

Shallow Text Generation

Stephan Busemann
DFKI GmbH
Stuhlsatzenhausweg 3
D-66123 Saarbrücken

`Stephan.Busemann@dfki.de`
`http://www.dfki.de/~busemann`



Application Systems for NLG Must be Developed Quickly and in a User-Oriented Way

- **Requirements placed by the application**
 - on the user: recognize and articulate needs
 - on the developer: make herself acquainted with the domain
 - on both: create and adapt a corpus of sample target texts
- **Requirements wrt the software**
 - Adaptability to new tasks and domains
 - Scalability (low costs of the next rule)
 - Modularisation (interpreter, daten, knowledge, interfaces)

High efficiency of development is difficult to achieve with traditional approaches to language generation



Non-Trivial Generation Systems are Expensive to Adapt to New Domains and Tasks

- **Examples**
 - KPML (Bateman 1997), systemic grammars, development environment
 - FUF/Surge (Elhadad/Robin 1992), functional unification grammar, interpreter
- **Features**
 - large multi-lingual systems
 - detailed, monolingual semantic representations as input
 - broad coverage of linguistic phenomena (goal: the more, the better)
- **Effort for adaptation**
 - Rich interface to the input language of the system (logical form, SPL)
 - Generation of sentences reflecting the distinctions covered

The excellent scope of services of generic resources can often not be utilised in practice

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 3

Deep* vs Shallow NLG

* This differs from the Chomskyan distinction between deep and surface structure, which is sometimes used to characterize deep and surface generation

- **Deep generation**
 - knowledge-based (models of the domain, of the author and the addressees, of the language(s) involved)
 - theoretically motivated, aiming at generic, re-usable technology
 - unresolved issue of general system architecture
- **Shallow generation**
 - opportunistic modelling of relevant aspects of the application
 - diverse depth of modelling, as required by the application
 - some methods viewed as „short cuts“ for unsolved questions of deep generation

Shallow generation can be defined in analogy to shallow analysis

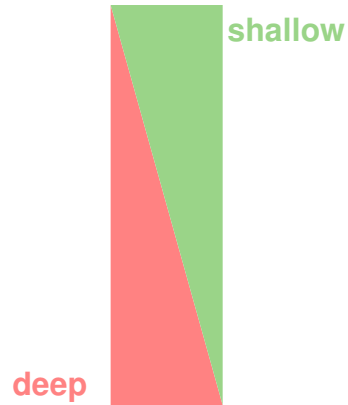
Source: Stephan Busemann



Language Technology I, WS 2012/2013, 4

There is a Smooth Transition Between Shallow and Deep Methods

- Prefabricated texts
- „Fill in the slots“
- with flexible templates
- with aggregation
- with sentence planning
- with document planning



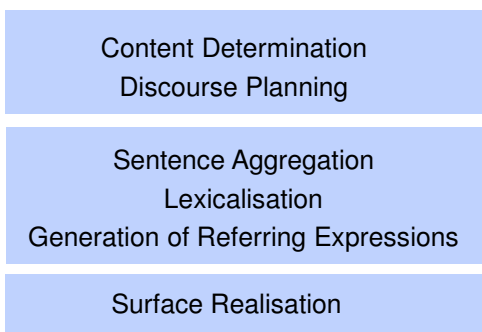
Source: Stephan Busemann



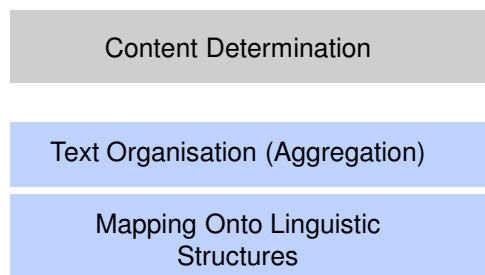
Language Technology I, WS 2012/2013, 5

Shallow Architectures Have a Simple Task Structure

“Deep” model with interaction
(cf. Reiter/Dale 2000)



„Shallow“ Model
(Busemann/Horacek 1998)



Source: Stephan Busemann



Language Technology I, WS 2012/2013, 6

Overview

- Motivation
- **The TG/2 Shallow NLG framework**
- Some major applications for shallow NLG
- Assessment and conclusions

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 7

Input for Air Quality Report Generation

```
[(COOP threshold-passing)
 (TIME [(PRED season)
        (NAME [(SEASON summer)
                (YEAR 1999)])])]
(POLLUTANT o3)
(SITE "Völklingen-City")
(DURATION [(MINUTE 60)])
(SOURCE [(LAW-NAME bimsch)
          (THRESHOLD-TYPE info-value)])
(EXCEEDS [(STATUS yes)
          (TIMES 1)])]
```

In summer 1999 at the measuring station of Völklingen-City, the information value for ozone – 180 µg/m³ according to the German decree Bundesimmissionschutzverordnung – was exceeded once during a period of 60 minutes.

