

**Is interactive learning also active learning?
A quantitative and qualitative study in
computer assisted language learning**

Jane Vinther

**Ph.D. dissertation
Institute of Language and Communication
University of Southern Denmark
2007**

Supervisors

Previous:

Fritz Larsen, Associate professor, Emeritus, University of Southern Denmark.

The late John M. Dienhart, Associate Professor, University of Southern Denmark.

Jane Pilgaard Vinther
Institute of Language and Communication
University of Southern Denmark
Campusvej 55
DK-5230 Odense M.

Tel.: +45 65501457

E-mail: jvinther@language.sdu.dk

© Jane Pilgaard Vinther 2007

Contents

<i>List of tables</i>	9
<i>List of figures</i>	13
<i>List of TA Excerpts</i>	14

<i>Acknowledgements</i>	15
-------------------------	----

PART ONE

1 Introduction and research questions	16
1.1 The paradigms	16
1.2 The methodological framework	20
1.3 Research questions	21
1.4 The structure of the thesis	23

PART TWO

2 Second language acquisition: Review	26
2.1 Instruction	27
2.2 The role of implicit and explicit knowledge in second language acquisition	28
2.2.1 Definitions and constructs	28
2.2.1.1 Metalanguage	30
2.2.2 Interface positions between implicit and explicit knowledge	33
2.2.2.1 The no-interface position	33
2.2.2.2 The strong-interface position	35
2.2.2.3 The weak-interface position	35
2.3 Immersion	36
2.4 Focus on form	39
2.4.1 Attention, noticing, and awareness	40
2.5 Teachability	45
2.5.1 Sequences of learning	45
2.5.2 Complexity of rules	51
2.6 Summary	55
3 Computer assisted language learning: Review	57
3.1 History of CALL	57
3.2 CALL Research	60
3.2.1 Research paradigms	61
3.2.2 Efficacy	62
3.2.2.1 Computer versus classroom instruction	62
3.3 CALL and grammar	65

3.3.1	Feedback	65
3.3.2	Syntax	69
3.4	CALL at the tertiary level	70
3.4.1	English and CALL	71
3.4.2	Availability	72
3.4.3	Needs	73
3.5	Summary	75
4	The quantitative study	76
4.1	Background and aim	76
4.1.1	The two advanced study programmes selected	78
4.1.2	University entrance qualification	80
4.1.2.1	“Gymnasium” (academically oriented secondary education)	81
4.1.2.2	“HF” (Higher Preparatory Exam)	82
4.1.2.3	“HTX” (Higher Technical Exam)	82
4.1.2.4	“HHX” (Higher Business Exam)	83
4.1.2.5	General comments	84
4.2	The subjects	85
4.3	Experimental groups	88
4.3.1	Practical constraints	89
4.3.2	Assignment to experimental groups	90
4.4	Treatment	91
4.4.1	Amount of input	91
4.4.2	The syntactical input	91
4.4.3	Procedure	94
4.4.3.1	The VISL groups	94
4.4.3.2	The NON-VISL groups	95
4.4.3.3	Observable differences between the VISL and the NON-VISL groups	96
4.5	Designing the tests: Problems and considerations	97
4.5.1	The purpose of the tests	97
4.5.2	The method of testing	99
4.5.3	Sequence of items and the level of difficulty	99
4.5.4	Test items	104
4.5.4.1	Composing the tests	106
4.5.4.2	Underlying principles	107
4.5.5	Reliability	108
4.5.5.1	Evaluating pre-test against post-test	110
4.5.5.2	Rating the tests	114
4.5.5.3	Inter-rater reliability	117
4.6	Results of quantitative study: statistical description	117
4.6.1	Total test levels	119
4.6.1.1	Total pre-test levels	119
4.6.1.1.1	The Cand. Negot. cohort. Totals pre-test	119
4.6.1.1.2	The English cohort. Totals pre-test	120
4.6.1.1.3	VISL, NON-VISL, and BASIS (Total pre-test) differences	122

4.6.1.2	Total post-test levels	124
4.6.1.2.1	The Cand. Negot. cohort. Totals post-test	124
4.6.1.2.2	The English Cohort. Totals post-test	125
4.6.1.2.3	VISL, NON-VISL, and BASIS (Totals post-test) differences	127
4.6.1.3	Total gain levels	129
4.6.1.3.1	The Cand. Negot. cohort. Total gains	130
4.6.1.3.1.1	Relative gains. Total	130
4.6.1.3.1.2	Percentage points gains. Total	131
4.6.1.3.2	The English cohort. Total gains	132
4.6.1.3.2.1	Relative gains. Total	132
4.6.1.3.2.2	Percentage points gains. Total	133
4.6.1.3.3	VISL, NON-VISL, and BASIS (Total gain) differences	134
4.6.2	Syllabus test levels	137
4.6.2.1	Syllabus pre-test levels	138
4.6.2.1.1	The Cand. Negot.cohort. Syllabus pre-test	138
4.6.2.1.2	The English cohort. Syllabus pre-test	139
4.6.2.1.3	VISL, NON-VISL, and BASIS (syllabus pre-test) differences	140
4.6.2.2	Syllabus post-test levels	142
4.6.2.2.1	The Cand.Negot.cohort. Syllabus post-test	142
4.6.2.2.2	The English cohort. Syllabus post-test	144
4.6.2.2.3	VISL, NON-VISL, and BASIS (syllabus post-test) differences	145
4.6.2.3	Syllabus gain levels	148
4.6.2.3.1	The Cand. Negot.cohort. Syllabus gains	148
4.6.2.3.2	The English cohort. Syllabus gains	149
4.6.2.3.3	VISL, NON-VISL, and BASIS (syllabus gain) differences	150
4.6.3	Non-syllabus test levels	152
4.6.3.1	Non-syllabus pre-test levels	152
4.6.3.1.1	The Cand. Negot. cohort. Non-syllabus pre-test	153
4.6.3.1.2	The English cohort. Non-syllabus pre-test	154
4.6.3.1.3	VISL, NON-VISL, and BASIS (non syllabus pre-test) differences	155
4.6.3.2	Non-syllabus post-test levels	156
4.6.3.2.1	The Cand. Negot. cohort. Non-syllabus post-test	156
4.6.3.2.2	The English cohort. Non-syllabus post-test	158
4.6.3.2.3	VISL, NON-VISL, and BASIS (non-syllabus post-test) differences	159
4.6.3.3	Non-syllabus gain levels	161
4.6.3.3.1	The Cand. Negot. cohort. Non-syllabus gains	162
4.6.3.3.2	The English cohort. Non-syllabus gains	163

4.6.3.3.3	VISL, NON-VISL, and BASIS (non-syllabus gain) differences	164
4.6.4	Success rate (post-test) differences in syllabus and non-syllabus results	166
4.6.4.1	VISL, NON-VISL, and BASIS syllabus success rate (post-test) differences	166
4.6.4.2	VISL, NON-VISL, and BASIS non-syllabus success rate (post-test) differences	169
4.6.5	Gain rate differences in syllabus and non-syllabus results	172
4.6.5.1	VISL, NON-VISL, and BASIS syllabus gain rate differences	172
4.6.5.2	VISL, NON-VISL, and BASIS non-syllabus gain rate differences	173
4.6.5.3	Correlating syllabus and non-syllabus gains	173
4.6.6	Results of discrete test item results	176
4.6.6.1	Initial knowledge	177
4.6.6.1.1	Initial knowledge: syllabus	177
4.6.6.1.2	Initial knowledge: non-syllabus	180
4.6.6.1.3	Initial differences in syllabus and non-syllabus knowledge	181
4.6.6.2	The efficacy of the treatment on post-test results of discrete items	185
4.6.6.2.1	Syllabus post-test item results across treatment groups	185
4.6.6.2.2	Non-syllabus post-test item results across treatment groups	188
4.6.6.3	The efficacy of the treatment on gain results of discrete items	192
4.6.6.3.1	Syllabus item gain results	194
4.6.6.3.2	Non-syllabus item gain results	196
4.6.6.4	VISL, NON-VISL and BASIS discrete item result differences	201
4.7	Attendance	204
4.8	Summary of results	207

PART THREE

5	Processes of learning: Review	210
5.1	Consciousness	210
5.2	Implicit vs. explicit learning	212
5.2.1	Implicit and explicit knowledge	214
5.2.2	The implicit-explicit interface	216
5.3	Artificial language learning	218
5.4	Automatisation	221
5.5	Summary	223

6	Introspection	224
6.1	Verbal report as a research tool	224
6.1.1	Levels of verbalisation	226
6.1.2	Character of reports	227
6.2	Retrospective reports	228
6.3	Concurrent reports	229
7	The qualitative study	233
7.1	The context	233
7.1.1	General	233
7.1.2	The design background	233
7.1.3	The preceding treatment	234
7.2	Methodology	234
7.2.1	Collecting the protocols	234
7.2.1.1	Rationale for the choice of type of protocol	234
7.2.1.2	Instruction	234
7.2.1.3	The recordings	235
7.2.1.4	The subjects	235
7.2.1.5	The subject-matter content	236
7.2.1.6	Task analysis	237
7.2.1.7	The role of the experimenter	239
7.2.1.8	Protocols	240
7.2.2	Content of the protocols	240
7.2.2.1	Elicitation of information	240
7.2.2.2	Categories of information	241
7.3	Interpreting the protocols	242
7.3.1	Dienes and Perner's theory of implicit and explicit knowledge	245
7.3.2	Verbal manifestations and content categories	248
7.3.2.1	Fully explicit knowledge	248
7.4	Results	250
7.4.1	Inferences in the interpretation process	250
7.4.1.1	Self	250
7.4.1.2	Attitude	253
7.4.1.3	Content	256
7.4.2	Explicit knowledge and consciousness	265
7.4.3	Knowledge, accessibility and control	277
7.4.4	The differential effects of the treatment	284
7.4.4.1	The high-achievers	286
7.4.4.2	The middle group	290
7.4.4.3	The low-achievers	292
7.4.5	Summary of VISL and NON-VISL results	294

PART FOUR

8	Discussion and conclusions	297
8.1	Discussion of the findings for the research questions	297
8.2	Conclusions and perspectives	302

<i>English summary</i>	308
<i>Danish summary</i>	315
<i>Bibliography</i>	320
<i>Appendices</i>	338
<i>I</i> Analysed and quantified sentence corpus	338
<i>II</i> Key to VISL symbol set	369
<i>III</i> Pre-test	370
<i>IV</i> Post-test	377
<i>V</i> Overview of entrance levels A and B	384
<i>VI</i> English supplement to “Gymnasiebekendtgørelsen”	388
<i>VII</i> Easiness of items at pre-score and post-score by groups.	401
<i>VIII</i> Ranking pre-, post-, gain scores, syllabus and non-syllabus	408
<i>IX</i> Gains(%) by experimental group	411
<i>X</i> Think-aloud excerpts with their full protocols	413

List of tables

Table 4.1	Years of instruction prior to university. By treatment group. Cand. Negot.	85
Table 4.2	Years of instruction prior to university. By treatment group. English.	86
Table 4.3	Type of exam. By treatment group. Cand. Negot.	87
Table 4.4	Type of exam. By treatment group. English.	88
Table 4.5	Easiness of items at pre-score over all groups.	102
Table 4.6	Easiness of items at post-score over all groups.	103
Table 4.7	Pre-experimental test of easiness of pre-test and post-test. Total result.	111
Table 4.8	Pre-experimental test of easiness pre-test and post-test. Detailed results.	111
Table 4.9	Inter-rater reliability. Pre-test and post-test.	117
Table 4.10	Number of subjects per treatment per cohort.	118
Table 4.11	Cand. Negot. cohort. Pre-test levels. Point scores. Total (max. 25).	119
Table 4.12	English cohort. Pre-test levels. Point scores. Total (max. 25).	121
Table 4.13	Cand. Negot.cohort. Pre-test levels (percentage success rates). Total (%).	122
Table 4.14	English cohort. Pre-test levels (percentage success rates). Total (%).	123
Table 4.15	Cand. Negot. cohort. Post-test levels. Point scores. Total (max. 25)	125
Table 4.16	English cohort. Post-test levels. Point scores. Total (max. 25).	127
Table 4.17	Cand. Negot. cohort. Post-test levels. Success rates. Total (%).	128
Table 4.18	English cohort. Post-test levels. Success rates. Total (%).	128
Table 4.19	Cand. Negot. cohort. Relative gain levels. Total.	130
Table 4.20	Cand. Negot. cohort. Percentage point gain levels.Total.	131
Table 4.21	English cohort. Relative gain levels. Total.	132
Table 4.22	English cohort. Percentage points gain levels. Total.	133

Table 4.23	Cand. Negot. cohort. Comparison of per cent difference. (Scheffe). Gain (%). Total.	136
Table 4.24	English cohort. Comparison of per cent difference (Scheffe). Gain (%). Total.	136
Table 4.25	Cand. Negot. cohort. Pre-test levels. Point scores. Syllabus (max. 16).	138
Table 4.26	English cohort. Pre-test levels. Point scores. Syllabus (max. 16).	139
Table 4.27	Cand. Negot. cohort. Pre-test levels (percentage success rates). Syllabus (%).	141
Table 4.28	English cohort. Pre-test levels (percentage success rates). Syllabus (%).	141
Table 4.29	Cand. Negot. cohort. Post-test levels. Point scores. Syllabus (max. 16).	142
Table 4.30	English cohort. Post-test levels. Point scores. Syllabus (max. 16).	144
Table 4.31	Cand. Negot. cohort. Post-test levels. Success rate. Syllabus (%).	145
Table 4.32	English cohort. Post-test levels. Success rates. Syllabus (%).	147
Table 4.33	Cand. Negot. cohort. Gain levels. Syllabus (%).	148
Table 4.34	English cohort. Gain levels. Syllabus (%).	149
Table 4.35	Cand. Negot. cohort. Comparison of per cent difference. (Scheffe). Gain (%). Syllabus.	151
Table 4.36	English cohort. Comparison of per cent difference (Scheffe). Gain (%). Syllabus.	152
Table 4.37	Cand. Negot. cohort. Pre-test levels. Point scores. Non-syllabus (max. 9).	153
Table 4.38	English cohort. Pre-test levels. Point scores. Non-syllabus (max. 9).	154
Table 4.39	Cand. Negot. cohort. Pre-test levels percentage. Non-syllabus (%).	155
Table 4.40	English cohort. Pre-test levels percentage. Non-syllabus (%).	156
Table 4.41	Cand. Negot. cohort. Post-test levels. Point scores. Non-syllabus (max. 9).	157

