“Burglars were broken into our house.” - English Passive Constructions in the Written Language of German Learners in Baden-Württemberg

Verena Möller

Institut für Informationswissenschaft und Sprachtechnologie
Universität Hildesheim

Centre for English Corpus Linguistics
Université catholique de Louvain

verena.moeller@uni-hildesheim.de

2nd Tübingen-Berlin Meeting on Analyzing Learner Language

Overview

1 Introduction

2 Input and Norm: The Teaching Materials Corpus

3 The Learner Corpus: Argumentative Essays and Experimental Task

4 The Learner Corpus: Linguistic Annotation

5 The Learner Corpus: Metadata

6 The Pilot Study: Passive Constructions in Learner Text

7 The Pilot Study: Passive Constructions in the Experimental Task

Introduction

The Passive and the German Learner - What Does the Curriculum Say?

Year 8:
Die Schülerinnen und Schüler können ...
... Geschehen aus der Sicht des Verursachers und des Objekts darstellen
(active/passive voice, verbs with two objects, verbs with prepositions, by-agent)

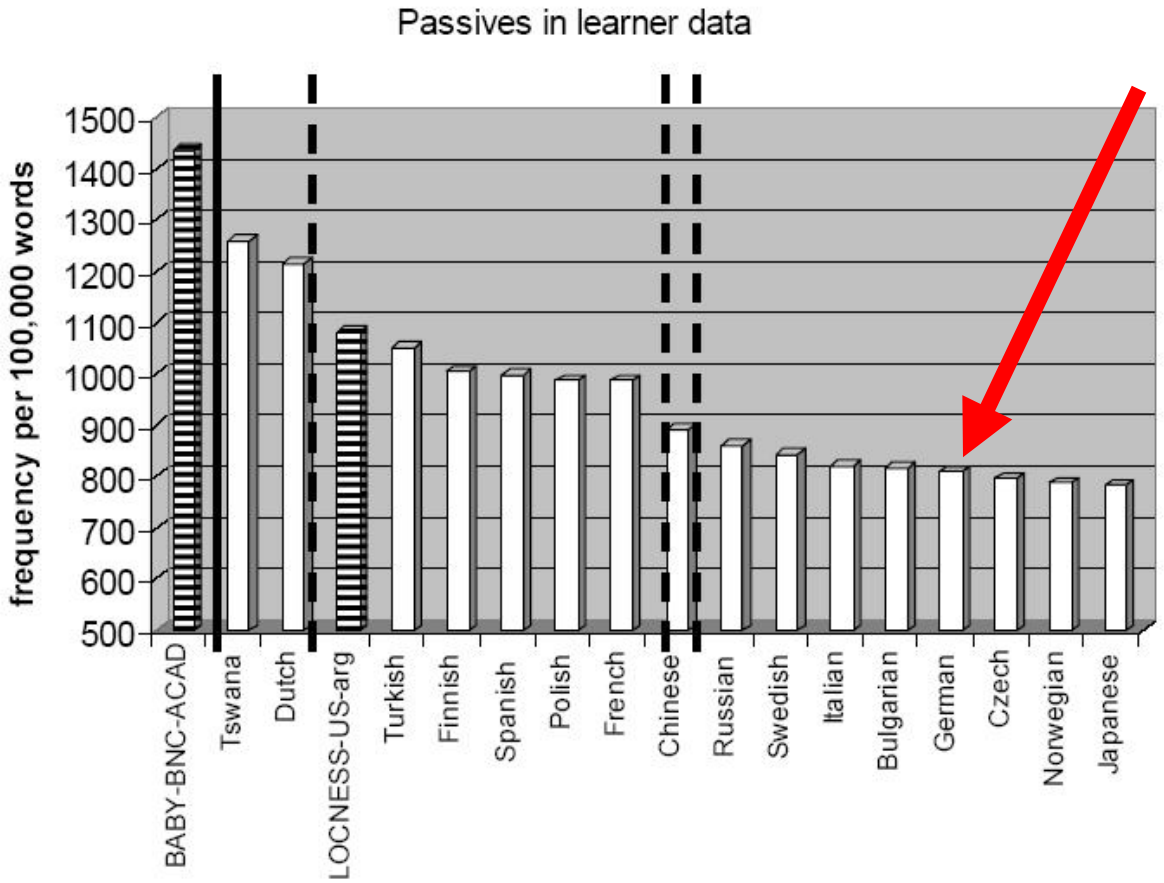
Year 10:
Die Schülerinnen und Schüler können ...
... Dauer/Wiederholung von Sachverhalten und Handlungen ausdrücken
(progressive forms: passive, [...])