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 8

Input for Air Quality Report Generation

```
[ (COOP threshold-passing)
  (TIME [ (PRED season)
          (NAME [ (SEASON summer)
                  (YEAR 1999) ] ] )
  (POLLUTANT o3)
  (SITE "Völklingen-City")
  (DURATION [ (MINUTE 60) ] )
  (SOURCE [ (LAW-NAME bimsch)
            (THRESHOLD-TYPE info-value) ] )
  (EXCEEDS [ (STATUS yes)
             (TIMES 1) ] ] ]
```

Im Sommer 1999 wurde der Informationswert für Ozon an der Messstation Völklingen-City während einer 60-minütigen Einwirkungsdauer ($180 \mu\text{g}/\text{m}^3$ nach Bundesimmissionsschutzverordnung) einmal überschritten.

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 9

Input for Air Quality Report Generation

```
[ (COOP threshold-passing)
  (TIME [ (PRED season)
          (NAME [ (SEASON summer)
                  (YEAR 1999) ] ] )
  (POLLUTANT o3)
  (SITE "Völklingen-City")
  (DURATION [ (MINUTE 60) ] )
  (SOURCE [ (LAW-NAME bimsch)
            (THRESHOLD-TYPE info-value) ] )
  (EXCEEDS [ (STATUS yes)
             (TIMES 1) ] ] ]
```

En été 1999, à la station de mesure de Völklingen-City, la valeur d'information pour l'ozone pour une exposition de 60 minutes ($180 \mu\text{g}/\text{m}^3$ selon le décret allemand (Bundesimmissionsschutzverordnung)) a été dépassée une fois.

Source: Stephan Busemann

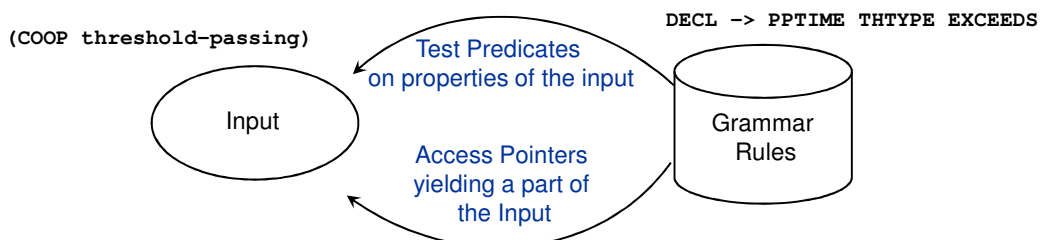


Language Technology I, WS 2012/2013, 10

TG/2 Offers a Flexible Framework for NLG

- TG/2 is a transparent production system
- TG/2 interprets a separately defined set of condition-action rules
- TG/2 maps pieces of input onto surface strings

TG/2 keeps grammars largely independent from input representations



Source: Stephan Busemann



Language Technology I, WS 2012/2013, 11

TG/2 Grammars Integrate Canned Texts, Templates and Context-free Rules

```

My category is DECL.
IF the slot COOP is 'threshold-passing
  AND the slot LAW-NAME is specified
THEN apply Pptime from slot TIME
  apply THTYPE from CURRENT-INPUT
  utter "("
  apply LAW from slot LAW-NAME
  utter ") "
  apply EXCEEDS from slot EXCEEDS
  utter "."
WHERE THTYPE AND EXCEEDS agree in GENDER
    
```

(Busemann 1996)

*En été 1999
la valeur limite autorisée
(
selon le decret ...
)
a été dépassée une fois
.*

```

My category is THTYPE.
IF there is no slot THRESHOLD-TYPE specified
THEN utter "la valeur limite autoris&e2e "
WHERE THTYPE has value 'fem for GENDER
    
```