Table 4.42	English cohort. Post-test levels. Point scores. Non-syllabus (max. 9).	158
Table 4.43	Cand. Negot. cohort. Post-test levels. Success rates. Non-syllabus (%).	160
Table 4.44	English cohort. Post-test levels. Success rate. Non-syllabus (%).	160
Table 4.45	All experimental groups overview. Non-syllabus mean gains (points).	162
Table 4.46	Cand. Negot. cohort. Gain levels. Non-syllabus (%).	163
Table 4.47	English cohort. Gain levels. Non-syllabus (%).	164
Table 4.48	Cand. Negot. cohort. Comparison of per cent difference (Scheffe). Gains. Non-syllabus.	165
Table 4.49	English cohort. Comparison of per cent difference (Scheffe). Gains. Non-syllabus.	165
Table 4.50	Overview mean values syllabus (in %).	168
Table 4.51	Overview mean values non-syllabus (in %).	171
Table 4.52	Cand. Negot. cohort. Difference of differences: syllabus % difference (post-pre) minus non-syllabus % difference (post-pre).	174
Table 4.53	English cohort. Difference of differences: syllabus % difference (post-pre) minus non-syllabus % difference.	174
Table 4.54	Total difference by gains (%) syllabus and non-syllabus.	175
Table 4.55	Easiness of syllabus items at pre-test.	178
Table 4.56	Easiness of non-syllabus items at pre-test.	181
Table 4.57	Word class (syllabus) pre-test (%).	182
Table 4.58	Syntax items (syllabus) pre-test (%).	182
Table 4.59	Syntax (non-syllabus) pre-test (%).	183
Table 4.60	Morphology (non-syllabus) pre-test (%).	183
Table 4.61	Tense/aspect (non-syllabus) pre-test (%).	184
Table 4.62	Easiness of syllabus items at post-test.	185
Table 4.63	Syntax items (syllabus) post-success rate (%).	187
Table 4.64	Word class (syllabus) post success rate (%)	188
Table 4.65	Easiness of non-syllabus items at post-test.	189
Table 4.66	Syntax (non-syllabus) post success rate (%).	190
Table 4.67	Morphology (non-syllabus) post success rate (%).	191

Table 4.68	Tense/aspect (non-syllabus) post success rate (%).	191
Table 4.69	Item gain all groups.	193
Table 4.70	Syllabus gain results.	195
Table 4.71	Word class (syllabus) gain (%).	195
Table 4.72	Syntax items (syllabus) gain (%).	196
Table 4.73	Non-syllabus gain results.	196
Table 4.74	Syntax (non-syllabus) gain (%).	198
Table 4.75	Morphology (non-syllabus) gain (%).	198
Table 4.76	Tense/aspect morphology (non-syllabus) gain (%).	199
Table 4.77	Differences between non-syllabus item groups.	200
Table 4.78	Differences (Scheffe) for item 6 between treatment groups.	201
Table 4.79	Differences (Scheffe) for item 23 between treatment groups.	201
Table 4.80	Distribution of attendance all subjects.	204
Table 4.81	NON-VISL treatment. Distribution of attendance.	204
Table 4.82	VISL treatment. Distribution of attendance.	205
Table 4.83	Cand. Negot. High attendance distribution.	206
Table 4.84	English. High attendance distribution.	206
Table 7.1	Entrance levels of TA subjects.	236
Table 7.2	Task analysis for the think-aloud tasks.	238
Table 7.3	Implicitness, explicitness and possible combinations.	257
Table 7.4	Think-aloud subjects pre-test ranking for syllabus.	285
Table 7.5	Think-aloud subjects post-test ranking syllabus and exam.	285
Table 7.6	Key figures for TA Subjects	289

List of figures

Figure 4.1	Experimental groups by study programme.	90
Figure 4.2	Overall comparison pre-test against post-test. Pre-experimental.	113
Figure 4.3	Pre-test first. Point score total by subject. Pre-experimental.	114
Figure 4.4	Post-test first. Point score total by subject. Pre-experimental	114
Figure 4.5	Hours of training per experimental group.	118
Figure 4.6	Scatterplot. Total scores pre-test by post-test per experimental groups.	126
Figure 4.7	Boxplot. Syllabus post-test results by experimental groups.	143
Figure 4.8	Scatterplot. Syllabus scores pre-test by post-test per experimental groups.	146
Figure 4.9	Boxplot. Syllabus gains percentage points by treatment groups.	150
Figure 4.10	Boxplot. Cand. Negot. cohort. Non-syllabus. Pre-test by post-test by experimental groups.	157
Figure 4.11	Boxplot. English cohort. Non-syllabus. Pre-test by post-test by experimental groups.	159
Figure 4.12	Scatterplot. Non-syllabus scores pre-test by post-test per experimental groups.	161
Figure 4.13	Item results pre-test ranking by post-test ranking.	190
Figure 4.14	Gains per item, non-syllabus versus syllabus combined for all groups.	194
Figure 4.15	Boxplot of attendance by treatment group.	205
Figure 5.1	Graphic presentation of the hierarchy of explicitness.	215
Figure 6.1	Memory model.	231

List of TA Excerpts

TA Excerpt 7.1	TAS4. POST-Sentence 2.	243
TA Excerpt 7.2	TAS7. PRE-Sentence 1a.	251
TA Excerpt 7.3	TAS7. PRE-Sentence 1b.	252
TA Excerpt 7.4	TAS2. PRE-Sentence 1a.	252
TA Excerpt 7.5	TAS2. PRE-Sentence 1b.	254
TA Excerpt 7.6	TAS9. POST-Sentence 1a.	256
TA Excerpt 7.7	TAS8. POST-Sentence VISL on paper. a.	258
TA Excerpt 7.8	TAS8. POST-Sentence VISL on paper. b.	260
TA Excerpt 7.9	TAS11. PRE-Sentence 1a.	261
TA Excerpt 7.10	TAS11. PRE-Sentence 1b.	262
TA Excerpt 7.11	TAS7. PRE-Sentence 1c.	257
TA Excerpt 7.12	TAS11. PRE-Sentence 1c.	273
TA Excerpt 7.13	TAS3. POST-Sentence VISL on paper.	279
TA Excerpt 7.14	TAS3. PRE-Sentence 2.	280
TA Excerpt 7.15	TAS5. PRE-Sentence 1.	282
TA Excerpt 7.16	TAS3. POST-Sentence 1.	286
TA Excerpt 7.17	TAS9. POST-Sentence 1b.	287
TA Excerpt 7.18	TAS10. POST-Sentence 2.	290
TA Excerpt 7.19	TAS15. POST-Sentence 1.	291
TA Excerpt 7.20	TAS5. POST-Sentence 1.	293

Acknowledgements

To the colleagues, researchers, friends and family who have encouraged and supported me in the course of this project I wish to express my gratitude and thankfulness for their enthusiasm on my behalf. I cannot name them all, but please know that your support was appreciated. My Kolding colleagues I want to give my warmest thanks: Dr. Roy Sellars for letting me take advantage of his native-speaker knowledge of the English language; Dr. Henrik Larsen for keeping me at it, and Lone A. Mikkelsen for helping me keep on track with my administrative duties.

My colleagues at universities abroad have supported me by including me in the research networks and communities of my field. In particular, I would like to mention Professor Carol Chapelle, Iowa State University, for her inspiring teaching at SDU's Ph.D. summer school in 2002; Professor Joan Jamieson, University of Northern Arizona, for her encouragement during her sojourn as a Fulbright Professor at the SDU Kolding Campus in 2006; Professor Nick Ellis for his inspirational comments which influenced the structure of my thesis; Professor Michael Sharwood Smith for welcoming me at Heriot-Watt University, Edinburgh.

I want to thank my initial two supervisors, Fritz Larsen and John M. Dienhart, who helped shape my project, and whose comments have been invaluable. I also thank the Dean of Humanities, Dr. Flemming G. Andersen for offering collegial advice.

My innermost feeling of gratitude goes to the two persons who have suffered the greatest sacrifices and borne with me over the years without making an issue of it, namely my son Christian and my daughter Christine.

PART ONE

1 INTRODUCTION AND RESEARCH QUESTIONS

The combination of computers and second language learning is relatively new. A beginning had been made in the time of main frame computers, but the development took off in earnest with the creation of the personal computer, which celebrated its 25th anniversary last year¹. The areas of computer assisted language learning (CALL) and second language learning (SLA) share the feature that they need to draw on several fields of research on which they depend for theory and from which they gain some of their insights and methods. Consequently, SLA and CALL also share the struggle to establish paradigms of their own.

CALL was initially driven by a fascination with the new technology and the new horizons it opened, but not until Carol Chapelle's (1997) appeal to direct attention to the issues of second language learning (SLA) did the field put learning and pedagogical issues in the forefront. By this time researchers at the English Department of the then Odense University were in the middle of developing a learning tool to support the formal aspects of the linguistic syllabus². This developed into the netbased Visual Interactive Syntax Learning tool, which is at the centre of this thesis.

1.1 The paradigms

Second language acquisition is a complex research area for a number of reasons, and a plethora of diverse avenues of exploration have helped inform the many strands of interests which underpin our knowledge in this field. SLA can be researched from a neurological, a processing or a pedagogical starting point. Each of these can be approached from a comprehension or a production angle. Further sub-direction can be seen from a teaching or a learning point of view. Within either field the focus can be on

¹ IBM released its personal computer, IBM PC, in 1981. Apple released the Mackintosh, the first computer with a 'mouse', in 1984.

² The driving force behind the group was Associate Professor John M. Dienhart.

linguistic items, on age issues, or individual learner differences, to name some. The actual areas of SLA research investigation constitute a long and varied range of topical issues. A characteristic feature of SLA research is that it is not a case of one generation of research findings standing on the shoulders of the previous one in the sense that there is only one direction and one goal which will come nearer with each new piece of evidence. SLA research forms new realisations and visions by adding new knowledge in the many parallel lines of research rather than proceeding in a hierarchical framework. The difficulty from a traditional research point of view, given the complexity and the number of variables, is that of control. Learning is in many ways an individual process, and the many variables in a volatile environment make it difficult to establish clinical conditions for experiments, so this calls for humility in the interpretation of results. Add to this the omnipresent philosophical schism between those who see innateness as the underlying condition of human language and those who see linguistic processes as no different from other cognitive activities and processes.

The desire to control as many variables as possible has led to experiments with artificial language learning³. This type of experiment is usually computer controlled to minimise the human factor and to secure control of input variation. Results have not led to any conclusive evidence, and some researchers, among them Hulstijn (1997), who has carried out some of these experiments himself, have warned that results from this type of research are transferable to real life situations only with caution. The missing factor in these experiments is the always present X-factor, the human mind, its versatility and plasticity. Even neuro-imaging results need interpreting, and gaining access to the workings of the human mind constitutes an enormous challenge and obstacle.

One of the areas of contentions in SLA is the development of proficiency and the relationship between implicit and explicit knowledge in facilitating such development. Krashen's Input Hypothesis (see for instance 1983, 1985) has been very influential in

³ There is also artificial grammar learning. The difference is vague. Artificial language learning may have more rules and more complex rules than artificial grammar learning which resembles morphology acquisition in that it often consists of learning suffixes which consist of combinations of letters.

promoting the view that explicit instruction⁴ is of limited value in that explicit knowledge could never develop into implicit knowledge, which, it is assumed, is what speakers rely on for fluency. In recent years, especially after the Canadian experience with immersion education showed that students failed to acquire a sufficient correctness level⁵, the focus has returned to the role and possibilities of explicit instruction and knowledge. The ‘middle-of-the-road’ stance, also known as the weak-interface position⁶, has long been widely held by SLA researchers, but recent research by VanPatten & Cadierno, Robinson, and DeKeyser has invigorated interest in explicit instruction, especially in grammar. One interesting new theory, which in a way short-circuits the implicit/explicit dilemma, has been put forward by Sharwood Smith (2004; 2005), and Sharwood Smith and Truscott (2005). They maintain that for adults there are two routes to fluency. On the one hand, there is the traditional implicit acquisition, which is available to children, and on the other hand, there is the explicit knowledge, which is an additional possibility available to adults, and the nature of which is to be compared to the acquisition of literacy. This knowledge can through practice be used almost as fluently as implicitly acquired knowledge.

In CALL, research in the pedagogical aspects is in its infancy and based primarily on theories related to SLA. One of the generally accepted tenets of CALL is that comparative research, i.e. between CALL methods and traditional classroom methods, is unnecessary. The question is not whether CALL should have a place in the curriculum and the classroom; rather, the question is how it should be deployed and what would be the most beneficial approach in order to promote learning and increase motivation. Consequently, the rigorous, controlled type of research involving CALL is scarce. The eagerness and the need to establish a paradigm of its own have resulted in a focus on CALL itself and the technological possibilities rather than deliberations on effectiveness in comparison to other methods of instruction. The bulk of CALL literature deals with descriptions of new software or courseware and how it has been used with students, or it deals with traditional pedagogical aspects of language learning such as motivation. The internet has played a large role in the development of CALL,

⁴ It should be noted that explicit instruction does not necessarily result in explicit knowledge.

⁵ See for instance Kowall & Swain (1997), Swain & Lapkin (1986), Swain (1998), and White (1991).

⁶ See chapter 2.2.2 for a more detailed explanation.

and the literature on how e-mail exchanges and chat rooms have been exploited in language learning features prominently. This was related to the new opportunity to enter into genuine communication with native speakers and the possibility of giving learners genuine tasks which would add authenticity to the learning situation. Perhaps as a natural consequence of the communicative approach to language learning in general, the actual interlanguage development and proficiency effects have been assumed rather than measured.

In recent years, a renewed interest in blended learning and curriculum development has expanded CALL in new directions. A new generation of interactive material and courseware has meant more individualised instruction and a focus on autonomy. The more 'intelligent' applications have led to better feedback, and the field is beginning to place stronger emphasis on a more structured learning interface. The old division of CALL materials into either communicative CALL or grammar drills is disappearing to the advantage of applications which take learner input into consideration, resulting in a rapprochement of CALL and SLA.

From a Danish perspective, the use of CALL is something which takes place primarily at pre-university levels, especially where English as a second language is concerned. This may, among other things, be due to the relatively advanced proficiency level in English of Danish university students. No doubt, most CALL materials are developed in the English speaking countries, especially the USA and Australia, and they are targeted at a different audience and a different level of competence than that of Danish university students of English. There is a shortage of high quality, high level CALL materials which can meet the need of students such as the Danish ones, and which can satisfy the academic, not just the pedagogical, requirements of study programmes. Such CALL materials need to be able to meet a variety of conditions: they must have linguistic and pedagogical credibility and they must further student autonomy, as a minimum. VISL is able to meet these criteria.

1.2 The methodological framework

The present thesis comprises an experimental study of two instructional methods, their effectiveness in quantitative terms, and their differential effects on task approach and cognitive processes from a qualitative perspective. The quantitative section of the thesis measures results in a pre-test/ post-test design, which means that each subject is measured against himself/herself for pre-test results, post-test results and gains. Consequently, descriptive statistics figure prominently in this section. Key results are investigated by means of statistical tests to gauge significant differences.

The qualitative section of the study is founded on think-aloud data which are interpreted and analysed within the framework of a theory of implicit and explicit knowledge developed by Dienes and Perner (1999). As emphasised above there is limited access to the mental processes involved in learning. One of the legitimate methods within the possibilities of humanistic research is introspection. It is a method which some would consider controversial, but there are several types of introspection, e.g. stimulated, retrospective, and concurrent. Each of these have qualities and problems which need to be weighed and considered in the light of the aim of the study in order to find the method best suited to a given purpose. Even under the best of circumstances it is a method which is 'messy' in the sense that it produces a huge amount of raw data that need to be ordered, categorised, interpreted - and inferences need to be placed in the given context. However, it is a window directly into the best source of information we have: the student's cognitive processes.

In order to rule out as many potential sources of interference as possible, and in order to get the data without too many filters, the concurrent think-aloud type of introspection was chosen for this study. The data were categorised and analysed in accordance with Dienes and Perner's theory, which is very well suited to be used with verbal data of this kind as it is possible to operationalise their theory, basically because the theory takes verbalisation into account, and because of the way it relates verbalisation to declarative and explicit knowledge. Dienes and Perner's theory stipulates a hierarchy of explicit knowledge, and when applied to the think-aloud data, the potentially available

information resulting from this analysis should give us an idea of the explicit knowledge held by the student. The think-aloud data further allow a view into the task approaches adopted by the students, which in principle should provide information about the learning processes which the quantitative data alone cannot reveal. The combination of the methods described should give a more complete view into the complexity of the learning situation than either method applied separately can do.