Year 11/12:
Die Schülerinnen und Schüler können ...
... sich vorwiegend sicher häufig verwendeter, auch komplexerer syntaktischer Strukturen bedienen, die auch besonders im schriftsprachlichen Englisch verwendet werden;
... Unterschiede zwischen Registern erkennen und diese angemessen verwenden.

KMBW (Ministerium für Kultus, Jugend und Sport Baden-Württemberg) [Eds.] (2004). Bildungsplan 2004. Allgemein bildendes Gymnasium. Ditzingen: Philipp Reclam Jun.

Introduction

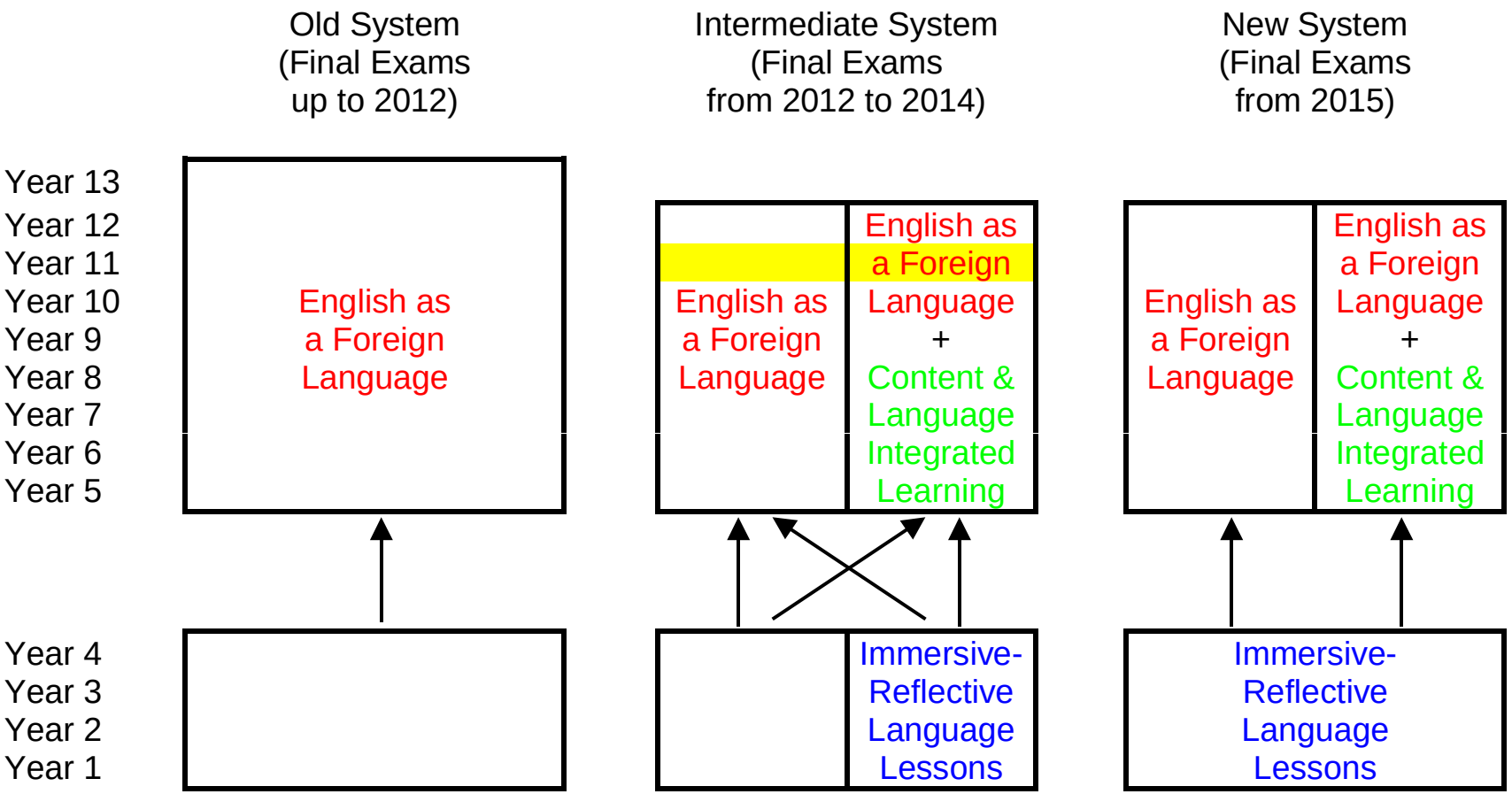
The Passive and the German Learner - What Does ICLE Say?