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 12

TG/2 Grammars Integrate Canned Texts, Templates and Context-free Rules

```

My category is DECL.
IF the slot COOP is 'threshold-passing
  AND the slot LAW-NAME is specified
THEN apply Pptime from slot TIME
  apply THTYPE from CURRENT-INPUT
  utter "("
  apply LAW from slot LAW-NAME
  utter ")"
  apply EXCEEDS from slot EXCEEDS
  utter "."
WHERE THTYPE AND EXCEEDS agree in GENDER
    
```

(Busemann 1996)

*En été 1999
la valeur limite autorisée
(
selon le decret ...
)
a été dépassée une fois
.*

```

My category is THTYPE.
IF there is no slot THRESHOLD-TYPE specified
THEN utter "la valeur limite autoris&e2e "
WHERE THTYPE has value 'fem for GENDER
    
```

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 13

TG/2 Grammars Integrate Canned Texts, Templates and Context-free Rules

```

My category is DECL.
IF the slot COOP is 'threshold-passing
  AND the slot LAW-NAME is specified
THEN apply Pptime from slot TIME
  apply THTYPE from CURRENT-INPUT
  utter "("
  apply LAW from slot LAW-NAME
  utter ")"
  apply EXCEEDS from slot EXCEEDS
  utter "."
WHERE THTYPE AND EXCEEDS agree in GENDER
    
```

(Busemann 1996)

*En été 1999
la valeur limite autorisée
(
selon le decret ...
)
a été dépassée une fois
.*

```

My category is THTYPE.
IF there is no slot THRESHOLD-TYPE specified
THEN utter "la valeur limite autoris&e2e "
WHERE THTYPE has value 'fem for GENDER
    
```

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 14

TG/2 Grammars Integrate Canned Texts, Templates and Context-free Rules

```

My category is DECL.
IF the slot COOP is 'threshold-passing
  AND the slot LAW-NAME is specified
THEN apply Pptime from slot TIME
  apply THTYPE from CURRENT-INPUT
  utter "("
  apply LAW from slot LAW-NAME
  utter ")"
  apply EXCEEDS from slot EXCEEDS
  utter "."
WHERE THTYPE AND EXCEEDS agree in GENDER
    
```

(Busemann 1996)

*En été 1999
la valeur limite autorisée
(
selon le decret ...
)
a été dépassée une fois
.*

```

My category is THTYPE.
IF there is no slot THRESHOLD-TYPE specified
THEN utter "la valeur limite autoris&e2e "
WHERE THTYPE has value 'fem for GENDER
    
```

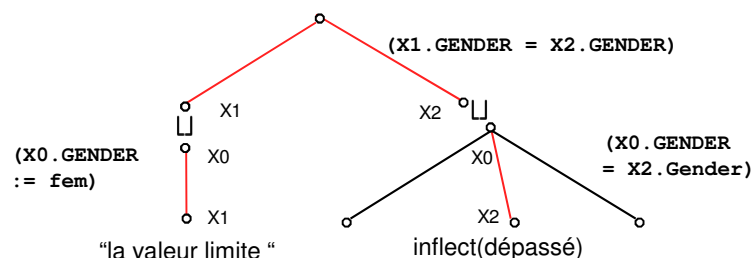
Source: Stephan Busemann



Language Technology I, WS 2012/2013, 15

Constraints are Percolated Across the Derivation Tree

- Feature unification (\sqcup) at tree nodes
- Every tree of depth 1 is licensed by a grammar rule
- A feature can be assigned a value ($:=$)
- Two features can be constrained to have identical values ($=$)



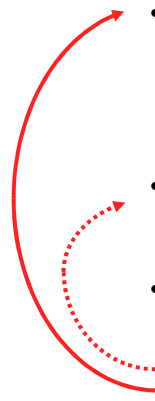
Source: Stephan Busemann



Language Technology I, WS 2012/2013, 16

The Interpreter is Based on the Context-Free Backbone of the Grammars

THREE-STEP EVALUATION CYCLE

- 
- **Matching**
 - Identify all rules with the **current category**
 - For each of them perform its tests on the **input structure** (“IF” part)
 - Add those passing the tests to the conflict set
 - **Conflict resolution**
 - Select an element of the conflict set (possibly by some preference mechanism)
 - **Firing**
 - Evaluate the rule’s constraints (if available, “WHERE” part)
 - For each element of the “THEN” part, read the new category and determine the new input structure by evaluating the associated access pointer