1.3 Research questions

The overall aim of this study is to investigate the effectiveness of VISL compared to the traditional classroom instruction. Furthermore, the study wants to evaluate the effectiveness of the instruction in different subject groups by looking at the degree to which explicit knowledge results from the instruction. The last issue which is of interest, but not at the centre of the study, is whether there would be any measurable or detectable effect of the syllabus instruction on the non-syllabus proficiency.

The research questions are as follows:

1. Is VISL as good as traditional classroom instruction?

Notes:

There was no pre-experiment hypothesis, one way or the other. It may be difficult to create exactly the same conditions for a comparison between computer-based learning and classroom-based learning. The computer provides instant feedback whereas feedback from the experimenter in the NON-VISL classroom is delayed or may even be absent despite attempts to be omnipresent. Equally, it is not possible in any comparable way to supply the classroom students with an equivalent to the supporting features of the VISL interface.

2. Is VISL equally good for English students and Cand. Negot. students?

Notes:

The pre-experiment hypothesis was that the two student groups would be affected equally. They have met the same entrance qualifications, but they may not have the same type of entrance exam. English students study English as a foreign language with the traditional emphasis on literature, social studies and linguistics whereas the Cand. Negot. study programme is a combination of language and economy. These differences (see Chapter 4.1 for further details) might give rise to a difference in interest and emphasis on the formal linguistic aspect of language learning between the two experimental cohorts, and it was therefore an interesting question to investigate.

3. Is VISL equally good for all achievement groups, i.e. high, middle, and low achievers?

Notes:

There was no pre-experiment hypothesis, since there was no pre-experimental hypothesis about the overall effectiveness of VISL. If anything, the expectation was that especially the high and the middle achievers would benefit from VISL as they, as a rule, benefit most from instruction generally. An ensuing expectation concerning the VISL courseware would be to hypothesise a detrimental effect for the low-achievers in that there is no chance to get individual assistance over and beside the general supportive features in the VISL interface.

4. Do students achieve full explicitness of the subject matter?

Notes:

The pre-experimental hypothesis was that full explicitness would probably only be achieved by a handful of students out of the 107 subjects. Due to its qualitative nature, an answer to this question was sought through the application

of Dienes and Perner's theory which allows for a hierarchy that can reveal to which stage in the acquisition process each student has progressed. The think-aloud protocols provide data which can be differentiated to give meaningful answers that a quantitative method cannot. A central issue is whether successful task completion is the manifestation of knowledge, qualified lucky guesses, or suppositions. A quantitative measure would ignore this difference.

5. Can the syllabus instruction affect the non-syllabus results?

Notes:

The pre-experimental hypothesis was that the experimental treatment would have no effect on non-syllabus items. The experiment was set up to test the effect of the syllabus instruction in a VISL and a NON-VISL context. The content of the experimental treatment was the syllabus content and the test was constructed to test syllabus items. For comparison, and in order not to alienate subjects, the test also contained non-syllabus items to which the respondents could react intuitively. It was therefore natural to want to follow the main research questions with an investigation into the non-syllabus sides of things. It should be made clear, however, that the number of test items is near the borderline of the required quantity to achieve significance. Intuitive judgements of grammaticality are generally considered to draw on implicit types of knowledge and as such these items could perhaps inform the SLA discussion of the interface between implicit and explicit knowledge. As outlined, this was only a side issue in the experiment, and it must be pointed out that the construction of the experiment should have followed a different pattern if this was to be tested in a reliable fashion. It is important to remember, therefore, that the results relating to this issue can only indicate tendencies.

1.4 The structure of the thesis

The thesis is divided into four parts. Part One, which contains Chapter 1, is constituted by the introduction and the research questions.

Part Two (Chapters 2-4) contains the quantitative section with the relevant reviews of research literature, the experimental structure and methodology, empirical data, their description, interpretation and results. Chapter 2 contains the review of SLA research literature, including a historical perspective and the issues of contention in the field, such as the relationship between explicit and implicit instruction, learning and knowledge. In the light of the research questions and the experimental treatment of this study the main focus is on the role of grammar instruction in SLA. Chapter 3 contains the CALL research review including a historical perspective. The emphasis is on grammar and the needs of the tertiary level. Chapter 4 contains the quantitative study, a description of the experimental treatment, the problems encountered in the execution of the experiment, the empirical findings, the results and their statistical descriptions, and what conclusions it is possible to draw from them.

Part Three (Chapters 5-7) contains the qualitative section with the relevant reviews of research literature, experimental setup and methodology, qualitative data, their description, interpretation and results. Chapter 5 gives a review of research literature on the role of consciousness in implicit and explicit learning and knowledge from a general and theoretical perspective. This supplements the section in Chapter 1 which dealt with the issue from a language learning perspective. It outlines the disagreement among researchers on the definition of central construct, such as consciousness, and whether or not learning without consciousness is possible. The chapter demonstrates how the same terms are sometimes used to designate one quality by some researchers and by other researchers to designate a different quality. Chapter 6 reviews the research literature on introspection, discusses introspection as a research method, discusses the various types of introspection, and gives the rationale for the choice of method adopted in this study. Chapter 7 contains the qualitative study, its methodology, results and interpretations. The chapter outlines the central concepts of Dienes and Perner's theory and its operationalisation. It demonstrates the method of inferences in the analysis of the think-aloud data and outlines the conclusion that can be reached with regard to the acquisition of explicit knowledge of the subject-matter in the experimental treatment. It attempts to demonstrate how this general theory is well-suited to the field of language learning, especially the metalinguistic side of language learning.

Part Four is constituted by Chapter 8, which is the final and conclusive chapter, and contains answers to the research questions and attempts to develop a unified conclusion in which the connection and interaction of the four strands of this study, i.e. SLA, CALL, quantitative and qualitative information come together to give new insight into the matters investigated.

PART TWO

2 SECOND LANGUAGE ACQUISITION: REVIEW

In second language acquisition (SLA) there are several issues which cause disagreement and discussion but few as much as the role of metalinguistic knowledge and instruction in linguistic rules. There are several reasons for this, partly the divergent underlying beliefs as to the nature and origin of language, and partly the difficulty of delivering tangible evidence of theoretical stances due to the multitude of variables which interact with each other and with each individual student. Furthermore, the diversity of the composite nature of the field has meant that SLA needs to draw on research results from a number of independent research paradigms, such as neuroscience, psychology, and sociology, in addition to related research in linguistics, psycholinguistics, and pedagogy. Furthermore, there are issues of reliability and validity concerning how to assess interlanguage developments in individual learners⁷.

R. Ellis (2005) emphasises that despite disagreements on learning processes there is a general agreement on the existence of two different types of knowledge, i.e. implicit knowledge and explicit knowledge: “It points to a common need, irrespective of one’s theory of linguistic knowledge and language learning, for empirical researchers to distinguish whether what individual learners know about a language is represented implicitly or explicitly” (p.143). The present review, which briefly presents the major research results and developments, will concentrate on that distinction as the role of metalinguistic knowledge and instructed grammar is of special interest to the studies in this thesis. In addition to the present chapter on theoretical and empirical studies of this dichotomy in relation to SLA, there will be a more general and theoretical discussion of implicit and explicit learning and knowledge in Chapter 6.

⁷ R. Ellis (2005) states: “Two of the major goals of SLA research are to define and describe second language (L2) linguistic knowledge and to explain how this knowledge develops over time by specifying the external and internal variables involved (R. Ellis, 1994). There is no agreement among SLA researchers regarding the theoretical model that should inform the first of these goals and, I will argue, there has been little real progress in achieving the second goal because of a general failure to address how learners’ L2 knowledge can be measured “ (p. 142).

2.1 Instruction

Instruction and research are different but related in that research results often have implications for the formation of pedagogical approaches to instruction and, conversely, classroom experience may lead to theoretical investigations that lead to new insights (see Widdowson, 1990). The traditional approach was the predominant method in the first half of the 20th century (see Fotos, 2005). The method is also known as the grammar-translation method, and the focus was on explicit explanation of rules (R. Ellis, 1994:569; Fotos, 2005:661). Prior to Chomsky's theories of universal innate structures, L2 instruction was influenced by behaviourist views on learning according to which learning is basically a question of habit formation. This instruction built on imitation and practice (Lightbown & Spada, 1999), and it was developed into the Audiolingual method (VanPatten, 1996).

Chomsky's theory of innate universal structures and Krashen's own theories of language learning (see section 1.1.2.1) led to the development of the Natural Approach in cooperation with Terrell (1983). The main idea is that even adults can 'pick up' a language if only they are exposed to the right input as they, like children, have access to the language acquisition device (LAD). Reading and communication were considered important sources of input whereas the opportunity to produce output, on the other hand, was considered to be of relatively little importance.

The development in second language teaching in the twentieth century was one from almost exclusive focus on form and explicit instruction to a communicative approach with a strong focus on meaning. The results from the Canadian immersion studies have shifted that focus again towards more focus on form (FonF). An increasing inclusion of psychological theories of learning has led to an inclusion of processing theories in teaching and research (VanPatten & Cadierno, 1993).

In a much acclaimed meta-analysis of experimental and quasi-experimental studies into the effectiveness of L2 instruction published between 1980 and 1998, Norris & Ortega (2001) investigated different types of instruction: implicit versus explicit, focus on form

(FonF) versus focus on forms (FonFS) versus focus on meaning (FonM). They also evaluated the reliability of the results of the studies included in the analysis. The latter aspect turned out to be difficult in that many studies (37%) did not report the necessary statistical information for a comparison to be possible. However, with the material available they were able to say that FonF treatments were slightly more effective than FonFS as measured in effect sizes and explicit treatments, but the same measures proved to have “substantially larger effect sizes than implicit treatments” (Norris & Ortega, 2001:178). The hierarchy for effect (ibid.) was as follows:

FonF explicit > FonFS explicit > FonF implicit > FonFS implicit.

However, as they state, the standard deviation in most studies was high, which may indicate heterogeneity in the observed effects. However, the observed effects for superiority of FonF and FonFS and explicit instruction over other types were so large that they are highly trustworthy (Norris & Ortega, 2001:195).

2.2 The role of implicit and explicit knowledge in second language acquisition

2.2.1 Definitions and constructs

The field of second language acquisition abounds with a number of overlapping and loosely applied terms and constructs. Often the term metalinguistic knowledge is used in a general way to cover the same concepts as the term explicit knowledge. In other cases metalinguistic ability or skill is used to refer to “cognitive control” and the ability to “analyse knowledge of language” (Bialystok, 1985:229). Bialystok goes on to concede “...there is no consensus regarding the precise domain of activities which properly may be called metalinguistic” (p.229).

The variance in the definition and application of the terms may be related to the general discourse concerning the role of conscious knowledge in the acquisition process. N. Ellis (1994b) states, “The role of conscious rules in language acquisition remains mysterious not only for want of further empirical investigation. Equally limiting are

fundamental conceptual confusions in the very language which we use to address the question” (p.5).

The term intentional (vs. incidental) learning, for instance, is now the preferred term and construct used by Jan Hulstijn, who has carried out several important studies with regard to grammar and vocabulary acquisition (1994, 1995, 1997, 2002, 2003).

Referring to Schmidt (1994), Hulstijn and de Graaf (1994) define explicit learning as “learning with awareness at the point of learning” (p.97). Schmidt (1994:20) in the same thematic issue of *AILA Review* devotes a full article to the discussion of constructs and terms. His article contains several recommendations to be adopted in applied linguistics, among them that incidental learning should replace the term ‘unconscious learning’, a policy which Hulstijn later adopted, and which he now uses consistently⁸ (see also Chapter 5). Schmidt does raise a warning about the understanding of the term ‘incidental’, namely that it should not be assumed that incidental learning takes place “unaccompanied by attention or awareness or that the knowledge gained cannot be expressed” (Schmidt, 1994:16). It seems that the introduction of the term ‘incidental learning’ may be more a pedagogical construct than one that can help clarify the issues of the role of consciousness in learning, although it has been operationalised as learning with a focus entirely on meaning and implicit learning as a condition where the learners memorise a number of sentences (see Hulstijn, 1989; Rosa and O’Neill, 1999).

The distinction between the learning processes and the product appears unclear.⁹

Bialystok (1994a) makes the connection between analysis of knowledge and explicitness:

The reason that thought evolves, or that language proficiency increases, is that mental representations develop. Analysis is the process by which mental representations that were loosely organized around meaning (knowledge of the world) become rearranged into explicit representations that are organized around formal structures. Relatively unanalyzed representations of language are based

⁸ Hulstijn notes that ‘Only three experimental L2 grammar-learning studies appear to have explicitly used the term “incidental”, [...] none of them pitted incidental against intentional learning’ (Hulstijn, 2003:359). Hulstijn’s 1989 study is one of the three but he states himself that “Theoretically the study is presented as one of implicit learning” (ibid.). The other two studies were carried out by Robinson (1997b).

⁹ Presumably, the resulting knowledge from ‘incidental’ learning would be categorised as implicit.

on the meanings and functions of language without the concern for how those meanings and functions are signified. Conversation is well served by these representations, but grammatical analysis is not. [...]. In a sense, analysis is the process underlying the phenomenological experience that implicit knowledge becomes explicit. In this way, explicitness is really a statement about the level of organization in the mental representation. (Bialystok, 1994a:159)

In contrast, implicit knowledge is often described as unanalysed knowledge, which is nondeclarative, stable and fast, but which cannot be applied in a controlled manner. This interpretation is in line with the theory developed by Dienes and Perner (1999) which is discussed in Chapters 4 and 6 and which posits that explicitness is hierarchical, i.e. explicitness can be observed in varying degrees ranging from full explicitness to no explicitness.

Bialystok (1994b) notes about the difference between implicit and explicit knowledge that explicit knowledge can be learned at any age, that it is language specific, and that “explicit knowledge dynamically develops from implicit knowledge through development” (ibid.:567). Implicit knowledge, she notes, is “the endowment out of which language grows” (1994b:567), but “through analysis some part of that [implicit knowledge] evolves into what we call grammar and permits the accretion of knowledge for the details of each language we know” (ibid.:567). Hulstijn (2002b) defines metalinguistic knowledge as “a kind of declarative knowledge” (p. 205; see also R. Ellis, 2004:236), different from implicit knowledge and residing in a different area of the brain. R. Ellis (2004) contends that “explicit knowledge cannot be defined without reference to implicit knowledge” (p.230), and he states that the main distinction lies in the processes involved with implicit knowledge being available for automatic use in contrast to explicit knowledge which can only be applied through conscious control. Declarability is another distinguishing feature (Hulstijn, 2002a,b) with explicit, metalinguistic knowledge being available for report whereas implicit knowledge is tacit and procedural.

2.2.1.1 Metalanguage

In a comprehensive investigation of explicit knowledge, R. Ellis (2004) equates explicit knowledge with “analyzed knowledge” and “metalanguage” (p.227). The role of

metalinguage is further discussed in relation to verbalisability and the ability to describe language rules and explanations for grammaticality and errors:

Although metalinguage is not an essential component of explicit knowledge, it would seem closely related. It is possible that an increase in the depth of explicit knowledge will occur hand in hand with the acquisition of more metalinguage, if only because access to linguistic labels may help sharpen understanding of linguistic constructs (R. Ellis, 2004:240).