Granger, S. (2009): More lexis, less grammar? What does the (learner) corpus say?
Paper presented at the Grammar & Corpora conference, Mannheim, 22-24 September 2009.

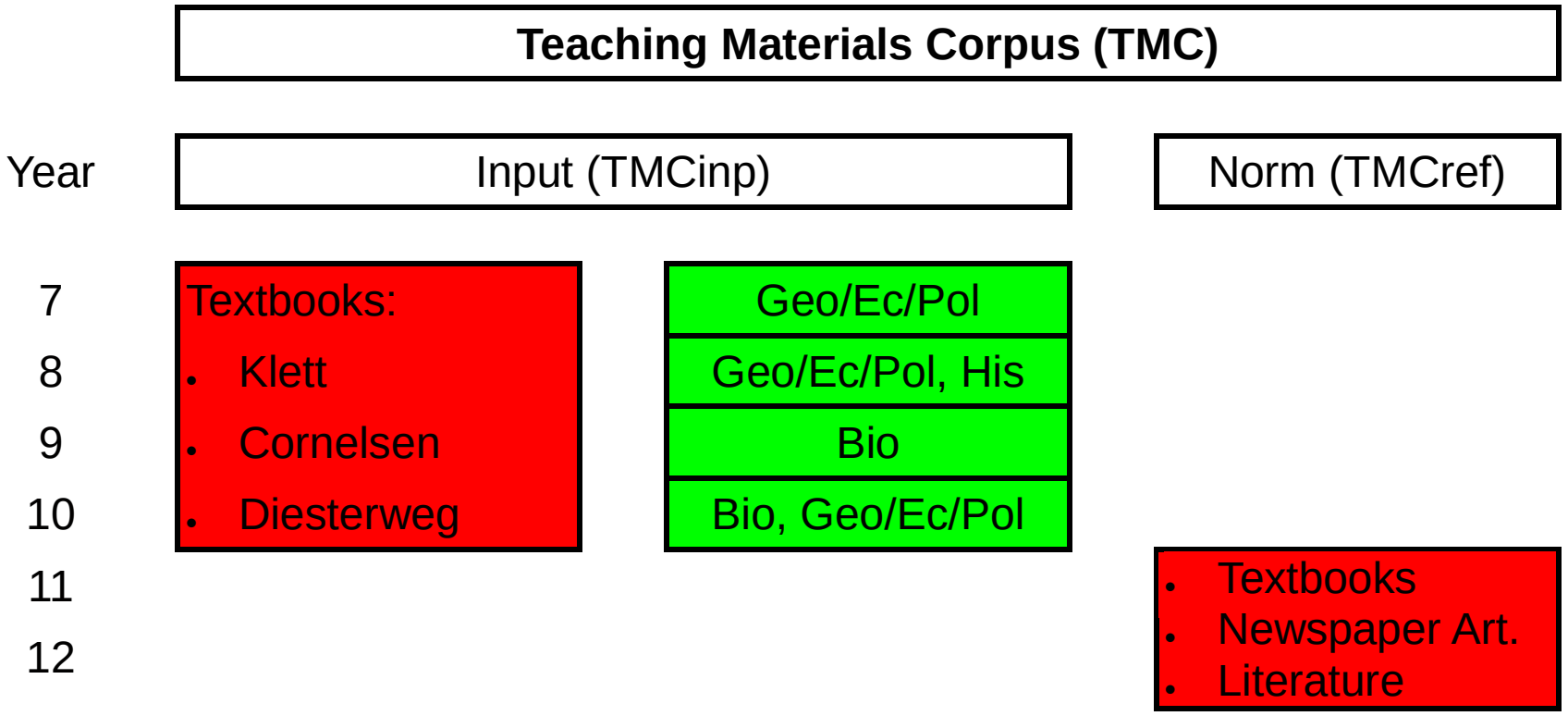
Introduction

Learning Environments - EFL and CLIL Programmes
at Secondary Schools *(Gymnasien)* in Baden-Württemberg