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 17

Overview

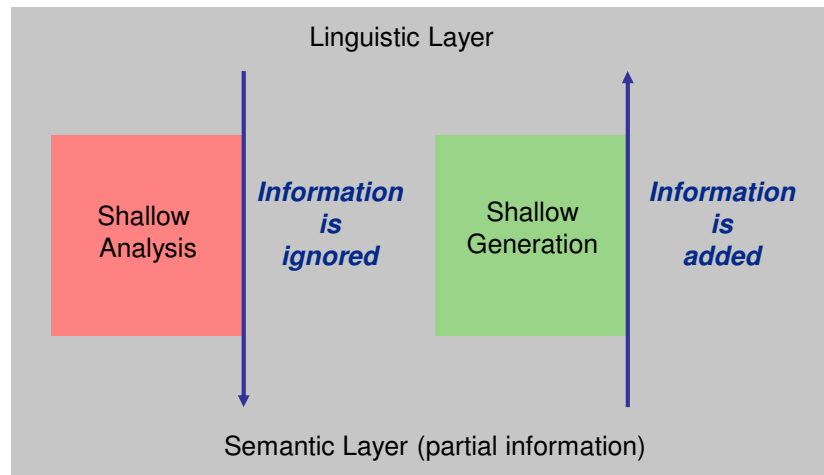
- **Motivation**
- **The TG/2 Shallow NLG framework**
- **Some major applications for shallow NLG**
- **Assessment and conclusions**

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 18

Shallow Processing Deals With Partial Information



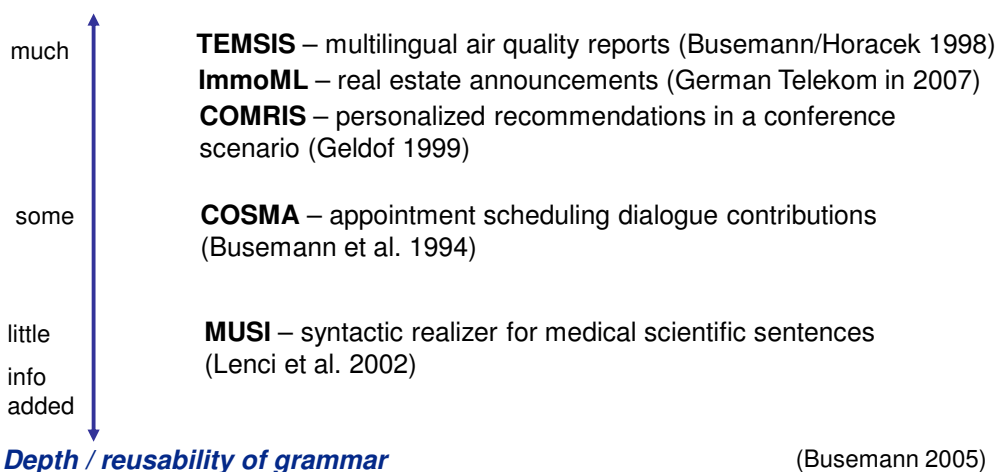
Source: Stephan Busemann



Language Technology I, WS2012/2013, 19

Some Major Applications with TG/2

Shallowness / Domain dependence of grammar



Source: Stephan Busemann

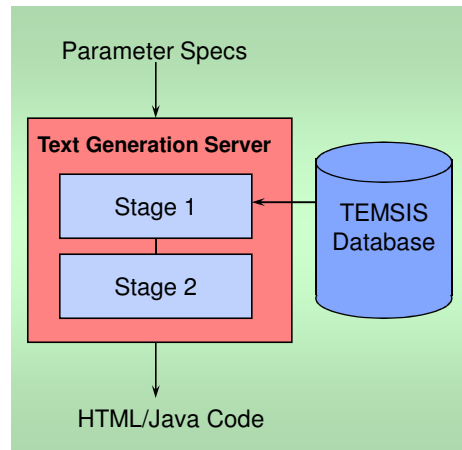


Language Technology I, WS2012/2013, 20

Text Generation in TEMSIS Occurs in Two Steps

GENERATION SYSTEM OVERVIEW

- **Parameter selection by the user**
 - language (D, E, F, P, C, J)
 - pollutant and measurement station
 - relevant period of time
- **Stage 1: Text schema construction**
 - querying the database
 - composition of report structure
 - elision of contextual redundancies
- **Stage 2: Linguistic realisation by TG/2**
 - selection of sentence patterns
 - wording, phrasing, grammar
- **HTML postprocessing**



Source: Stephan Busemann



Language Technology I, WS 2012/2013, 21

The Texts Vary According to the User's Preferences

EXAMPLE

- **Parameters selected within the TEMSIS Navigator menus:**
 - French text about a German situation
 - ozone data, exceeding thresholds according to decree
 - measurements at Völklingen-City in summer 1997 (to be confirmed)

Vous avez choisi la station de mesure de Völklingen-City afin de consulter la pollution atmosphérique relevée en été 1997.