The largest (n=509) investigation of the role of metalinguistic knowledge was carried out by Alderson, Clapham and Steel (1997), comprising students at five British universities. Their study has metalinguage as well as linguistic concepts as a substantial element in their definition and testing of metalinguistic knowledge. Their metalinguistic Assessment test includes identification of parts of speech and linguistic concepts such as subject, predicate, direct object, and indirect object. The study included students of English, French and Linguistics, and the testing was carried out three times over a time span from 1986 to 1994. The rationale behind the study and the argumentation for the importance of metalinguage is that “university methods of teaching foreign languages [...] are still based on the assumption that students have a knowledge about language” (p.94).

Alderson et al. found no convincing evidence for a correlation between metalinguage and language proficiency in their tests, but did find a moderate correlation between the MLAT¹⁰ test results (words in sentences section) and language proficiency (p.116). The students reported that they frequently experienced a need for metalinguistic knowledge but also that the terminology and labels used by teachers were inadequately explained (p.109). Alderson et al. comment that perhaps teacher expectation was that the students knew the label and concept in question.

The roles of metalinguage and grammatical labels have been investigated from a teaching point of view by Borg (1999) in order to throw light on teaching practices and teacher cognition in order better to understand the role of metalinguistic terminology in L2 instruction. The study concludes that more research is necessary since the study was

¹⁰ Modern Language Aptitude Test.

limited to four teachers and the activities involved can best be described as focus on form in grammar lessons. However, the main conclusion is that the application of metalanguage is very individualised and builds on subjective experiences and interaction in the classroom (p.118). Færch (1985) is an eloquent proponent of the use of grammatical vocabulary as he finds it to be “an important heuristic tool” (p.190). The relationship between focus-on-form instruction and the role of metalanguage is the object under investigation by Basturkmen, Loewen and R. Ellis (2002).

Where Borg focussed on the teachers, Basturkmen et al. include the students as well and the relationship to learner uptake. The study was based on recordings of communicative lessons and the focus on form episodes (FFE) were analysed for the effect of learner uptake in pre-emptive and reactive teacher-initiated and student-initiated events. The results showed no significant relationship between metalanguage and uptake when the situation was reactive or teacher-initiated, but interestingly there was a correlation between metalanguage and student-initiated events (p.9). Basturkmen et al. conclude that “metalanguage appears to be an important means through which students can initiate discourse about language forms in the classroom” (p.10). They propose that “The more explicitly a linguistic item is addressed, the more likely students are to notice and incorporate it in their production” (Basturkmen et al., 2002:11), but they also concede that very little research has been done so far in this area, and they recommend that further research is done about the influence of metalanguage in different language learning contexts.

Berry (2000) investigated the form of the metalanguage and the influence it had on learners of English. He operated with three groups which were subjected to the regular, impersonal style of grammatical description, a user-friendly style (Michael Swan, *Practical English Usage* is given as an example), and a hybrid of the two, respectively. The group which was subjected to the hybrid grammatical description did worse than the other two groups, and there is no evidence that user-friendly¹¹ language promotes learning; perhaps it may even have a detrimental effect (p.200). A consistent style is

¹¹ Berry (2000) calls the style ‘youuser-friendly’ because of the style’s frequent use of ‘you’ in active sentences rather than the impersonal style of the traditional style of grammatical description with passive sentences and nominalisation.

preferable to a mixed one (p.205) and “a mixed style may appear easier to read and more appropriate, but this does not guarantee understanding and may actually detract from it” (p.205). The student groups were asked about their attitude to the various styles and the hybrid style was rated positively, but the students did worse on the test than the other two groups. This study was followed up by a more comprehensive study (Berry, 2004) comparing *Collins Cobuild Grammar* and the scholarly *A Comprehensive Grammar of the English Language* by Quirk, Greenbaum, Leech and Svartvik. One conclusion is that ‘friendliness’ is not considered by the students to be “an issue for the genre” (p.13), and students appear to have a higher trust in the correctness of the scholarly grammar. Han and Ellis (1998) investigated the relationship between implicit and explicit knowledge, including metalanguage (in this case the ability to explain verb complementation either of –ing or the infinitive), and introduced a Metalingual Comments score which was incorporated as a measure of explicit knowledge. They found a correlation between metalanguage and explicit, analysed knowledge as measured by a delayed grammaticality judgement test (i.e. judgements made 15 seconds after a sentence had been read), but found that metalanguage is not related to language proficiency. The study was conducted with the overall purpose of finding a way to measure the different types of knowledge and the focus was consequently very narrow on a grammatical item for which it is very difficult to state a definite rule.

2.2.2 Interface positions between implicit and explicit knowledge

The question of whether the application of metalanguage has any bearing on subsequent production and L2 acquisition is closely related to one’s position on whether there is or can be an interface between implicit and explicit knowledge.

2.2.2.1 The no-interface position

The prominent proponent of this position is Krashen (1981, 1985, 1991), who introduced the notion of ‘acquisition’ versus ‘learning’ (1981). Krashen considers ‘acquisition’ to be an unconscious, implicit process in contrast to ‘learning’ which is a conscious, explicit process. In Krashen’s monitor theory the teaching of grammar is of little importance since he holds that explicit knowledge can only be used to monitor output but never enter into the system of proficiency and fluency of production itself:

“Error correction and explicit teaching of rules are not relevant to language acquisition” (1981:1). According to Krashen children utilise basically the same mechanisms for their L1 and the L2 and the same is true for adults, who apply the Language Acquisition Device as outlined by Chomsky’s theory and the principles of Universal Grammar (Chomsky, 1957, 1964, 1965). All that is needed for acquisition is exposure and comprehensible input (Krashen, 1985): “Only comprehensible input is consistently effective in increasing proficiency; more skill-building, more correction, and more output do not consistently result in greater proficiency” (Krashen, 1991:410). For development the input must contain input that the learner is ready to acquire, the so-called *i+1* hypothesis (ibid.). Krashen does concede, though, that adults, given enough time to monitor, may find ‘learned’ knowledge useful in self-correction (Krashen, 1994:46).

However, as Færch and Kasper state, “Positing two types of learning, ‘acquisition’ and ‘learning’, has little explanatory adequacy without some detailed specification of the assumed differences between the two” (1986a:260). Generativists and universal grammar adherers have criticised Krashen for his reference to Chomsky because they see Krashen’s arguments as a misrepresentation of Chomsky’s ideas: “Krashen is confusing two different kinds of rules [universal grammar and pedagogical grammar] and this confusion renders illegitimate his use of Chomsky to endorse his learning/acquisition distinction” (Gregg, 1986:119). Also Lightbown and Pienemann (1993) have voiced their criticism and presented convincing arguments for what they perceived as the weaknesses of the stance against instruction; in particular they took exception to Krashen’s claim that the effects of grammar teaching appear “peripheral and fragile”, a view he maintained in his response (Krashen, 1993:725).

It is remarkable that the criticism comes from Pienemann since his research findings on acquisition sequences (see Chapter 1.4) lend some support to Krashen’s Natural Order Hypothesis (Krashen, 1978a, 1985). Pienemann’s research demonstrated that certain linguistic features are acquired in a certain order (Pienemann, 1987, 1999). Hulstijn and de Graaf (1994) comment that the non-interface position is defensible from a pedagogical position, but “from an empirical perspective it is a fruitless position, as

long as the distinction between acquisition and learning has not been operationalized” (p.99). In a more recent article, however, Hulstijn (2002b) declares himself to be a supporter of the non-interface position (p.208). Unlike Krashen, he does not see this as ruling out the usefulness of explicit knowledge; in fact, he considers “explicit knowledge to be a worthwhile, sometimes, indeed indispensable, form of knowledge to be used as a resource where and when implicit knowledge is not (yet) available” (Hulstijn, 2002b:209). Regardless of the little support he has found, Krashen’s theories have been very influential, but they were and are controversial although many second language acquisition researchers agree that fluency draws on implicit rather than explicit knowledge; the disagreement arises on the issue of knowledge building and what can and cannot enter into the interlanguage grammars of learners.

2.2.2.2 The strong-interface position

According to this position, which builds on cognitive psychology, explicit knowledge can be converted into implicit knowledge through practice and conversely, implicit knowledge can become explicit, declarative knowledge through inferences based on noticing and linguistic awareness. Bialystok’s (1978) theoretical model of second language learning was an expression of this position (1978:71). Anderson’s ACT-R¹² model (1983, 1993) maps out an automatization process in which declarative knowledge gradually turns into procedural knowledge¹³ and controlled processes become automatic (see also Chapter 5). In line with this, DeKeyser (1997) hypothesised that explicitly learned grammar rules become gradually automatized, and subsequently found his hypothesis largely confirmed (p.207). Hulstijn and de Graaf maintain that “For many years, the strong interface position was the accepted view of most practitioners in the field of L2 pedagogy” (1994:99).

2.2.2.3 The weak-interface position

R. Ellis argues the case for a structural syllabus (1993) in support of the weak-interface

¹² Adaptive Control of Thought-Rational.

¹³ In second language acquisition research an unresolved issue is which type of knowledge comes first. Anderson claims that declarative knowledge comes first; Bialystok (1978) claims unanalysed knowledge comes first. R. Ellis (1994) comments: “the claim that language must begin with unanalysed knowledge seems unwarranted in the case of L2 acquisition. Many instructed L2 learners begin with explicit knowledge” (R. Ellis, 1994:358-59).

position in that it can serve as intake facilitation through attention to linguistic properties in the input. He posits that the ‘model’ (i.e. the weak-interface position) shows that “implicit knowledge can be internalized in two ways. The main way is by deriving intake from the input. A secondary way is directly from explicit knowledge that is learned through formal instruction” (R. Ellis, 1993:98). The conditions under which the intake can come directly from explicit knowledge are that the learner is at the developmental stage that makes the process possible, and that the learner is able to ‘notice’ features in the input (for ‘noticing-the-gap’ see Schmidt and Frota, 1986).

Sharwood Smith (1981) and Sharwood Smith and Rutherford (1985) saw consciousness-raising¹⁴ as “a facilitator for the acquisition of linguistic competence” (1985:281); that is, some attention to form could influence competence. Sharwood Smith (2004) has outlined the relation between metalinguistic knowledge and performance as one between universal grammar and a metagrammar. He sees these two as separate but related states that both need to work together in production of L2. “For a metagrammar to operate and indeed for the rawest, simplest kind of metalinguistic awareness (without any knowledge of formal grammar) there still needs to be link of sorts with the core language system” (2004:270). In his discussion he further states that though the two types may be separate they are linked by a system of interfaces (p.275). He disagrees with Krashen’s view that explicit knowledge is of no use in fluent performance in that, according to the theoretical framework he and Truscott (2004) have described,¹⁵ “metalinguistic knowledge is as open to automatization as any other domain of knowledge so it therefore makes sense to talk of metalinguistic or metagrammatical fluency” (Sharwood Smith, 2004:276). Thus adults will have two sources available for L2 processing and fluency.

2.3 Immersion

The largest experiment with focus on meaning and communication has been the

¹⁴ Sharwood Smith (1991) recommended that the term ‘consciousness-raising’ be replaced by ‘input enhancement’ as there was no certain way of knowing the cognitive effects of making input more salient to the learner.

¹⁵ MOGUL or Modular On-line Growth and Use of Language.

French immersion classes¹⁶ in Canada (see for instance Harley, 1989; Swain, 1985, 1998; Swain & Lapkin, 1986; Lightbown & Spada, 1990). Students in immersion programmes quickly gain communicative abilities and communicative confidence, but accuracy levels do not proceed to the same degree: “while children learn to speak French fluently and confidently, their accuracy in French syntax and morphology is still below what one might expect of learners who have spent several years immersed in a second language” (Lightbown & Spada, 1990:431).

Swain and Carroll (1987) found through analysis of observational data that grammar was taught at particular times separate from meaning. The content-based instruction was not planned in such a way that would naturally lead to a focus on form. The teaching concentrated on the subject-matter and the linguistic information was not naturally attended to nor: “did these methodologies permit learners to learn about structural features of form” (Segalowitz & Lightbown (1999). Lyster & Ranta (1997) investigated the uptake of six different types of corrective feedback in four immersion classrooms at primary level. They found that teachers use recasts more than any of the other types despite these being less effective. Their results showed that uptake was greatest if the feedback led to negotiation of form, i.e. elicitation, metalinguistic feedback, clarification requests, or teacher repetition of error followed by peer or self-repair (p. 58).

Lightbown (2001) in her survey of input filters states that the ability to attend to meaning as well as form is subject to individual differences and that some learners have the ability to shift quickly between form and meaning and still keep up with the ongoing lecture or lesson but that not all learners have this ability. Lightbown concludes that learners in communicative and content-based instruction classes need more guidance than they are getting “if they are to overcome the effects of the various filters¹⁷ which prevent them from noticing input when their attention is focused on meaning in oral interaction or in reading” (2001:93).

¹⁶ About three hundred thousand students are enrolled in these programmes. The students are either Anglophones enrolled in French immersion classes (e.g. in Montreal) or Francophones enrolled in English immersion classes (e.g. in Vancouver) (Kowal & Swain, 1997).

¹⁷ Lightbown discussed affective filters, auditory/phonological filters, cognitive filters, developmental filters, and L1 filters.

One possible L1 interference or transfer problem between English and French concerns the placement of adverbs in sentences, and this particular issue was investigated by White (1991) in a pre-test, post-test, delayed post-test design with francophone learners of English. The placement of adverbs between verb and object is not allowed in English but is a correct placement in French. One group was given explicit instruction on adverb placement and a second group was given instruction in question formation plus a flood of sentences with adverbs placed in the correct position. Only the students in the group receiving explicit instruction on adverb placement understood that the SVAO was not a possibility in English. The instruction took place in an intensive programme¹⁸ but for a limited period in all they received five hours of instruction in adverb placement over two weeks. A follow-up test after a year showed no lasting effect of the instruction, and thus no difference between the two groups, which may be due to the short period of instruction. In a related study White, Spada, Lightbown, and Ranta (1991) investigated question formation with a control group instructed in adverb placement. The study found that the instruction on question formation had an effect on syntactic accuracy in that the students in the instructed groups appropriately used inversion in questions to a much higher degree than the uninstructed students; however, the instructed students were still far from native-speaker accuracy (p.424).

Adverb placement was further investigated by Trahey and White (1993), whose study on the effect of an input flood consisting of huge numbers of sentences with correct adverb placement showed that the students had not learned to place adverbs correctly. Lightbown concludes (2000) that corrective feedback and explicit instruction have been demonstrated to be necessary for structural accuracy.

The slightly disappointing results from the communicative-based teaching in immersion classes as far as correctness levels and non-target-like structures in production are concerned, combined with the results of studies which showed that corrective feedback and explicit instruction had an effect even though the main focus was still on meaning, has

¹⁸ For five months of the year the students in these programmes receive five hours of instruction in English as a second language for five days a week. The instruction is communicative-based.

led to renewed interest in focus on form. In her 1998 article Swain sums up the situation:

This research related to the French proficiency of immersion students, makes clear that an input-rich, communicatively oriented classroom does not provide all that is needed for the development of target-like proficiency (Swain, 1985). It also makes clear that teaching grammar lessons out of context, as paradigms to be rehearsed and memorized, is also insufficient (Swain, 1998:65).