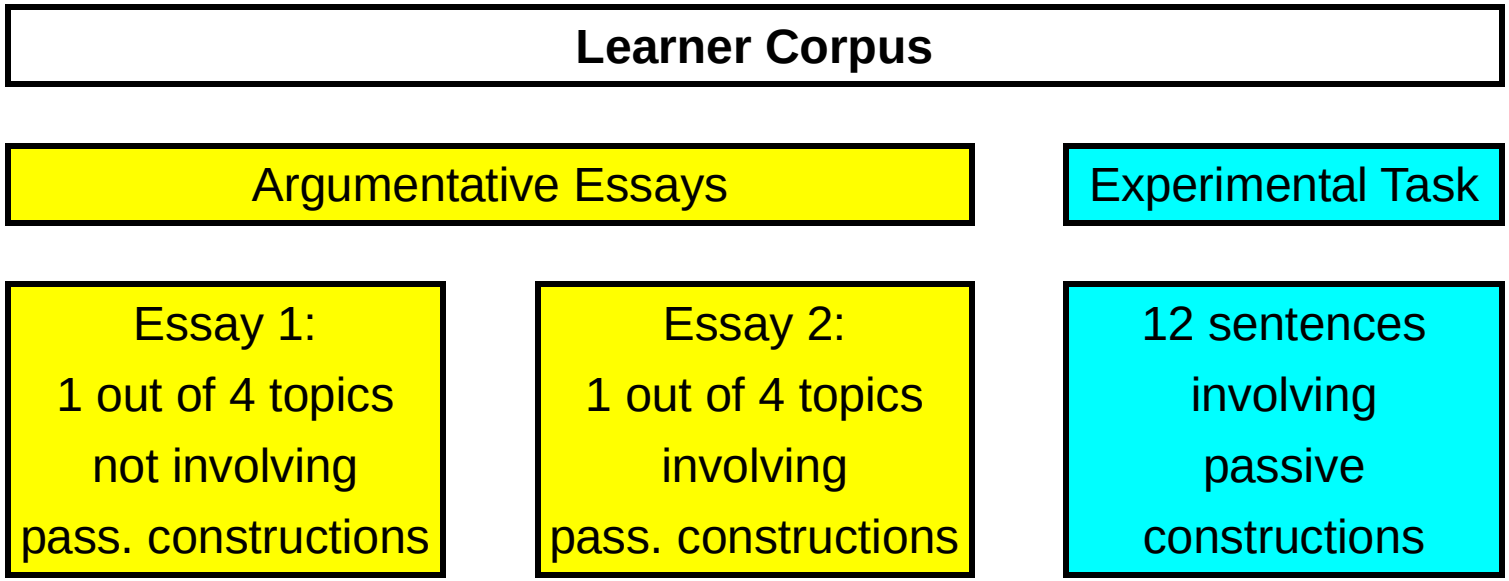


Input and Norm

The Teaching Materials Corpus (TMC)



The Learner Corpus



The Learner Corpus

Argumentative Essays

The Learner Corpus

Experimental Task

The Learner Corpus

Linguistic Annotation - Pilot Study

TreeTagger (Schmid 1994): POS-tagger, lemmatizer, tokenizer

- 423 <UNKNOWN>-tags in the pilot corpus
- > 50 % of <UNKNOWN> words received a correct POS tag

Example:

```
if      IN      if
the     DT      the
alcohol NN      alcohol
can     MD      can
be      VB      be
buyed   JJ      <unknown>
```

The Learner Corpus

Linguistic Annotation - Pilot Study

CLAWS (Garside/Smith 1997): POS-tagger, tokenizer

- no <UNKNOWN> tags,
  but some incorrect forms receive an <ERROR> tag
- 5.255 ambiguities (~17.000 words);
  88,4 % received a correct POS tag as a first alternative
  with a probability of ≥ 80 %

Example:

```
0000613 030 if          93 [CS/96] CSW@/4
0000613 040 the         93 AT
0000613 050 alcohol     93 NN1
0000613 060 can         93 [VM/100] NN1%/0 VV0%/0
0000613 070 be          93 VBI
0000613 080 buyed       06 [VVN@/99] JJ@/1 VVD/0
```