A la station de mesure de Völklingen-City, la valeur d'information pour l'ozone pour une exposition de 60 minutes ($180 \mu\text{g}/\text{m}^3$ selon le décret allemand (Bundesimmissionsschutzverordnung)) a été dépassée une fois.

La valeur d'interdiction du trafic ($240 \mu\text{g}/\text{m}^3$) a aussi été dépassée une fois.

En été 1996 la valeur d'information ($180 \mu\text{g}/\text{m}^3$) n'a pas été dépassée .

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 22

The Reports Consist of Several Statements

SAMPLE SCHEMA FOR SUMMER OBSERVATION, THRESHOLD PASSING

- **Confirm pollutant, measurement station, and time interval**
- **Number the values exceeding the lowest threshold**
- **Number the values exceeding the next threshold**
- **Compare with values of preceeding year**
- **Repeat the core statement („Summary“)**

A schema is instantiated on the basis of the input parameters and the retrieved data

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 23

Instantiating a Schema Leads to a Report Structure

TEXT ORGANISATION

- **Achieves text coherence by**
 - removing redundant information
 - inserting particles („also“)
 - simple techniques of aggregating information
- **Yields canned texts or intermediate content representations**
- **Intermediate representations are independent of particular languages**
 - TG/2 generates German, French, English, Portuguese, Chinese and Japanese text from them

Shallow generation can do without explicit knowledge representation and text planning

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 24

Non-Linguistic Input for Air Quality Report Generation in TEMSIS

```
[ (COOP threshold-passing)
  (TIME [ (PRED season)
          (NAME [ (SEASON summer)
                  (YEAR 1999) ] ] )
  (POLLUTANT o3)
  (SITE "Völklingen-City")
  (DURATION [ (MINUTE 60) ] )
  (SOURCE [ (LAW-NAME bimsch)
            (THRESHOLD-TYPE info-value) ] )
  (EXCEEDS [ (STATUS yes)
             (TIMES 1) ] ] ]
```

In summer 1999 at the measuring station of Völklingen-City, the information value for ozone – 180 µg/m³ according to the German decree Bundesimmissions-schutzverordnung – was exceeded once during a period of 60 minutes.

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 25

Multilingual Generation in TEMSIS

- **Grammar size about 100-120 rules**
- **Written with standard text editors (emacs)**
- **Six languages: German, French, English, Chinese, Japanese, Portuguese**
- **Grammar is the only language-specific part** (except for canned texts about pollutants etc. and error messages)
- **Adding a new language required little effort: 2-4 weeks, depending on skills** (incl. getting familiar with the system)

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 26

Generated Texts Are Not Invented

CORPUS-BASED GRAMMAR DEVELOPMENT (Ehud Reiter)

- **User provide examples for target texts - the more, the better**
 - Texts produced manually by domain experts
- **Initial analysis of user-generated corpus**
 - Identify the knowledge used by the authors
 - Clarify with users any underlying semantic and rhetoric relationships
 - Discuss with users how the texts can be improved
- **Analysis of the revised corpus**
 - Definition of linguistic coverage
 - Correlate surface chains and underlying relations
 - Test of revised corpus (Wizard of Oz) and iterate the whole process, if necessary
- **Generalisation from Corpus Samples to Prototypical Examples (Templates)**
 - Basis for shallow grammar development

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 27

Input and Grammars

- **Most NLG systems expect a specific, fixed kind of input**
 - Linguistic vs non-linguistic
 - Course-grained vs fine-grained semantic specifications
 - The TG/2 framework does not constrain input
- **TG/2 grammars usually are domain-dependent**
 - The TEMSIS input needs domain-dependent interpretation
 - Grammar development in TEMSIS was cheap (~120 rules, ~20 lexemes)
- **Deeper applications require a more generic approach**
 - MUSI input is language-specific
 - MUSI grammar to cover German sentence structures (DUDEN grammar)
 - Coverage requirements are considerably higher than with shallow NLG (>800 rules, several thousand lexemes)
 - TG/2 grammar editor eGram to improve maintainability