She further concludes that few teachers and researchers involved in immersion would disagree with the view that more attention to grammatical accuracy is needed.

According to Swain (2005) it is conscious reflection and the ensuing deeper processing which will lead to successfully modified output.

2.4 Focus on form

Long & Robinson (1998; see also Long, 1991) draw up a contrast between what is known as focus on forms (FonFS) and focus on form (FonF). The FonFS approach they describe as synthetic syllabi; that is, different parts of the language are presented to the learner who must synthesise its different and discrete pieces for use in production¹⁹. In contrast, focus on form is based on interaction between learners, between more proficient and less proficient speakers, between texts and the learners, i.e. the focus on form takes place in a meaningful context and Cand. Negotiation of meaning is central. The focus on form takes place when the need or opportunity arises and, in contrast to FonFS, there is no pre-planned linguistic focus. The pedagogical planning is task-based and “exploit[s] opportunities that arise naturally from the interaction of learners and tasks” (Long and Robinson, 1998:23). Task-based instruction attempts to create authentic tasks in which genuine communication will take place and through the noticing of gaps in their own interlanguage language or through negotiation of meaning, attention will be directed to linguistic features in the input, subsequently leading to interlanguage development. Pedagogically there are several ways of instructionally focusing on form but central to them all is the notion of attention.

¹⁹ Grammar-translation and the Audiolingual method fall under this category.

2.4.1 Attention, noticing, and awareness

The relation between the constructs is described by Schmidt (1995b) in this way:

A low level of awareness, called here “noticing” is nearly isomorphic with attention, and seems to be associated with all learning. A higher level of awareness (“understanding”) is involved in contrasts between explicit learning (learning on the basis of conscious knowledge, insights, and hypotheses) and implicit learning (learning based on unconscious processes of generalization and abstraction (1995b:1).

The role of attention and noticing has been under investigation since Schmidt’s influential article (1990) on noticing, in which he observed that noticing is the necessary and sufficient condition for learning. Schmidt defined the operational level of noticing as being available for report.

The constructs have been operationalised in a variety of ways with the general purpose of understanding the role of awareness in the learning processes and ultimately in the effects of instruction. The discussion on attention was sparked by Sharwood Smith and his focus on ‘consciousness-raising’ (1981; Rutherford & Sharwood Smith, 1985). Rutherford & Sharwood Smith defined consciousness-raising as “the deliberate attempt to draw the learner’s attention specially to the formal properties of the target language” (1985:274). Their thesis was that the creation of a ‘natural’ environment in the classroom communication situation would not be enough and it would be a mistake to disregard the importance of attention to grammar and the formal aspects of language. They emphasised explicitness, but simultaneously pointed out that there are degrees of explicitness which should be adapted to the level of the learner. They warned against regarding attention to form and the communicative aspects of language use as opposites that preclude each other.

Rutherford and Sharwood Smith’s framework is that of universal grammar and the attention to form, they claim, is necessary for the language-specific parameter setting. They promote the view that “contrastive linguistics can be taken seriously and that it has an important role to play not only in L2 acquisition research but in language pedagogy as well” (1985:280). They regard consciousness-raising as “a facilitator for the acquisition of linguistic competence” (ibid.). They do recognise that communicative

interaction is a central part of L2 acquisition but not sufficient, and they see consciousness raising as providing the necessary attention to form.

Sharwood Smith elaborated on the role of crosslinguistic, or transfer, influences in his 1996 article, not by promoting any particular theory, but by establishing the area as worth looking into, but also taking the position that any theory needs to discard simple behaviourist assumptions, and by accepting that the creation of the L2 in the learner's head makes use of UG constraints as well as processing input of the L2. Sharwood Smith sees a connection between crosslinguistic influences and noticing (2000b). The logical problem of language acquisition, also known as Plato's problem, was dealt with by Chomsky (1965, 1981, 2000), and the puzzle is how the poverty of input can lead to acquisition; this idea is reversed by Sharwood Smith who sees that ample input and information are being ignored or not noticed by the L2 learner; therefore the question is how we can make the learner notice or utilise the information available through making it more salient.

The 'input enhancement' approach, which Sharwood Smith (1991) advocated should replace consciousness-raising, is one way of making input more salient to the learner. If learners notice formal aspects of the language, they may be in the process of adding to or restructuring their underlying linguistic competence (Sharwood Smith, 1993:176), but would directly enter into the (inter-)language system of the learner. In fact the role of noticing constitutes an area which needs to be further investigated empirically:

One major challenge for second language research is to work out when and how on-line performance processes actually trigger changes in the underlying system, changes that survive when the learner goes off-line. Part of this has to do with attention, that is, "noticing" particular relevant features in the input. (Sharwood Smith, 2000:30).

In a study of input enhancement and rule representation, Alanen (1995) investigated the learning of semi-artificial Finnish under different conditions. The target structures were locative suffixes and consonant alternation. There were three treatment groups and one control group. One group received enhanced input, i.e. the target structures were given in italics, one group received explicit rule presentation and one group received a

combination of enhancement and rule. The study was supplemented by a think-aloud procedure during the study phase and this showed that the learner's focus of attention played a role for acquisition of the target structures; that is, those students who have paid attention to the structures all learned something regardless of treatment. Alanen concludes that noticing the learning structures seems to have been "sufficient for at least some learning to take place" (1995:294) and further that "The present study found little evidence to support a claim that there can be learning without noticing" (ibid.). Interestingly, the study also found that the explicit knowledge developed by the learners and which they could verbalise in statements of rules was not always present in their performance. It became clear that students sometimes have their own agenda, which means that students do not always behave as expected so the experimental groups do not necessarily correspond to learner behaviour. For instance students may have special interests or motivations which will make them pay attention to form even though they were asked to concentrate on meaning. In the Alanen (1995) study the think-aloud procedure revealed this and allowed for the conclusion that the strategy employed by the students did affect results, albeit not always as foreseen by the experimental group to which they had been assigned.

As far as the studies are comparable, the results from the Alanen study were supported by Gass, Svetics & Lemelin (2003). In their study of learners of Italian they operated with two conditions, namely focused attention and non-focused attention, and three linguistic areas, namely syntax, morphosyntax²⁰ and lexicon. Their hypothesis was that students in the group of plus focused attention would outperform students in the non-focused attention group on lexicon but not on syntax. The opposite was the case. The reason for assuming that attention would have the greatest effect on lexicon was that "attention is a limited-capacity processing system and that because one can only attend to a limited amount of material at a time, those learning a language will pay more attention to some aspects than to others" (p.107). Items of lexicon are easy to isolate and thus easy to focus attention on. Syntax, it was presumed, would be more abstract and complex and therefore not the obvious choice for focal attention. Following these predictions, it was also assumed that the effect of attention on proficiency would be

²⁰ Agreement between indirect and direct object.

greatest for lexicon and least for syntax. The underlying beliefs of the whole study were:

...that attention is relevant in all instances of learning, but that it is not a sufficient condition for learning. This is so because complex rules are not immediately apparent from the input data; not only does one have to "notice" the rules, but one also has to understand them in the sense of figuring them out. (Gass, et al. 2003:109)

The results showed that attention results in learning. This was true for syntax, morphosyntax and lexicon, but the greatest effect was for syntax and morphosyntax, and it was a substantial effect. In the non-focused condition there was little effect on learning in general but most on lexicon and least on syntax and morphosyntax. The results of the impact on proficiency showed that the focused attention condition also resulted in the greatest effect on proficiency for all three structures as far as first-year learners are concerned. For second-year learners there was an effect for lexicon but not for syntax and morphosyntax. For third-year learners there were no significant effects, but the effects on syntax were the greatest and close to a significant level. The non-focused attention condition showed no effect from pre-test to post-test, except for first-year learners and lexicon.

That the results were in contrast to expectations with regard to the effects of attention on complex structures such as syntax and morphosyntax the authors explain with a comment which is very similar to the one proffered by Alanen: "One cannot rule out the possibility that learners, because of their own individual needs and interests, paid attention to something despite the fact that we did not experimentally draw their attention to that something" (Gass et al., 2003:526). They conclude, though, that "focused attention is better utilized in more complex areas" (p.527). Since syntax and morphosyntax are such complex structures, it may be necessary to direct the attention of learners to these structures as the task demands are so heavy that learners may not be able to do this efficiently of their own accord; consequently, learners can be aided to become more efficient learners through pedagogical planning of tasks.

The findings of the Gass et al. study (2003) study are in accordance with Hulstijn and de Graaf (1994) in indicating that for learning complex structures what Gass et al. call “internal devices” (p.528) are not enough, and both the Hulstijn & de Graaf study and the Gass et al. study seemed to indicate that complex structures need particular attention (see also Chapter5). Syntax is one such complex structure, and Gass et al. suggest that syntax learning may be different from lexicon learning and perhaps even from morphosyntax. With regard to proficiency, Gass et al. suggest that the benefits of focused attention may be related to the issue of ‘readiness’ of the learner, and especially high-level learners may be ready to learn the complex structures, and conversely, this ‘readiness’ makes focused attention the appropriate learning strategy.

In his study (1990) of how students learn under different processing conditions, VanPatten found that it is difficult to attend to form and meaning simultaneously. VanPatten and Cadierno (1993) in a follow-up study investigated ideas developed on ‘input processing’, which constitutes an attempt to pre-emptively get the learners to understand the significance of the metalinguistic information and thus make them capable of incorporating the input into their knowledge of the language. The study comprised three experimental groups, and the topic of instruction was object pronouns in Spanish; group one received no instruction, i.e. the focus was on communication according to the Natural Approach; group two received ‘input processing’ instructions, i.e. the subjects were presented with information on objects in an OVS structure, nouns as well as pronouns, subject information as well as object information and additional instruction in important points to remember about the structure, including contrasting English and Spanish; group three received ‘traditional’ grammar instruction, i.e. explanation of the form and position of direct object pronouns in Spanish. The traditional group’s instruction was followed by practice in producing sentences with these pronouns. The practice for the input processing group was not production but interpretation of the sentences in order to create form-meaning connections. Results showed that processing instruction did have an impact on the way the subjects processed input and subsequently on the learner’s interlanguage system, and it did have an impact on production. Over time the traditional group differed little from the no-instruction group in interpretation tasks, but both groups did improve on production

tasks, and the conclusion is that traditional instruction does not add to the learner's language but only to the learned language knowledge in accordance with Krashen's distinction between learned and acquired knowledge. The contention is that instruction according to the input-processing method adds to the underlying competence and can transfer to new items, judged by the fact that the processing group could produce items not included in their instruction.

In continuation of VanPatten and Cadierno, similar research was carried out by Cadierno (1995) on the acquisition of past tense morphology in Spanish. The main results were that the traditional groups and the input processing group achieved results at equal levels, but the difference was that the traditional instruction group acquired metalinguistic knowledge whereas the group receiving structured processing input was able to improve their proficiency through restructuring their mental grammars. Both these groups were significantly better than the non-instruction group.

2.5. Teachability

A major difference between L1 and L2 learning is the measure of success that can be expected. In L1 success is more or less a given thing in contrast to L2. Despite ample input in the form of exposure and instruction, L2 acquisition is fraught with difficulties. The learner has to be cognitively ready in order to be receptive to new structures in the input. With regard to teaching, it seems that the complexity of structures is of relevance where the effectiveness of instruction is concerned.

2.5.1 Sequences of learning

Some aspects of language are acquired in the same sequence by all learners regardless of their L1 and individual intellectual ability and diligence, while other parts of language are acquired depending on individual ability. The former are referred to as developmental features, while the latter are referred to as variational features (Sharwood Smith, 2000b). Naturalistic L2 acquisition follows particular developmental sequences (Wode, 1976). Pienemann's research indicated that the same was the case for instructed L2 acquisition. Results indicated that instruction in developmental features could not

alter the sequence of acquisition but could affect the rate of acquisition. The developmental features which are subject to particular sequences of acquisition are word order and grammatical morphemes.

In 1981 Meisel, Clahsen, and Pienemann developed the multidimensional model of second language acquisition based on research into German and English (see also Clahsen, Meisel, & Pienemann, 1983²¹) as a second language, establishing the difference between features which are subject to individual differences (variational) and those which are not (developmental). Clahsen et al. identified 14 features for which a general characteristic is that the direction of acquisition is from simplified towards norm-oriented (see Pienemann, Johnston & Bindley, 1988, Figure 1, p. 223). “The basic line of argument in this explanatory approach is that the psychological complexity of a structure is dependent on the degree of reordering and rearrangement of linguistic material in the process of mapping underlying semantics onto surface forms” (Pienemann et al., 1988b:223). Pienemann & Johnston (1987) found that even when a feature was present in a learner’s L1 it may not initially be available to the L2 learner, and “initially learners organize their interlanguage around nonlinguistic processing devices and gradually build up language specific and target-language-specific processing devices” (Pienemann et al., 1988b:224).

On the basis of the ZISA research results Pienemann developed the Teachability Hypothesis (1986) which holds that developmental features are acquired in an invariant sequence immune to external influences such as teaching. However, it “does not imply that learning is *guaranteed* by the mechanisms internal to the learner. It also does not imply that teaching has no influence on SLA whatsoever” (Pienemann et al., 1988:226. Italics in original.). The morphological and syntactic features appear to be acquired in five stages according to the following order:

Stage 1: single words, formulae

Stage 2: SVO, plural marking

Stage 3: *Do* fronting, Topicalization, Adverb Preposing, Neg+V

²¹ This is often referred to as the results of the ZISA (Zweitsprachenerwerb Italienischer und Spanischer Arbeiter) group.

Stage 4: Pseudo-Inversion, Yes/No-Inversion

Stage 5: 3rd-Sgl-S, Aux-2nd, Do-2nd, ²²

(Pienemann et. al., 1988:228).

In detail, the developmental sequence of structures for English as a second language is as follows (see also Pienemann & Johnston, 1987):

Structure	Example
1. single words, formulae	How are you?
2. SVO, SVO?	*The tea is hot?
3. ADVERB PREPOSING	* <i>Yesterday</i> I work
4. DO FRONTING	*Do he work?
5. TOPICALIZATION	<i>This</i> I like
6. NEG + V (don't)	*He <i>don't</i> eat meat
7. PSEUDO-INVERSION	Where is my purse?
8. YES/NO INVERSION	* <i>Have</i> he seen it?
9. PARTICLE SHIFT	*He <i>turn</i> the radio <i>on</i>
10. V –“TO” –V	We like <i>to</i> sing
11. 3RD-SG-S	She <i>comes</i> home
12. DO-2ND	They <i>did</i> not buy anything
13. AUX-2ND	Where <i>has</i> he seen you?
14. ADV-LY	They spoke <i>gently</i>
15. Q-TAG	It's expensive, <i>isn't it?</i>
16. ADV-VP	He has <i>often</i> heard this

(Pienemann et al., 1988:226).

The sequence of negation placement was studied by Hyltenstam (1978), who for Swedish found sequences supporting Pienemann's findings. The sequences for placement of negations in Swedish were found to be invariant and to follow the same order regardless of the learners' L1:

Stage 1: neg V

²² This implies Wh-inversion and internal negation.

Stage 2: (AUX) negation V

Stage 3: V negation

Stage 4: sub clause: negation V

Stage 5: negation AUX V

Pienemann (1998) comments on Hyltenstam's results: "Hyltenstam's findings support our predictions extremely well. The only proviso on this is that currently the Swedish processability hierarchy does not differentiate between all the stages he found. Instead it conflates some of them into one" (Pienemann, 1998:205).