The Learner Corpus

Linguistic Annotation - Pilot Study

MATE (Bohnet 2010): parser, POS-tagger, lemmatizer, tokenizer

- no <UNKNOWN> tags

Example:

```
14 if      if      _   _   IN   _   _   10   10   NMOD   NMOD   _   _
15 the     the     _   _   DT   _   _   16   16   NMOD   NMOD   _   _
16 alcohol alcohol _   _   NN   _   _   17   17   SBJ    SBJ    _   _
17 can     can     _   _   MD   _   _   14   14   SUB    SUB    _   _
18 be      be      _   _   VB   _   _   17   17   VC     VC     _   _
19 buyed   buy     _   _   VBN  _   _   18   18   VC     VC     _   _
```

The Learner Corpus

Linguistic Annotation - Pilot Study

Target-like *be Ved* Constructions			
	TT	CL	MA
be + participle (n=129)	129	128	123

Erroneous *be Ved* Constructions			
	TT	CL	MA
Correct tag for *be* (n=16)	12	12	11
Correct tag for participle (n=22)	11	15	15
Correct tag for *be* and participle (n=16)	4	8	8

The Learner Corpus

Linguistic Annotation

Error Annotation:

e. g. UCLEE (Université catholique de Louvain Error Editor)

```
[...] if the alcohol can be (FM) buyed $bought$ [...]
```

Target Hypotheses:

cf. e. g. FALKO

```
[...] if the alcohol can be buyed [...]
[...] if the alcohol can be bought [...]
```

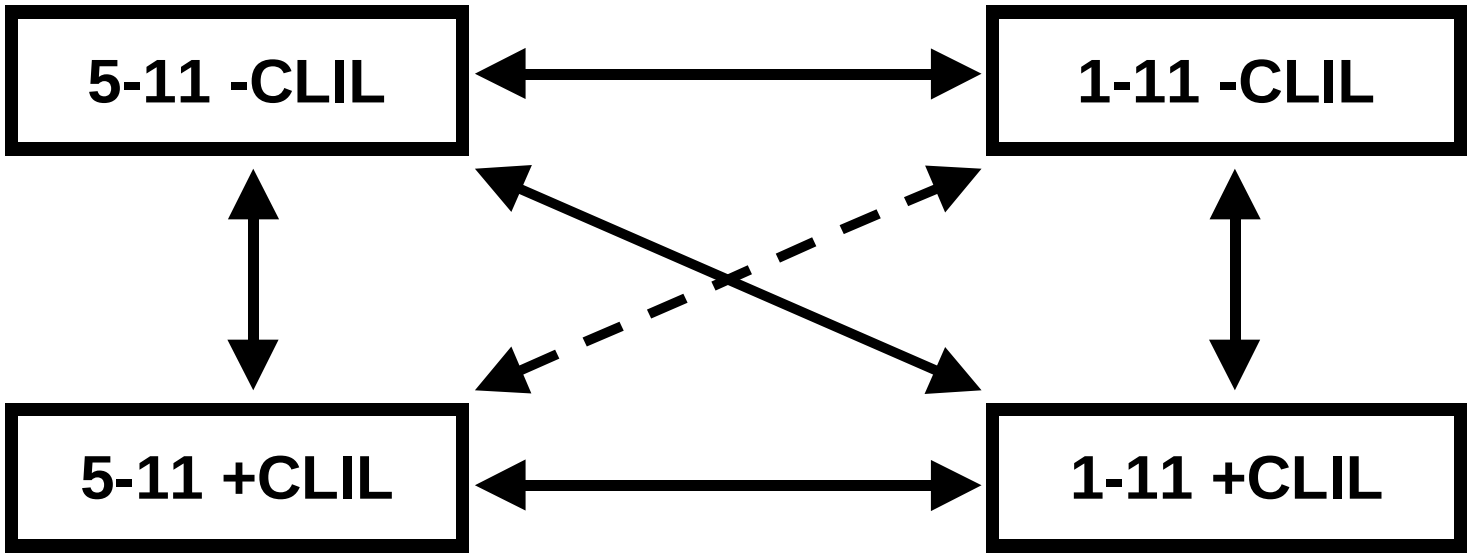
The Learner Corpus

Linguistic Annotation

```
0000003 010 Burglars   93 NN2
0000003 020 were       93 VBD
0000003 030 broken     03 VVN
0000003 040 into       93 PRP
0000003 050 our        93 DPS
0000003 060 house      93 NN1
0000003 061 .          03 .
```

The Learner Corpus

Metadata



Problem:

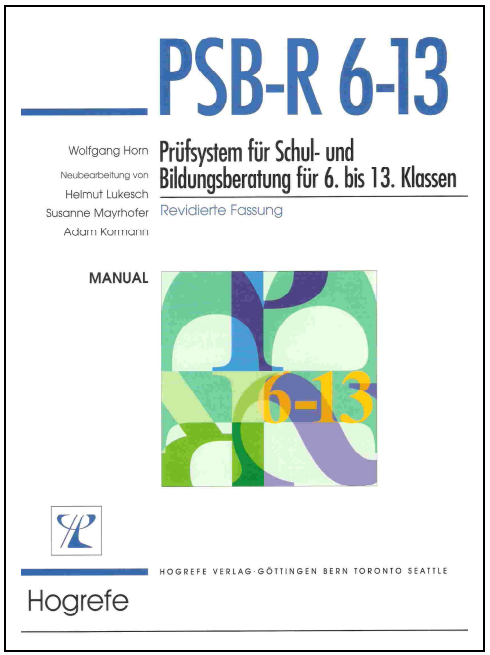
CLIL programmes are not compulsory

⇨ differences might be due to intervening variables

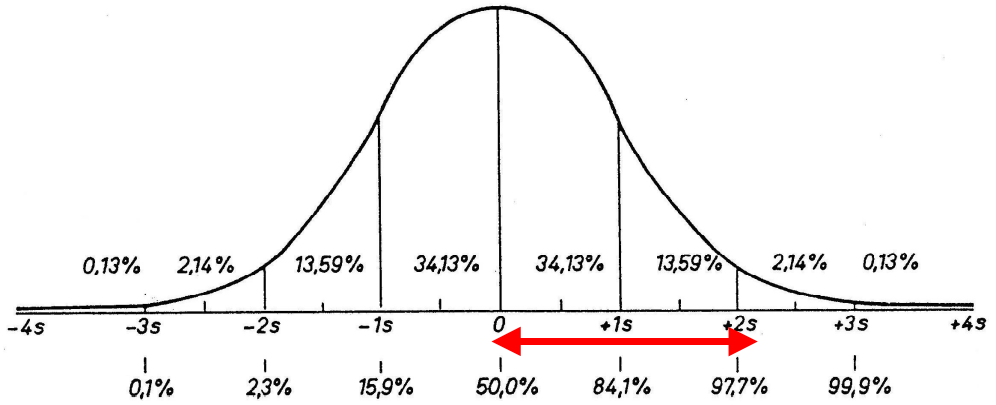
(e. g. cognitive capacities, motivation)

The Learner Corpus

Metadata

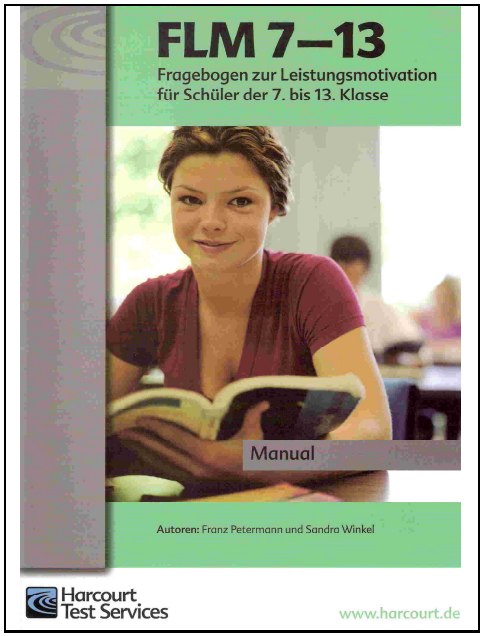


- Overall cognitive capacities
- Verbal cognitive capacities
- Word fluency (German)
- Language-related logical thinking
- Concentration