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 28

MUSI: Language-Specific Input to TG/2

```
[ (SENTENCE DECL)
  (VC [(VOICE PASSIV) (MOOD IND) (TENSE PRAESENS) (SBP S2)
        (STEM "verursach")])
  (DEEP-SUBJ [(TY GENERIC-NP)
              (NUMBER SG) (DET INDEF) (NR V2) (GENDER MAS)
              (STEM "antagonismus")
              (PP-ATR [(PREP MIT) (DET WITHOUT) (NUMBER SG) (GENDER NTR)
                       (STEM "Acetylcholin")
                       (LOCATIVE ...)])
              (ADJ [(STEM "kompetitiv")
                    (POS ADJECTIVE) (DEG POS)]])]
  (DEEP-AKK-OBJ [(TY GENERIC-NP)
                 (NUMBER PLUR) (DET DEF) (GENDER FEM)
                 (STEM "wirkung")])]
```

*Die Wirkungen werden durch einen kompetitiven Antagonismus
mit Acetylcholin ... verursacht.*

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 29

MUSI: Language-Specific Input to TG/2

```
[ (SENTENCE DECL)
  (VC [(VOICE PASSIV) (MOOD IND) (TENSE PRAESENS) (SBP S2)
        (STEM "verursach")])
  (DEEP-SUBJ [(TY GENERIC-NP)
              (NUMBER SG) (DET INDEF) (NR V2) (GENDER MAS)
              (STEM "antagonismus")
              (PP-ATR [(PREP MIT) (DET WITHOUT) (NUMBER SG) (GENDER NTR)
                       (STEM "Acetylcholin")
                       (LOCATIVE ...)])
              (ADJ [(STEM "kompetitiv")
                    (POS ADJECTIVE) (DEG POS)]])]
  (DEEP-AKK-OBJ [(TY GENERIC-NP)
                 (NUMBER PLUR) (DET DEF) (GENDER FEM)
                 (STEM "wirkung")])]
```

*Die Wirkungen werden durch einen kompetitiven Antagonismus
mit Acetylcholin ... verursacht.*

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 30

Realization of German Sentences in MUSI

- **Size of hand-written grammar: about 950 rules**
- **Written with standard text editors (emacs), then dedicated editor eGram**
- **CFGs do not support encoding of word order variation etc.**
- **Metarule formalism within eGram (Rinck 2003)**
- **Size of derived grammar about 2.500 rules**
- **Processing slows down with huge conflict sets**
- **Take decisions on sentence structure and lexical choice outside of TG/2**

Performance loss on backtracking is low

Size of grammars and conflict sets matter

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 31

Overview

- **Motivation**
- **The TG/2 Shallow NLG framework**
- **Some major applications for shallow NLG**
- **Assessment and conclusions**

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 32

TG/2 is a Single Pass Mapper

- **TG/2 is often combined with other NLG components**
 - TEMSIS: Text structuring depending on database content; TG/2 generating at paragraph level
 - MUSI: Lexicalization and syntactic choice, avoiding huge conflict sets in TG/2; TG/2 as a sentence realizer
- **For interdependencies between subtasks, as in sentence planning, the rule set must spell out all alternatives and quickly becomes unwieldy**

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 33

Shallow Generation Has Pros and Cons

ASSESSMENT

Possible advantages Possible drawbacks

- Low development effort
- Reusable interpreter and subgrammars
- Very fast processing
- Easy introduction of additional languages
- Easy extension with alternative formulations (through a preference mechanism in TG/2)

- Knowledge representation depends on application
- Implicit dependencies
- Scalability is inherently lower than with deep generators
- Maintaining transparency of grammars can become a cost factor

