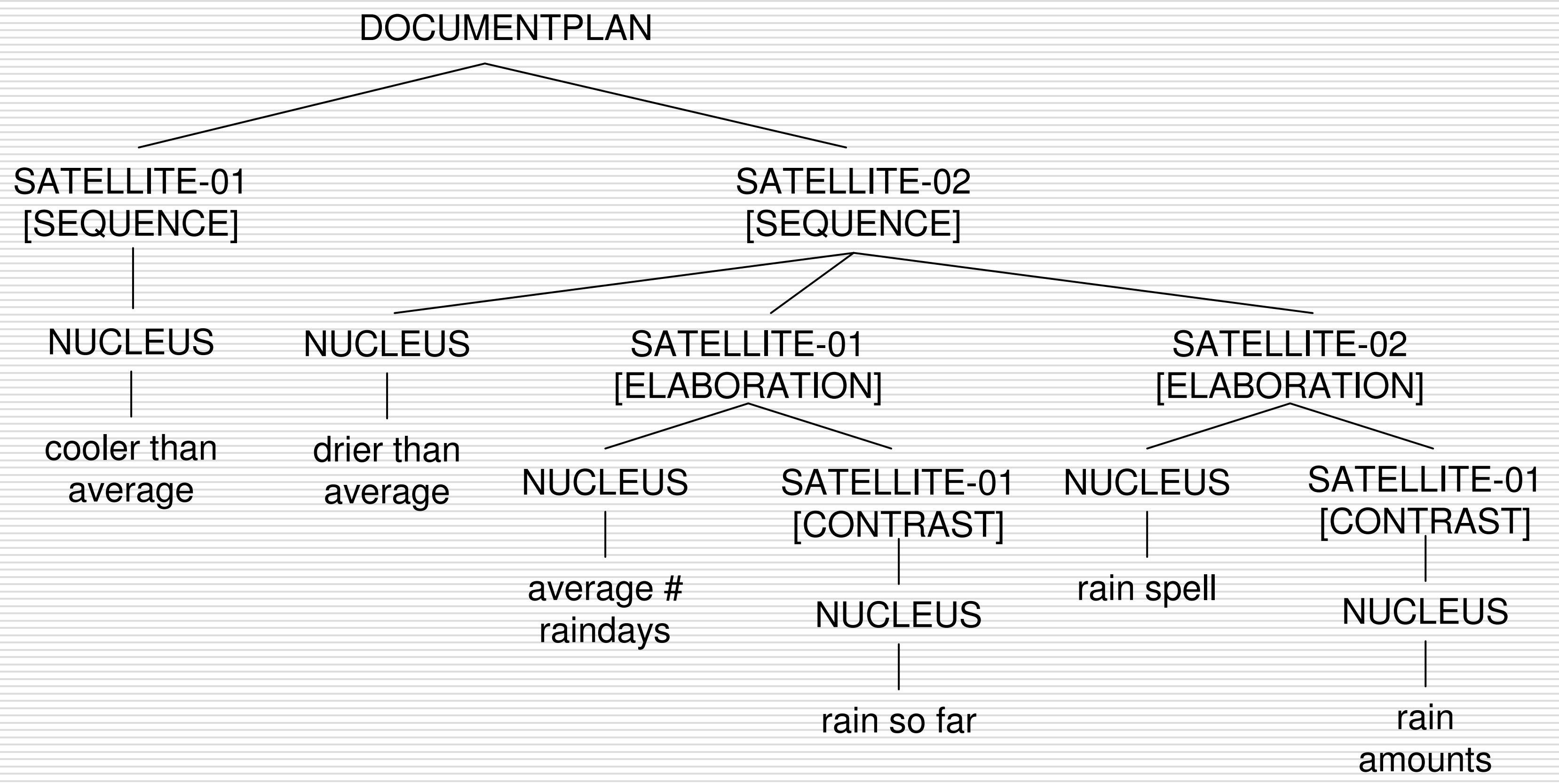
Rhetorical relations from RST

Elaboration, Contrast, Condition, Purpose, Sequence, Result, …
(inventory of 23 relations)



Content plan (aka. document plan)

- Tree structure with messages at its leaf nodes
- Example from *WeatherReporter* system (Reiter et al.):



2. Microplanning

A lot of (older) systems leave this out, but then *generation gap:* discrepancy between the actual content plan and the expected input of the surface realization process

Goal:

- Close that gap! - convert a content plan into a sequence of sentence or phrase specifications

Tasks:

- Aggregation
- Lexicalisation
- Reference

2a. Aggregation

Example – noun phrase design

- A noun phrase can convey an arbitrary amount of information:
  - *someone* vs
  - a *designer* vs
  - an *old designer* vs
  - an *old designer with red hair* vs
  - …

Example – clausal structure construction

- Multiple messages can be conveyed by one sentence
  - *Heavy rain fell on the 27th. Heavy rain fell on the 28th.* vs
  - *Heavy rain fell on the 27th and 28th*. vs
  - …

2a. Aggregation - some possibilities

- aggregation via simple *conjunction*:
  - `Heavy rain fell on the 27th and heavy rain fell on the 28th.`
- aggregation via *ellipsis*:
  - `Heavy rain fell on the 27th and [] on the 28th.`
- aggregation via *set introduction*:
  - `Heavy rain fell on [the 27th and 28th].`

Without aggregation:

- March had a rainfall of 120mm. It was the wettest month.

- aggregation via *embedding*:
  - March, which was the wettest month, had a rainfall of 120mm.

2a. Aggregation - some heuristics

There are usually many ways to aggregate a given message set: how do we choose?

- conform to genre conventions and rules
- observe structural properties
  - for example, only aggregate messages which are siblings in the document plan tree
- take account of pragmatic goals
  - be emphatic, make the text friendlier
  - make the text easier for poor readers
  - …

2b. Lexicalisation

- The process of choosing words to communicate the information in messages
- If several lexicalisations are possible, consider:
  - user knowledge and preferences
  - consistency with previous usage
    - in some cases, it may be best to vary lexemes
  - interaction with other aspects of microplanning
  - pragmatics
    - `It is encouraging that you have many reasons to stop.` (more precise meaning)
    - `It's good that you have a lot of reasons to stop.` (lower reading level)

2c. Referring expression generation

How to refer to specific domain objects and entities?

Two issues:

1. *Initial* introduction of an object into the discourse
   - use a full name: *Jeremy*
   - relate to an object that is already salient: *Jane's goldfish*
   - specify location: *the goldfish in the bowl on the table*

   (not really well understood)

2. *Subsequent* references to an already known and salient object
   - Pronouns: *It swims in circles.*
   - Definite NPs: *The goldfish swims in circles.*
   - Proper names, possibly abbreviated: *Jeremy swims in circles.*

2c. Choosing a form of reference

Some suggestions from the literature:

- use a pronoun if it refers to an entity mentioned in the previous clause, and there is no other entity in the previous clause that the pronoun could refer to
- otherwise use a name, if a short one exists
- otherwise use a definite NP

Also important to conform to genre conventions – for example, there are more pronouns in newspaper articles than in technical manuals

3. Surface realisation

Goal:

to convert text specifications into actual text

Purpose:

to hide the peculiarities of English (or whatever the target language is) from the rest of the NLG system

Tasks:

- *Structure realisation*
  - Choose markup to convey document structure
- *Linguistic realisation*
  - Insert function words
  - Choose correct inflection of content words
  - Order words within a sentence
  - Apply orthographic rules

3b. Linguistic realisation - techniques

Use handcrafted templates

- Ex: "shallow" generation with *TG/2* (DFKI) (Busemann, 1996,98)
  - Canned text, templates and context free rules
  - All expressed as production rules whose actions are determined by conditions on input structure (written in TGL)
  - Input structures specified in the *Generation Interface Language* (GIL)
  - Three-step processing cycle as in AI production systems on the available TGL rules:
    - identify all applicable rules,
    - select an applicable rule (e.g. according to preferences),
    - fire that rule
  - Output can easily include formatting elements

3b. Linguistic Realisation - techniques

Utilize *grammars* tuned for generation

- Provides a set of choices for realisation, made based on input text spec
- Grammar can *only* be used for NLG
- Important approaches
  - Systemic grammar
  - Functional unification grammar
  - Tree-adjoining grammar

Speech Output

- the NLG Perspective: enhances output possibilities
  - communicate via spoken channels (e.g. telephone)
  - add explicit information (e.g. emotion, importance)
- the speech synthesis perspective: intonation carries information
  - need information about syntactic structure, information status, homographs, focus, contrast etc.
  - currently obtained by text analysis
  - can be obtained from an NLG system automatically, the idea of *concept-to-speech*