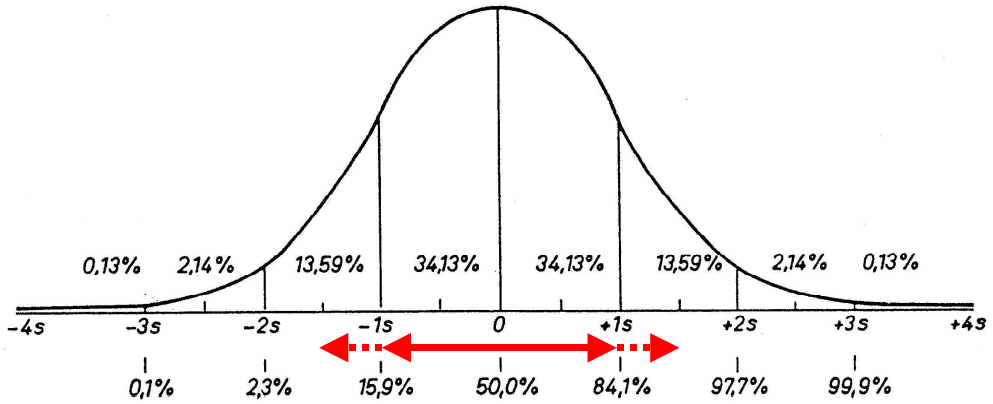
The Learner Corpus

Metadata



Aspects of motivation:

- Orientation towards performance and success
- Perseverance and effort



The Learner Corpus

Metadata

```
<meta idstudent="186" idschool="10" age="17" sex="f"
l1a="ge" l1b="x" lhomea="ge" lhomeb="x" stay="1"
l2a="en" l2b="fr" l2c="x" l2d="x" l2e="x" l2no="2“
l2noen="0" l2enyears="7" l2encomp="4" l2gecomp="x"
l2frcomp="3" l2lacomp="x" l2itcomp="x" l2spcomp="x"
double="0" skip="0" primger="4" primefl="1"
textbook="g20" clilyears="0" clilsubjects="x"
speak="3" read="3" watch="3" surf="3" psb1="94"
psb2="109" psb3="90" psb4="105" psb2-4="100"
psb5="117" psb6="114" psb7="105" psb8="104"
psb9="100" psbv="103" psbr="107" psbk="101"
psbgl="104" flmls="55" flmaf="52" flmae="42"
flmap="46" flmhp="69" exprat="28" topic="7">
[...]
</meta>
```

The Pilot Study

Passive Constructions in Learner Text

***be Ved:* 22 out of 151 erroneous**

- Omission of *be* (6 instances):

  *Should the death penalty _____ reintroduced in Germany?

- Morphological and/or orthographic errors

  in the form of *be* or related clitics (3 instances):

  **You arent forced to post anything in the internet.*

- Morphological and/or orthographic errors in the past participle (11 instances):

  **[...] if the alcohol can just be buyed by 21 old people.*

- Lexical errors (1 instance):

  **[...] so he is already prisoned by the police.*

- Combination of different types of error (1 instance):

  **[...] because it´s forbideden.*

The Pilot Study

Passive Constructions in Learner Text

***get Ved:* 3 out of 9 erroneous**

- Morphological and/or orthographic errors in the form of *be*:

  **You arent forced to post anything in the internet.*

- Morphological and/or orthographic errors

  in the form of *be* and the past participle :

  **If you imagine that your daughter gehts raped and murderd by a person, ...*

The Pilot Study

Passive Constructions in Learner Text

lemma	nlemma (LC)	npassive (LC)	passive ratio (LC)	passive ratio (TMC)
allow	22	21	95.5 %	25.5 %
(re)introduce	35	18	51.4 %	36.6 %
raise	48	13	27.1 %	25.3 %
kill	31	6	19.4 %	26.0 %
discuss	12	6	50.0 %	20.8 %
say	47	5	10.6 %	0.7 %
integrate	18	4	22.2 %	22.2 %

The Pilot Study

Passive Constructions in Learner Text

lemma	nlemma (LC)	npassive (LC)	passive ratio (LC)	passive ratio (TMC)
allow	22	21	95.5 %	25.5 %
(re)introduce	35	18	51.4 %	36.6 %
raise	48	13	27.1 %	25.3 %
kill	31	6	19.4 %	26.0 %
discuss	12	6	50.0 %	20.8 %
say	47	5	10.6 %	0.7 %
integrate	18	4	22.2 %	22.2 %

	Cornelsen English G 2000	Klett Green Line New	Diesterweg Camden Town
be allowed to	vol. 2	vol. 2	vol. 2
allow	vol. 5	---	vol. 3