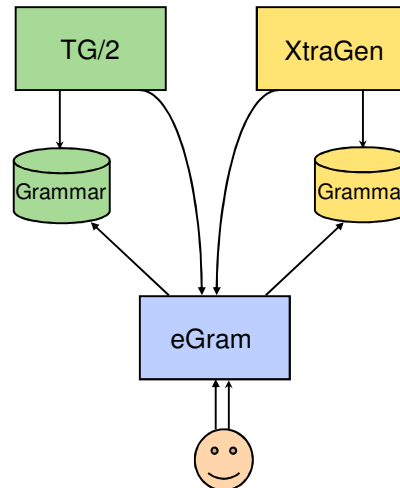
Source: Stephan Busemann



Language Technology I, WS 2012/2013, 34

Conclusions

- TG/2 is a framework that can implement shallow NLG tasks as well as deep realization
- Grammar writing for TG/2 and XtraGen is supported by eGram
- TG/2 has been licensed to more than 30 sites for commercial, educational and research purposes



Source: Stephan Busemann



Language Technology I, WS 2012/2013, 35

References

Note: a rather comprehensive bibliography of the field is available from <http://www.fb10.uni-bremen.de/anglistik/langpro/bibliographies/>

- **Bateman 1997** John Bateman: Enabling technology for multilingual natural language generation: the KPML development environment. In *Natural Language Engineering* 3(1), pp. 15-55.
- **Busemann et al. 1994** Stephan Busemann, Stephan Oepen, Elizabeth Hinkelman, Günter Neumann, Hans Uszkoreit: *COSMA - Multi-Participant NL Interaction for Appointment Scheduling*. Research Report RR-94-34, DFKI, Saarbrücken
- **Busemann 1996** Stephan Busemann: Best-First Surface Realization. In Donia Scott (ed.): *Proc. 8th International Natural Language Generation Workshop (INLG'96)*, pp. 101-110, Sussex, UK
- **Busemann 2005** Stephan Busemann: Ten Years After: An Update on TG/2 (and Friends). In *Proceedings 10th European Workshop on Natural Language Generation*, Aberdeen.
- **Busemann/Horacek 1998** Stephan Busemann and Helmut Horacek: A Flexible Shallow Approach to Text Generation. In: Eduard Hovy (ed.): *9th International Natural Language Generation Workshop*, pp. 238-247, Niagara-on-the-Lake, Canada
- **Elhadad/Robin 1992** Michael Elhadad and Jaques Robin: Controlling content realization with functional unification grammars. In R. Dale et al. (eds.): *6th International workshop on natural language generation*. LNAI 587. Berlin, Heidelberg, New York: Springer, pp. 89-104.
- **Geldof 1999** Sabine Geldof: Templates for wearables in context. In T. Becker and S. Busemann (eds.): *"May I Speak Freely?" Between Templates and Free Choice in Natural Language Generation*. Workshop at the 23rd German Annual Conference for Artificial Intelligence (KI'99), Bonn, DFKI Document D-99-01, DFKI, Saarbrücken
- **Lenci et al. 2002** Alessandro Lenci, Roberto Bartolini, Nicoletta Calzolari, Ana Agua, Stephan Busemann, Emmanuel Cartier, Karine Chevreau, José Coch: Multilingual Summarization by Integrating Linguistic Resources in the MLIS-MUSI Project. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC'02)*, May 29-31, Las Palmas, Canary Islands, Spain
- **Reiter/Dale 2000** Ehud Reiter and Robert Dale: *Building Natural Language Generation Systems*. Cambridge University Press
- **Rinck 2003** Matthias Rinck: *Ein Metaregelformalismus für TG/2*. Master's thesis, Diplomarbeit, Computerlinguistik, Universität des Saarlandes, Saarbrücken, Germany

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 36

Questions Answered by Slideset

- **How does shallow generation differ in principle from (standard) deep generation?**
- **Give advantages and disadvantages of shallow generation.**
- **How are sample corpora used to ensure the required coverage is available and the correct wordings are generated?**
- **Should an NLG problem be addressed by using as a resource clause-length pieces of prefabricated text with gaps to be filled during generation? Justify your decision considering both the complexity of the problem and the complexity of the generation process.**

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 37

Design an NLG System Based on TG/2

- **Domain: stock values in comparison to previous day**
- **Language: English**
- **Data: 100 stock values on two subsequent days**
- **Output: tendency, changes not covered by the tendency**
- **To do**
 - What concrete kinds of texts will you generate? – Choose the grainsize, with which you will inform the user.
 - What useful pre-calculations would you make on the data?
 - How would you represent the TG/2 input?
 - How would you define TG/2 rules to generate the output?
- **As always in NLG there is no single or best solution. Try to cover all imaginable developments of the data.**

Source: Stephan Busemann



Language Technology I, WS 2012/2013, 38