Karen Lund's longitudinal study of four adult learners of Danish as a second language concerned inversion and negation. She found some of the same stages as Pienemann and Hyltenstam but found that functional and semantic considerations were influential in the acquisition process whereas Pienemann's research was exclusively structurally based. Lund found that learners acquire inversion in questions before inversion in declarative sentences. It seems that salience of the features plays a role in the acquisition process. As she points out there is no semantic difference between *Igår kom han hjem klokken 6* [Yesterday came he home at 6] and *Igår han kom hjem klokken 6* [Yesterday he came home at 6] and therefore the function of inversion in declarative sentences is not salient to the learner in the same way that inversion in questions is (Lund, 1997:4).

With regard to negation she has found that salience plays a role that has been overlooked by other studies; in fact she finds that variation in acquisition depends on the learners' L1, L2 or the relationship between them, and that this would explain some of the variation found in previous studies which were more or less ignored. Specifically, she raises the question of whether the acquisition of negator placement is related to the application of auxiliaries rather than the acquisition of negation (ibid., p.5). She concludes that the learner goes through the following stages (ibid.):

1. how to use the auxiliary system
2. how to use the correct form of the negator (no/not/don't)
3. how to conjugate verbs

Lund's stages would explain the stages given by Larsen-Freeman and Long (1991:4):

<i>Stage</i>	<i>Sample utterance</i>
1. External	No this one/No you playing here
2. Internal, pre-verbal	Juana no/don't have job
3. AUX+neg	I can't play the guitar
4. Analysed don't	She doesn't drink alcohol

Lund suggests that “the negator is always placed in front of the semantically salient verb, whether that verb is finite or non-finite” (ibid., p.6), and for Danish the most difficult feature will be placement of the negator after the finite verb in main clauses (e.g. *Jeg svømmer ikke* [I swim not]) and before the auxiliary verb in subordinate clauses (e.g. *Hvis jeg ikke kan svømme* [If I not can swim]) because this presupposes that the learner is able to distinguish finiteness from non-finiteness and subclauses from main clauses (see also Lund, 1998). Thus, she has found that some learners add a step, some learners skip a step (depending on their L1 and the nature of the target language), and furthermore, she has demonstrated that the semantics and the salience of the target structures play a role. Pienemann (1998:74) does not disregard the influence of L1 and L2 (for an extended exposition see also Pienemann, 1998, 1999), but his contention is that learners of a given L2 will pass through the same stages.

Sharwood Smith & Truscott (2005), in their discussion of stages versus continua in the acquisition process, ask: “Should evidence of variation and instability in learner behaviour be treated as noise in the data or short-term phenomena of no theoretical interest?” (p.219). This question, of course, is also related to the issue of when a structure has been acquired. Like Pienemann (e.g. Pienemann et al., 1988:223), Sharwood Smith and Truscott regard the moment of emergence as the moment of acquisition. From the first appearance it may be a while before mastery is achieved, if ever. They suggest that previously the issue of stages has given rise to confusion, and they attempt to find a solution which could reconcile the stages with the continuum:

“Development proceeds in steps but it also proceeds gradually, suggesting a continuum. Development moves the learner from one structural solution to another but development also allows different solutions to co-exist” (Sharwood Smith & Truscott, 2005:221). Thus Sharwood Smith and Truscott try to establish a model of acquisition in which frequency plays a role for acquisition, in that the number of ‘recognized’ structures (p.234) in the input will raise the activation level and thus make it more likely for these structures to be present in the production of the learner. The idea of a threshold for acquisition, in their opinion, is an arbitrary boundary; for instance, is a structure acquired when the structure in question is used correctly 50% of the time, 60% of the time, etc.? (pp.235, 236). They also make a point of clarifying that the stages or thresholds they talk about are internal (individual), not subject to externally observable production of features (p.235).

The question of whether syntax can be taught was discussed by R. Ellis (1984). His initial observations of the children in his study was that two children had begun to apply interrogative WH utterances but most of the time they relied on routines (i.e. formulae), intonation or inverted yes/no questions, which is quite in accordance with Pienemann et al.’s stages (for an overview of negation and inversion see Lightbown and Spada, 2003:77-80). Half the interrogative utterances of these two children (R and T) contained no verb and only R used subject-verb inversion. This illustrates the points made by Sharwood Smith and Truscott (2005) that several structures co-exist and to set a threshold (or determine a stage) would seem arbitrary. In this experiment the formal instruction for three lessons and included 13 children. The first lesson taught them the meaning and application of ‘who’, ‘what’, ‘where’, ‘and ‘when’ and lessons two and three taught them subject-verb inversion equal to stage four of Pienemann et al.’s sequences.

R. Ellis found that there was a correlation between the application of subject-verb inversion and the salience of the pronoun, with ‘what’, ‘who’, ‘where’ and ‘when’ as the sequence of accuracy. This supports Lund’s finding of salience of structure as being of importance in sequence of acquisition. Overall the result of the experiment was, though, that there was no significant increase in the children’s correct application of

subject-inversion in WH questions, and R. Ellis concludes that the route of acquisition has not been influenced by the instruction (R. Ellis, 1984:146); with regard to the rate of acquisition the issue is more complex. “It has been shown that some of the children made conspicuous improvement, but this improvement does not appear to be related to the amount of direct teaching specifically addressed at the children” (R. Ellis, 1984:146). It is also evident that practice has no relation to development, and this gives rise to various speculations by R. Ellis. One question he poses is: why is it that the children who had least opportunity for practice were those who progressed the most? One might comment that in fact the unrelatedness of amount of instruction and practice to the acquisition is foreseen by Sharwood Smith and Truscott’s theories of activation levels, continua and co-existence of structures. This would explain why some students make spectacular progress while others do not, and it appears to be in accordance with their point that thresholds are internal and individual, i.e. cognitive readiness (Sharwood Smith, 2000) is a factor to be considered.

2.5.2 Complexity of rules

As seen in the Alanen (1995) and the Gass, Svetics, & Lemelin (2003) studies, the complexity of the grammatical structures to be learned plays a role for learnability. Several other studies have found support for similar positions. This was evident in the Hulstijn and de Graaf study (1994) in which they examined nine hypotheses about the conditions under which explicit grammar instruction would facilitate the acquisition of implicit knowledge. They hypothesised that the advantage of explicit instruction would be greatest for complex rules with large scope and high reliability and in cases of rule learning rather than item learning. They further hypothesised that instruction would have a greater effect on comprehension than on production, and this would be true for grammatical features which carried semantic content over redundant grammatical features (see also VanPatten, 1994). For production the hypothesis was the reverse, namely that instruction would be more beneficial for redundant grammatical features than for grammatical features with semantic content. In order to control variables and concurrently have a ‘natural’ setting they recommend a ‘twin’ approach of computer controlled language learning of a semi-artificial language (a modified Esperanto) combined with computer controlled instruction of Spanish; that is to say, both types of

instruction take place in a laboratory setting. A number of the hypotheses are included in de Graff 1997b with a detailed analysis of the results whose full context is given in de Graaff 1997a. The hypotheses are tested in two groups of subjects under different conditions: one group received explicit explanation of the target structures and one group received exposure and the opportunity to practise. The target structures in the study combined complexity and morphology/syntax, and the structures were: 1. plural noun form, 2. inflection of imperative mode, 3. position of negation forms, and 4. position of object.

In his eXperanto study de Graaff (1997b)²³ had found a significant effect for explicit instruction and complexity with respect to complex syntactic structure but not for complex morphological structures. The effect of explicit instruction was greatest for what DeKeyser (1995) terms ‘categorical’ rules, which always apply and thus have high reliability, unlike ‘prototypical’ rules which apply no more than 90% of the time²⁴ and whose reliability is much lower. The lack of effect of explicit instruction on complex morphological structures, de Graaff explains as a possible lack of salience of the items concerned as two of the items were distinguished by redundant markers which carried no distinguishing semantic function. This would be in line with R. Ellis’s (1984) suppositions discussed above in relation to question formation and inversion.

The results of the experiments involving explicit instruction in a computer-controlled setting based on the eXperanto programme²⁵ and computer-controlled Spanish were in part outlined in de Graaff (1997b) and fully expanded in de Graaff’s Ph.D. thesis (1997a). The eXperanto section of the experiment confirmed first of all that explicit instruction had an effect; secondly, that explicit instruction has a greater effect on complex syntactical structures, but not on complex morphological structures. The overall effect of explicit instruction on morphological and syntactical structures,

²³ de Graaff 1997a is the ph.d. dissertation on which de Graaff 1997b is founded and thus the semi-artificial eXperanto programme is the basis of both 1997a and 1997b.

²⁴ Hulstijn & de Graaff (1994:103) give the following measures of ‘high’ and ‘low’ categories of scope and reliability: “the scope of a rule is said to be high or low when the rule covers more or fewer than 50 cases; the reliability of a rule is said to be high or low when the rule applies in more than 90% of all cases”.

²⁵ The computer application was based on Esperanto, a constructed language, which was modified to fit the same target structures in Spanish.

however, showed no greater effect on syntactic features compared to morphological features due to the high effect it had on the simple morphological features. Interestingly, the study also investigated the effect of explicit instruction on testing with and without time pressure, and it turned out that there was no effect for time pressure on the test results (sentence judgement). The explicit instruction groups did better under both conditions. It was hypothesised that, given time to monitor, the results would be significantly higher without time pressure (de Graaf, 1997a:95).

It is often assumed that judgement tests without time pressure will tap into explicit knowledge whereas judgement tests with time pressure is thought to tap into implicit knowledge (R. Ellis, 2005). The Spanish section of de Graaff's 1997a study was based on the same structures and the same distribution of them as the eXperanto section of the experiment. This study also supported the effect of explicit instruction in general and for the acquisition of complex structures, morphological as well as syntactical. With regard to the effect on the acquisition of morphological in comparison to syntactical structures, it appeared that the effect was greater for morphological structures. In parallel to the eXperanto experiment there was no effect on the judgement test for time pressure or the absence of time pressure, with the explicit instruction groups doing better under both conditions. In both the eXperanto and the Spanish experiment it was investigated if the two instructional conditions would lead to explicit knowledge. This was operationalised as the ability to verbalise the acquired knowledge.

The overall result was that the explicit instruction groups in both the eXperanto (73%) and the Spanish (70%) sections were able to verbalise their knowledge, which is to say that the explicit instruction had resulted in explicit knowledge on both the syntactical and the morphological structures, with the lowest correlation for the Spanish complex morphological structures (50%). For the non-explicit condition the explicit knowledge of syntactic structures was almost the same as for the explicit instruction group, but the knowledge on morphological structures was lower. The conclusion reached by de Graaff is that for syntactical structures it is possible to build up explicit knowledge without explicit instruction but that this is not the case for morphological structures (de Graaff, 1997a:159). In the eXperanto experiment there was a high correlation between

accuracy and explicit knowledge (i.e. the ability to verbalise the knowledge) for both the explicit instruction group and the non-explicit group. But the lowest correlation was for syntactical structures in the non-explicit condition.

In the Spanish experiment the results were more diverse. Explicit knowledge was correlated to accuracy for the complex structures to a higher degree than for the simple structures in the explicit instruction group. However, it turned out that nearly all subjects could verbalise rules on both complex and simple structures, but for the simple structures verbalisation did not always manifest itself in accuracy in production. For the non-explicit Spanish group, explicit knowledge was highly correlated to accuracy (de Graaff, 1997a:160) for simple structures but not for complex structures for which the procedural knowledge was higher than their declarative knowledge. The general conclusion is:

...it was found that the ability to verbalize explicit knowledge relates to the accuracy as measured by the accuracy tests. This suggests that explicit knowledge is facilitative for the acquisition of the target structures. However, it was also possible to reach a certain level of accuracy without explicit knowledge.” (de Graaff, 1997a:160)

The study also led to the conclusion that explicit instruction may be most useful when target structures are not immediately meaningful or salient to the learner (ibid., p.166).

The idea that explicit instruction is more useful for some structures than for others was investigated and discussed by Hulstijn (1995) on the basis of Dutch as a second language, and in his paper he gives a rationale for grammar teaching. In line with suggestions by de Graaff (1997ab) and Hulstijn and de Graaff (1994) and VanPatten (1994) and VanPatten and Cadierno (1993), explicit instruction may not immediately transform into improved proficiency, but:

- it helps the learner to bring order to the input
- it facilitates the understanding of the (written or spoken) input
- it therefore may boost or support the ‘natural acquisition process’ (i.e. the development of implicit knowledge)

(Hulstijn, 1995:383)

A point made by Hulstijn is that, since not all grammar rules are the same, it is the role of the teacher to evaluate the structures in relation to their teachability and thus determine the appropriate approach. VanPatten (1993) made the point: “It is our responsibility as instructors to ensure that the various aspects of language teaching blend together as some harmonious whole” (p.435).

The discussion of explicitly teaching pedagogical rules was continued by Dietz (2002) who focuses on what a 'rule' is, and how to reduce complexity through various strategies. First of all, he points out that SLA literature has tended to equate 'complex' and 'difficult' and suggests that a distinction should be made between structurally complex and psychologically complex, i.e. difficult to understand or explain (p.264; see also Hulstijn and de Graaff, 1994:103). A case in point could be the input processing instruction advocated by VanPatten and Cadierno (1993), which proved to help learners cope with complex pedagogical rules. Dietz (2002) suggests that known problem areas are re-analysed from this perspective in order to reach new conclusions in relation to instruction.

2.6 Summary

Instruction in grammar and its effectiveness is subject to continual debate in the field of SLA and only gradually is the evidence from empirical research beginning to yield some answers. The question is not only whether instruction should be given, but also under what circumstances and in what form. The number of variables is high and, in addition, most variables are very difficult to control. Consequently, some researchers have chosen to operate in computer-controlled settings with semi-artificial languages.

Results from classroom settings as well as computer-controlled settings appear to point in the same direction, that is, towards more weight to some sort of form-focused instruction. Explicit instruction has proven highly effective, but the focus-on-form (FonF) approach in which formal instruction is given in a meaningful context appears to

give better results than focus-on-formS (FoFS), which is formal instruction out of meaningful context. Least effective is the implicit type of instruction.

The issue of effectiveness, however, is related to the structures in question and the stage reached by the learner. Instruction appears to be more effective for some structures than for others. The learning of complex structures is a case in point. For a given L2 the learning of structures seems to come in a certain order or sequence which all learners follow, and instruction can only affect the rate of learning, not the sequence, although some aspects may not be completely resolved, and there may be variation with regard to the various target languages. The interaction between the many variables - some controllable, some uncontrollable - in combination with many composite strands of research paradigms adds to the complexity and opaqueness of this research area, which does not proceed in linear progression.

3 COMPUTER ASSISTED LANGUAGE LEARNING: REVIEW

3.1 History of CALL

The introduction of the personal computer in the early 1980s meant a surge of initiatives in computer assisted instruction including language learning. Initially, this field was very much technology driven and its proponents technology enthusiasts, many of whom were programmers as well as or perhaps more than educators. New developments in technology since the early beginnings have had a tendency to foster research projects with a focus on that particular technology rather than the learning or pedagogical implications, but acceptance of the importance of the integration of technology, learning and pedagogy is now pervasive, although CALL material continues to differ with respect to which of the pillars of the triad is in focus.