The Pilot Study

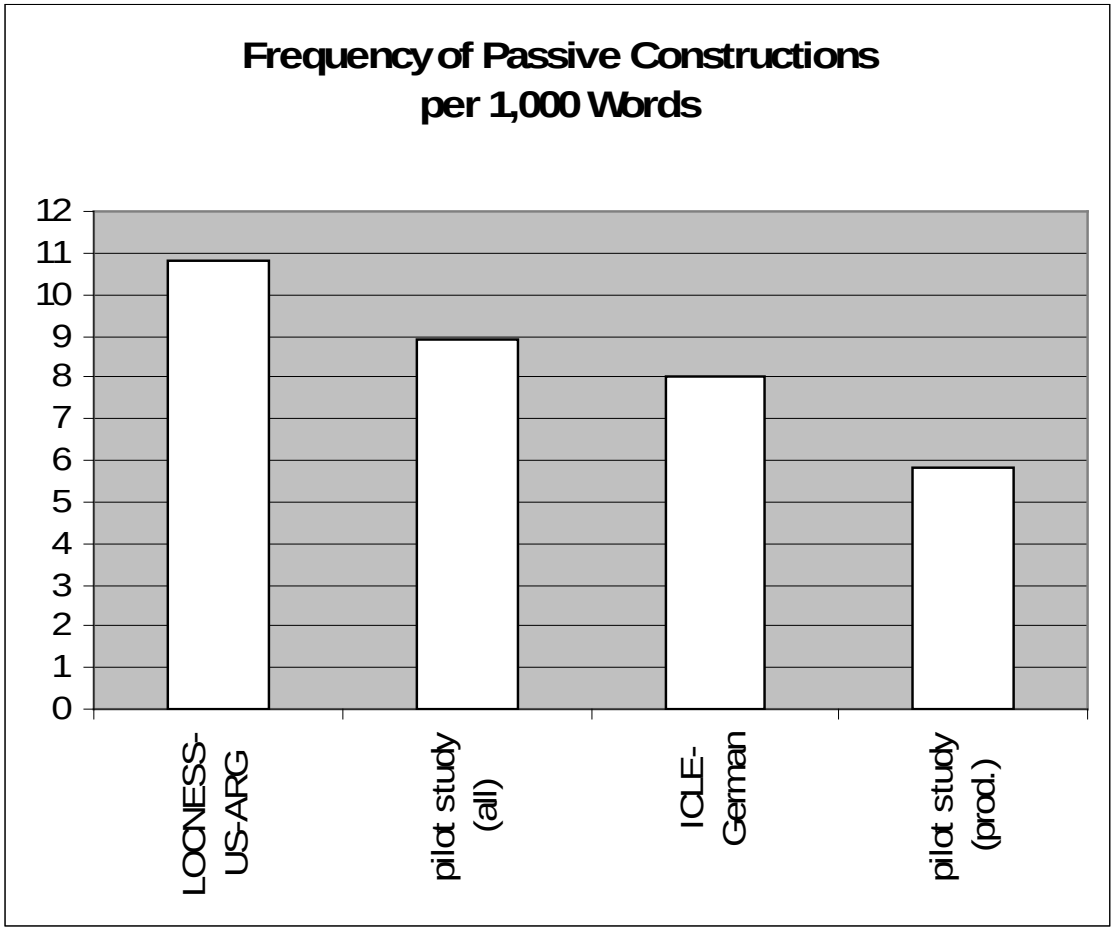
Passive Constructions in Learner Text

lemma	nlemma (LC)	npassive (LC)	passive ratio (LC)	passive ratio (TMC)
allow	22	21	95.5 %	25.5 %
(re)introduce	35	18	51.4 %	36.6 %
raise	48	13	27.1 %	25.3 %
kill	31	6	19.4 %	26.0 %
discuss	12	6	50.0 %	20.8 %
say	47	5	10.6 %	0.7 %
integrate	18	4	22.2 %	22.2 %

(re)introduce	The death penalty should be reintroduced in Germany.
raise	In order to fight teenage drinking, the legal drinking age should be raised to 21.

The Pilot Study

Passive Constructions in Learner Text



The Pilot Study

Passive Constructions in the Experimental Task

	Average self-rated reliability of response (1-5)	(Average) number of correct responses (n=28)	
monotransitive verb (6 sentences)	3,0	15,7	Structures treated explicitly by textbooks
ditransitive verb (2 sentences)	3,1	4,5	
prepositional verb (2 sentences)	2,6	9,5	Structures not treated explicitly by textbooks
complex-transitive verb (1 sentence)	3,0	10,0	
impersonal passive (1 sentence)	1,8	7,0	