With e-mail and the internet came tandem learning²⁶ and MOOs²⁷, chat, concordancing and corpus work, and distance learning. The CD-ROM made multimedia CALL possible before fast, powerful computers made it possible to download video-clips and films. Speech recognition technology found its uses in learning tools for the blind. Most technological advances were developed outside Europe, but the EU and European educators were in the forefront from the start in their appreciation of the possibilities of these new tools.

The EU funded large pan-European research projects and other development projects to further the learning of European languages and intercultural understanding. The founding of the EUROCALL²⁸ organisation in 1993 was a natural development in the light of the continued dissemination of technology and the improvements in applications. An already existing fledgling organisation was, with a financial support grant from the European Commission, turned into a viable professional organisation. This meant in return that research projects in the field found a supportive forum

²⁶ The international E-Mail Tandem Network at the Ruhr Universität Bochum:
<http://slf.ruhr-uni-bochum.de>

²⁷ Multi-user domains, Object Oriented; operates with synchronous communication.

²⁸ <http://eurocall-languages.org>

resulting in an upsurge of initiatives which facilitated the recognition of this new area as a field of study in its own right. EU-funded projects have helped the field develop continually in a variety of ways.

The tandem network sustained by Ruhr University, for instance, was one such project whose outcomes are both supportive of the European integrative efforts and the pedagogically innovative possibilities opened by this new technology. The LINC (Language Interactive Culture) project is another example which took advantage of the CD-ROM technology to develop a very ambitious learning tool combining pedagogical aims with the newest multimedia technology. The LINC project used authentic material from news media combined with exercises for reading, writing, speaking and listening, and an email facility which enabled communication among learners, with teachers connected as consultants. The CD-ROM contained viable links to further material which might be of interest to students in their further studies, which encouraged autonomous learning. This project was a truly European project in that contributors from all member states, and a few besides, worked together to contribute authenticity and quality to the material²⁹. The CD-ROMs are available for beginning, medium, and advanced levels in fifteen languages. Furthermore, the LINC project had the inbuilt capacity that it could be used by individuals as distance learning due to the consultative function which could be added, and it could be embedded in a more formalised instructional environment³⁰.

These project descriptions illustrate that CALL is not one thing but is in fact a many-faceted phenomenon, and, despite the European angle in the above projects, it is in essence a borderless, global enterprise which takes advantage of a set of very varied technologies; these may originate in the USA, but Europe has been in the forefront with the application of this technology in a pedagogical setting, and for this reason the yearly conferences arranged by EUROCALL have been attended by many American researchers in CALL. The field's diversity is illustrated in the many different acronyms which all involve a computer and instruction, e.g. CAI – Computer Assisted Instruction, CMC – Computer Mediated Communication, CALL – Computer Assisted

²⁹ <http://linc-www.uia.ac.be/linc>

³⁰ For other European projects see Chambers & Davies, 2001.

Language Learning, ICALL – Intelligent Computer Assisted language Learning, ICT – Information and Communications Technology, CASLA – Computer Assisted Second Language Acquisition, TELL – Technology Enhanced Language learning, etc. In the following sections the term CALL will be used as a generic term³¹.

The early CALL applications were mainly of the gap-filling sort, for instance the so-called ‘drill-and-kill’ exercises which were behaviouristic in nature³². These primarily grammatical exercises were followed by applications of the cloze type, i.e. text (re)construction, such as John Higgins' STORYBOARD from 1981, later ECLIPSE. This programme is typical of a number of similar applications which enable manipulation or modification of text but within a limited framework. Like other similar programmes, e.g. Gapmaster, it comprised authoring tools which allowed instructors to insert their own text, i.e. the teacher needed to master authoring as well as instructional pedagogical competence. The programme was evaluated in 1990 by Legenhausen & Wolff, who recognised the problem-solving and hypothesis-testing which learners would engage in but questioned the language learning efficiency, which they found would be better served in task-based activities.

By the mid-nineties the paradigm shifted to what Warschauer and Healey (1998) refer to as integrative CALL, which was able to take advantage of the multimedia technology to enable authentic language use in task-based approaches, e.g. video conferencing, and the incorporation of authentic material and information via the internet or on CD-ROM. Today, technological development has reached a stage where the world-wide web provides availability and reliability equal to that of the CD-ROM and offers the further advantage of a much wider range of opportunities for choice and thus more individualised approaches. At the tertiary level in particular, the opportunities for embracing CALL have increased with the myriad of web-based programmes and materials.

³¹ Chapelle (2001:3, note 1) explains that this was "the expression agreed upon at the 1983 TESOL (Teachers of English to Speakers of Other Languages) convention in Toronto".

³² For an overview see Warschauer & Healey, 1998. Warschauer & Healey divide CALL into behavioristic, communicative, and integrative CALL.

3.2 CALL research

A characteristic feature, which emphasises the inherent diversity in CALL, is the fact that CALL draws on research in a wide range of areas, particularly second language acquisition theory, psycholinguistics, neurolinguistics, human-computer interaction, natural language processing, and pedagogy. This diversity has many advantages but may also accentuate the need for explicit formulation of the field, such as the Joint Policy Statements³³ of CALICO³⁴, EUROCALL, and IALLT³⁵:

Today, CALL activities exploit improved technology to produce highly interactive learning environments, providing effective support for the acquisition of listening, speaking, reading, and writing skills. High-speed networks allow access to authentic cultural materials and link learners to speakers around the world. When integrated into a pedagogical plan, these new technologies enhance learning opportunities beyond anything previously possible. CALL researchers explore and evaluate these new instructional options to establish how they can best integrate them into effective pedagogy. They also research what these new instructional approaches can tell us about language learning processes. (Introduction; retrievable from www.eurocall-languages.org)

This diversity has resulted in a search for a definition of the research field which was also the theme explored by Chapelle (1997) in her influential article ‘CALL in the Year 2000: Still in Search of Research paradigms?’, in which she refers to several researchers who have pointed out this lack of theoretical underpinning (cf. Oxford, 1995; Holland, Kaplan, & Sams, 1995), and she continues:

...there is a need to specify the particularly relevant questions about CALL and to identify ways that can be investigated through empirical research. In this paper, I suggest that our understanding of CALL would benefit from addressing questions similar to those posed about other L2 classroom learning and from applying the methods used to study L2 learning in other types of classroom activities. (1997:19)

Chapelle finds that: "What is needed then is a perspective on CALL which provides appropriate empirical research methods for investigating the critical questions about

³³ The full Policy Statement (1999) is attached as an appendix to Davies (2001).

³⁴ CALICO, Computer Assisted Language Instruction Consortium.

³⁵ International Association of Language Learning Technology.

how CALL can be used to improve instructed SLA" (1997:21)³⁶. In other words, research and theories about instructed SLA should also be in the forefront of CALL.

3.2.1 Research paradigms

The commonality of research paradigms between CALL and SLA means that ideally the same issues could be investigated under CALL as well as SLA. Some areas feature more prominently in one than the other depending on the relevance in either field. Common ground can be found for instance in the investigation of communicative competence, focus-on-form³⁷, vocabulary learning, the role of feedback, autonomous learning, and in specific and detailed issues which aim to investigate topics within each of these fields in more depth. Natural language processing and human-computer interaction are two further fields which to some extent constitute shared common ground but on a more marginal and general level than the previously mentioned topics, which are more directly connected to learning and pedagogy.

In a very critical assessment of CALL research Professor Richard Towell (1999, document retrieved 17-03-2000), on the basis of a submission to the CILT Research Forum, points out that most of the submissions were not research in accordance with the traditional research definitions. The typical submissions he characterises as:

- a set of grammar exercises with answers
- a set of grammatical explanations and/or definitions
- a set of texts with questions of various kinds plus comments on features in the texts
- a set of combined texts, particularly translations with a source text and various versions in the target language where the users could attempt a translation and/or consult the various versions included. (p.3)

³⁶ Chapelle's article led to a discussion with Salaberry (1999), who argued that her interactionist approach was too narrow and that she ought to have considered the socio-cultural approach in her discussion of computer mediated communication just as she ought to have considered computational linguistics as well as SLA. Chapelle's response (1999) contains a very clear matrix of how CALL is related to other research disciplines (p.109; Table 1) and an equally illustrative one for interactionist SLA and related research questions and methods (p.109; Table 2). Chapelle welcomes the discussion, which she sees as central to the development of the field, but she maintains that "What is needed to appropriately draw from other disciplines is a clear notion of exactly what they have to offer to the development, use, and evaluation of CALL". (1999:108).

³⁷ As opposed to focus-on-forms; cf. Doughty & Williams (Eds.) 1998.

They may be worthwhile and valuable, he states, but perhaps more direct at development than research. Towell's evaluation points to a state of affairs which still needs addressing in the linguistic disciplines of CALL.

3.2.2 Efficacy

3.2.2.1 Computer versus classroom instruction

When new and untried methods are introduced, especially ones which require a not unsubstantial investment, the case often comes to rest on the efficacy of such methodologies. This has also been the case with CALL. In line with the present study project, Kettemann (1995) employed a traditional design with parallel groups of CALL versus traditional instructional groups in his investigation of CALL efficiency in Austrian secondary modern (Hauptschule), grammar (Allgemeinbildende höhere Schule) and vocational (Berufsbildende höhere Schule) schools. The experiments concerned regular English lessons, and the computer groups worked with a variety of CALL programmes, e.g. Eclipse, Storyboard, Gapmaster, London Adventure, and others. He found significant differences in outcome with the largest gains in the CALL groups (p.52). The weakness of the experiment lies in the fact that apparently a variety of CALL programmes were used in the different schools and at the different levels. The names of the mentioned CALL materials indicate that the content primarily must have been vocabulary and text reconstruction. However, the reporting is not particularly precise or elaborate.

A detailed investigation of the issue is to be found in Nagata (1996), who compared workbook and computer assisted language instruction of Japanese particles and sentence construction. She states that "...when it comes to utilizing computers for second language instruction, the question of whether and when computer programmes can be more effective than traditional non-computer instruction is still a basic question to be addressed" (p.55). In her introduction Nagata reviews previous research on the matter finding that many studies from the mid-sixties to the mid-eighties showed no or little difference in efficiency, and she refers to Kleinmann's (1987) comment that the programmes investigated seemed to be electronic textbooks.

Later studies, including Nagata's own, point to the importance of feedback with regard to both content and timing. If the programme is able to provide immediate and relevant feedback, computer assisted instruction seems to achieve better results than non-computer assisted instruction. Teichert, (1985) in his study of learning German vocabulary and grammar, reported a difference of 10%, and Nagata's own investigation (1996) of the Nihongo-CALI programme and comparable workbook instruction showed results similar to those achieved by Teichert with regard to production, whereas comprehension tests showed no significant difference between the two groups. Both Nagata's groups of first-semester students of Japanese at the University of San Francisco received four lessons of instruction containing the same grammar notes, exercises and vocabulary hints. The only difference consisted in the nature of the grammatical feedback, which for the computer group was ongoing and immediate, whereas for the non-computer group the feedback consisted of answer sheets that listed the correct answers. Nagata's work is especially relevant for the present study as the Nihongo-CALI programme was intended for self-study outside the classroom in order to supplement the classroom instruction, which parallels the intentions and possibilities envisioned for VISL. Nagata demonstrated a significant difference in achievement in the post-test scores between the workbook group and the Nihongo-CALI group (significance 0.02) with regard to production. The workbook group had a mean of 44.6 and the CALI group a mean of 55.2.

In order to investigate the importance of the nature of the given feedback, Nagata compared the results of an earlier study, conducted in 1995, with her 1996 results. In her 1995 study she compared two versions of the Nihongo-CALI programme, one was the same as the one used in the 1996 study whereas the other was a more traditional type which only informed the students of the location of the missing particles. The other version of the programme was the same as the one used in the 1996 study comparing the workbook based instruction with the CALL based instruction. When the traditional CALI programme group of the 1995 study was compared to the workbook group of the 1996 study there was no significant difference in results. When the post-test results of the workbook group of the 1996 study was compared to the post-test results of the intelligent Nihongo-CALI programme from the 1995 study, there was a significant

difference favouring the CALI group (1996:64). A retention test after three weeks showed a significant difference at the 0.005 level in favour of the Nihongo-CALI group. Nagata herself says that "the use of a medium (i.e. computer) alone does not bring better effects; rather the quality of the messages produced by the medium affects the results" (1996:67). The number of studies comparing traditional classroom instruction based on textbooks and exercises with CALL programmes is limited. This may be motivated by a variety of arguments, primarily pedagogical, it seems.

The paradigm shifted in the 1990s from research-based studies, like Nagata's 1996 study mentioned above, to learner centred approaches. It was no longer a question of choosing between computer instruction and traditional instruction. Rather, it became a matter of how best to apply the computer, often with the underlying premise that the computer was a given, among other things because it enhanced learner autonomy, which was and is a goal that serves as a benchmark in pedagogical discussions. Chapelle has addressed the issue at some length (2003), examining the viewpoints and interests of the language teacher, the administrator and the publisher:

People wishing to see results of research comparing CALL with classroom study seem to assume that a case needs to be made for using technology in English language teaching. This interest is shared by some language teachers, administrators responsible for budgeting decisions, and commercial publishers even though their ultimate use of research results might differ (p.70).

It is her experience that "The results of research comparing "the computer" to "the classroom" are not conducive to developing principles of language learning and teaching" (p.76), a view which is supported by Garrett (1998) although the latter concedes: "This is of course not to say that efficacy studies are not possible or valid or worthwhile, when they are appropriately constrained" (Garrett, 1998:8).

The direction of CALL research is now headed towards "best practices"³⁸, which is a process that makes the field less technology centred, in that pedagogy, learner attitudes, and autonomous learning are devoted more attention. Chapelle states:

³⁸ See for instance "Best Practice - Best Language Teaching Methods", A Leonardo II Project (EU funded), at <http://www.languages.dk/methods/index.html>.

In view of my experience suggestion technology does now and will in the future play a significant role in teaching and learning, it seems to me that the priority in the field should be research that addresses questions that can inform teachers and learners about the best ways to design and use technology. (2003:76)³⁹.

Studies like Nagata's (1996) study of two different versions of the Nihongo-CALI programme⁴⁰ signify the direction of CALL research today. Indeed, even studies of comparison between two programmes or two applications of the same programme are rare. Research today tends to revolve around issues such as learner behaviour, learning issues, learner attitudes, or testing.

3.3 CALL and grammar

3.3.1 Feedback

The Nihongo-CALI programme investigated by Nagata was further developed and eventually replaced by "BANZAI" which was the basis of a study by Nagata (1997) in which she compared two types of feedback, rule-based and example-based. BANZAI, like VISL, is based on a parser, and this parser is able to parse the students' input and compare it to the correct answer coded into the computer. Thus, the feedback will be 'intelligent' in the sense that it will be specifically related to student input. In this experiment (n=30) Nagata compared the results of two groups according to the type of feedback they received, i.e. deductive (rule-based) or inductive (example-based). The results showed a significant difference between the two groups in the post-test, favouring the deductive group with regard to production⁴¹. A comprehension test, however, showed no significant difference between the two groups, which was in line with the earlier result from the Nihongo-CALI study (1996). Nagata explains, referring

³⁹ The author has later modified the particular lines quoted:

“In view of my experience which is suggesting that technology plays a significant role in teaching and learning, it seems to me that the priority in the field should be research that addresses questions that can inform teachers and learners about the best ways to design and use technology” (Personal communication 19.02.2007).

⁴⁰ One programme which gives 'intelligent' feedback, and one which gives an indication of an error, in this case for instance a missing particle, but does not say where the error is made.

⁴¹ The test consisted of two parts, a fill-in-the-blank section, and a sentence production test; both were significantly in favour of the deductive group.

to Flynn (1986), that "grammatical competence is less critical for comprehension than for production" (Nagata, 1997:528).

Parsing and grammar is something most people, perhaps unknowingly, use in the form of grammar checkers installed as standard elements in the word processing software of every computer. Grammar checkers can be used deliberately, pedagogically and didactically in language learning. Some universities are working at developing different and improved versions of the traditional standard grammar checkers of the big multinational computer corporations, among them UMIST, whose researchers have developed a grammar checker for English learners of German, *Textana*, by modifying an existing grammar checker. Schultze (1999) describes the difference between a commercial grammar checker and *Textana* as being mainly a question of the type of feedback given to the users, or in this case the learners. *Textana* provides feedback on morpho-syntactic errors rather than on style, which is typically what the commercial grammar checkers do.

Schultze outlines four types or levels of feedback incorporated in *Textana*: error warning, when a possible error has been detected, highlighting the part of the sentence which contains the error, providing a general explanation of a rule, and specific explanations of errors. The final level, correction, could be a potential addition. Schultze argues that this type of grammar checker would provide useful feedback at the production level for language learners and that it would be in accordance with theories of focusing on form rather than focusing on forms (p.123).

The traditional commercial editions of grammar checking programmes are probably of little use to language learners as they are directed towards native speakers whose ability of instinctive evaluation of the grammar checker's proposed modifications is not matched the non-native language user. Furthermore, at the learning stage the feedback given would probably be unintelligible to the language learner. Being told by a grammar checker, for instance, that one's writing contains too many passive sentences, would not be directly transferable into the learning of how to construct active declarative sentences. The learning aspect is beyond the purpose and scope of the commercial grammar checker.

In a paper presented at the EUROCALL 1996 conference, N. Ellis et al. (1996) presented results from research based on a spell and grammar checker for Welsh which was not based on a parser but on programming specifically addressing grammatical patterns which traditionally pose difficulties for learners of Welsh. The programme had two dictionaries of about 3000 words each, one for English-Welsh and one for Welsh-English, with a definition of Welsh words and examples of their use in context. Subjects were divided into a vocabulary group with access to the dictionaries and the spell checking part of the programme, and a grammar group with access to the grammar checking facility of the programme and the two bilingual dictionaries. Results showed that the grammar groups outperformed the vocabulary group with regard to grammar performance although both groups improved. However, there was no significant difference in performance on vocabulary, both groups improved their performance. The experiment was conducted in an environment where the subjects were using the CALL programme in relation to free text production. Unfortunately, there seemed to be little long-term retention in either group. One interesting finding was that the weaker students tended to use the look-up facilities less than the higher performing students (p.13 of first draft of 16 October 1996⁴²).

The grammar checker *Grammatik V* was investigated in a study by Wei and Davies and the results presented at the EUROCALL 96 conference in Hungary. Original texts produced by learners of English as a second language were subject to processing by the grammar checker which could react to three categories of errors: mechanics (e.g. spelling), grammar (word classes and S-P agreement), and style (according to a chosen formula, e.g. business letter). *Grammatik V* was not particularly effective in detecting parts of speech: "...the programme would fail to identify subject and verb not immediately adjacent to each other but with modifiers in between; and it would sometimes fail to identify parts of speech, i.e. perceiving nouns, adjectives or adverbs as verbs, or vice versa. This may have been caused by the limitation of the programme's parsing ability; nevertheless, some of the commonly used sentence structures and

⁴² Received in personal communication from N. Ellis, 26.06.2000.

collocations were neither recognised nor accepted."⁴³ (p. 5 of the retrieved document). This emphasises the need for users to have a level of attainment which allows them to make discriminating judgements of the responses given by the computer programme. Hence a tool like *Grammatik V* might be an interesting and valuable instrument for very advanced learners but quite unsuitable for learners whose interlanguage is less sophisticated. Student interviews revealed that most participants were dissatisfied with the programme, probably due to expectations which they tacitly had had about the capabilities of the computer's 'intelligence'. "It also reveals students' perceptions and expectations of the computer, i.e. as a machine which not only delivers information about words and grammar when required, but which also provides information for formulating ideas that they want to put across but do not know how" (p.13 of the retrieved document). This raises the issue of intelligent CALL programmes, their capabilities, and feasibility.

The above quotation points towards a state of computer technology which is beyond the feasibility of the present capabilities of parsers and natural language processing (NLP) which are at the core of intelligent CALL.

The use of parsers in language instruction is commonly referred to as intelligent CALL or "ICALL". It might more accurately be described as parser-based CALL, because its "intelligence" lies in the use of parsing – a technique that enables a computer to encode complex grammatical knowledge such as humans use to assemble sentences, recognize errors, and make corrections. (Holland et al., 1993:28)

Holland et al. go on to say that: "The distant goal of NLP is to have the computer analyze and respond to language much as a human would" (p.29).

The role of feedback is at the core of the successfulness of an application, be it from a

⁴³ Examples of input and computer response:

- a) Text input: People become infected by the virus everyday
Computer response: Normally an adjective like 'everyday' doesn't modify a verb.
- b) Text input: How much cholesterol do I need?
Computer response: The singular subject 'cholesterol' takes a singular verb, not the plural verb 'do'.
(p. 6 of retrieved document).

learning point of view or from a user's point of view. Generally speaking, though, ICALL and NLP in CALL have not been particularly successful. "Even the most concerted attempts to involve researchers, teachers, publishers, and industry, such as the EC-funded ReCALL-project, have not been continued. Until now channelling NLP into CALL by way of ICALL has had little success" (Jager, 2001:102; see also Matthews, C, 1993).

3.3.2 Syntax

The views of the students referred to above point in the same direction as that towards which researchers in the field work, and despite constant achievements and improvement, it is still not accomplished even though some NLP systems are remarkably good, VISL's own being one of them. Good NLP systems – basically parsers - require literally thousands of rules, which makes them costly to develop, and even at the best of times such systems will find it hard to match the flexibility and plasticity of the human brain, which is able to act and react on intuitions.

The ad hoc adaptability of the human input processing system is hard to transform into algorithms. Kempen (1998) refers to research into sentence parsing, which has moved from the assumption that there were two stages in human processing, one initial set of syntactic strategies followed by a second and extrasyntactical process which would modify earlier input. He states that:

...this two-stage model has been seriously challenged by experimental data, collected by means of increasingly refined and sensitive techniques, indicating that the parser is affected by extrasyntactic information immediately upon receiving new syntactic input. The most likely conclusion from this work is that, even if a short purely initial parsing stage exists, it is relatively inconsequential. A more fruitful research strategy is predicated on the assumption that syntactic and non-syntactic factors can operate concurrently and determine parsing preferences in mutual interaction...
(Kempen, 1998:9 of retrieved document)⁴⁴.

Kempen sees 'interactivity' as the "key feature of the control structure of the human sentence parser" (ibid. p.3 of retrieved document). The interactivity being referred to is

⁴⁴ Retrieved 18 February 2007 from
<http://www.gerardkempen.nl/publicationfiles/KempenSentenceParsing.pdf>

that between lexical, conceptual, and referential factors (ibid.). Despite the obvious discrepancies in performance level between human processing capabilities and electronic processing capacities, it is still worthwhile to work towards tools and applications that can help pedagogical innovation and support students in reaching their full learning potentials. Apart from VISL there seems to be only one other CALL tool with a focus on the pedagogical application of syntax learning, that is, an application which can be used in an actual learning framework, and that is the visual grammar developed by Kempen and associates (Kempen 1999)⁴⁵. The Performance Grammar Workbench under development by Camiel van Breugel is an application which is comparable to VISL, but which has the attractive feature of being able to determine correct word order and give interactive and corrective feedback. The interface with the ghost family (Spookjes), which was created for use in elementary schools, makes dependency structures and the syntactical hierarchy fun to work with and the cute little creatures are bound to appeal to children. The lexical frames on which the system is founded will form families which are visualised as ghost father, ghost mother, and ghost children (for illustrations see Kempen 1999:232-233). One feature of the system is that the representations of sentences can work with as well as without linguistic terminology, which increases its potential applicability to several levels of proficiency.

All too often such considerations are lacking despite the obvious possibilities of the medium. The pedagogical aspect is also a concern voiced by Salaberry (2001): "...it is possible that the most important challenge posed by technology-assisted language learning will be the identification of the pedagogical objective that technology-based teaching is intended to fulfil" (p.50). McBride and Seago (1997) propose that many grammar resources fail to demonstrate their full learning potential because they assume or presuppose a fuller knowledge of basic grammatical concepts and appropriate metalanguage than is actually the case for most students (see also Garrett, 1995).

3.4 CALL at the tertiary level

There is no unitary systemic framework within which CALL could or should be

⁴⁵ See <http://www.wi.leidenuniv.nl/home/cvbreuge/bin/spookjes> and <http://www.liacs.nl/~cvbreuge/ToKeN2000/>

discussed and evaluated at the tertiary level⁴⁶; first of all, because the conditions and principles of language instruction itself are context-dependent and in many cases language-specific. In countries where the obligatory status of foreign language learning requires children to receive instruction from an early age, students will have achieved a relatively high level of proficiency at the school leaving age. This is the case for Denmark and other Nordic countries but also in the Netherlands and Germany. Countries in which English is the native language may not see the same need for foreign language teaching and learning, and hence this type of instruction is initiated later and often on a selective basis. Consequently, the proficiency level will be lower when the transition from pre-university to university level takes place. This, among other things, appears to make a difference in the application and availability of CALL⁴⁷. As a consequence, or perhaps in addition to these factors, the reluctance among faculty to engage in this mode of instruction is common.

3.4.1 English and CALL

When students are enrolled in the English study programme at a Danish⁴⁸ university, they are no longer regarded as language learners in the fundamental sense of the term. Rather, they are users, albeit with a need for further development and guidance towards higher accuracy, which seems to be lacking to a particular degree. The curriculum bears comparison to literature, history, and social studies programmes in English-speaking countries. The use of CALL applications and methodology in these comparable study programmes in English-speaking countries appears limited just as it does in English departments at Danish universities. A look at language programmes at university departments in English-speaking countries, be they ESL⁴⁹ programmes or other language learning programmes, reveals a different situation. Here CALL is fast becoming an integral part of the range of tools and methodologies available to the language teacher. Very few teachers of say, Spanish, German, French, etc. at university levels in English speaking countries do not avail themselves of CALL in one form or

⁴⁶ Chapelle (2001) states: "... a range of methods can be used for evaluating CALL because methods need to be chosen to address particular research questions." (p.9).

⁴⁷ CALL materials proliferate in all major languages, but I will be speaking from the point of view of English as a foreign language.

⁴⁸ It is assumed that this also applies to other Nordic universities.

⁴⁹ English as a Second Language.

another. The same is true for teachers of English as a foreign language. Of course, their students are learners at a much lower level of proficiency than are Danish university students of English, and therefore both the need for and the availability of good CALL materials are higher. At issue here is finding the right grounds of comparison and the ensuing selection of strategy. These will depend on context, learner fit, didactical concerns and resources, among other things. In the following the perspective will be that of very advanced English students, and the focus will be on grammar or linguistic issues.

3.4.2 Availability

Multimedia CD-ROMs and other commercially marketed CALL materials are usually very good, very expensive and available at levels ranging from beginning, over intermediate to advanced. They vary in their didactic approach, but their scope is usually general proficiency, or they may be aimed at English for special purposes (ESP), but hardly ever at the high level of proficiency and specialisation required for the very advanced university programmes in the humanities.

A variety of freely available possibilities can be located on the internet. They seem to fall into two categories: either multiple choice formats of stand-alone applications or ones that connected to courses and home pages of individual teachers. A good example of the former is the internet-grammar website offered by University College London (URL: <http://www.ucl.uk/internet-grammar>). The site is thorough with a pleasing layout, easy to use, and it gives feedback. The drawback is that it cannot be customised, and the feedback is a standard text which is the same irrespective of your correctness level or the type of errors you make. The learning outcome would depend on the dedication and analytic abilities of the individual user. From an overall point of view it is a 'drill-and-kill' type of tool, and although it does accept student input in the form of ticking off one of two choice possibilities, its feedback makes it resemble a digital textbook. The type of website available is the one connected to a course book – or a particular course offered by individual teachers⁵⁰. Typically, the site would contain probing questions related to the various chapters or topics of the book. The purpose is

⁵⁰ See for instance <http://people.uncw.edu/veit/DEG/exercises/Answers5.8.2c.htm>

that you can test yourself, and how well you have been able to understand or absorb the subject matter of the course, but it is not a didactic tool aiming to play an active role in the learning process. The similarity to the textbook is striking and typically content-based, i.e. it is not aimed at language learners nor is its aim that users improve their interlanguage grammar; rather, its context is a linguistics course whose students need to master particular linguistic features.

Finally, there are the "Ask Miss Grammar" types of sites which allow you to type in your question which will be answered after a short period of time. This can be very useful, no doubt, but it is more like an e-mail exchange and not a systematic learning tool. This does not mean that it has no learning potential; in fact, it might be quite efficient from a learning point of view since it involves student input, interaction and speech acts. However, its scope is limited, there is no quality assurance, nor is the feedback immediate. There are sites created with the EFL learner in mind which contain good features for specific issues such as prepositions or verb tenses, but generally the level is too low for the very advanced university student in the Nordic countries⁵¹.

3.4.3 Needs

On the basis of the above-mentioned factors, it might be claimed that there is no need for incorporating CALL at university level since the commercially available programmes are not suitable for most university level courses. Languages other than English might benefit from the available prevalent materials of the kind described above, but for the English language programmes the perspectives and considerations would take on different evaluative qualities. The English study programmes offer content-based courses at a near-native level in literature, history of the English speaking world, social studies, linguistics and communication. Expectations are that students are competent users of English at enrolment. The commercially developed programmes frequently come under the category of 'tutor' based programmes (see section 3.1), a category which as a rule cannot be meaningfully applied in the English study programmes in the Nordic countries. This said, recent years have shown that there are

⁵¹ See for instance <http://a4esl.org/q/h/grammar.html>

two issues which must give rise to concern. One is that the level of accuracy in language use seems to have been on a downward route, i.e. some awareness raising needs to occur. The other, and probably related issue, is the fact that students have very little metalinguistic knowledge; metalanguage as well as linguistic concepts need to be established or enhanced. There is a very obvious need here. The question remains how best to meet this need, and there is no ready answer. The VISL project at the University of Southern Denmark represents only one answer to this particular problem, but there appears to be only a limited number of relevant solutions to choose from.

The overall requirements for the potentially appropriate CALL material for English at Danish and Scandinavian universities would be that it could be embedded into the curriculum and that students and teachers find that it would add value to their common goal of achievement. The situation today is that, despite the plethora of materials, very few would meet those criteria. An indication of the need for devoting more attention to the integration of computer based material and its integration in the instructional setting, including training of instructors, can be found in the results of recent research on this issue. Arnold (2006, manuscript submitted for publication) found in her survey of the use of CALL at the postsecondary level that 88% of foreign language teachers at this level “used CALL only infrequently” (as cited in Arnold and Ducate 2006: 6). One study by Bell (2005) investigated the attitudes of language teachers, and 83% were of the opinion that computer technology should be integrated into the language classroom. Ducate and Arnold (2006) relate Bell’s findings of attitudes to Arnold’s findings (2006) and conclude that, given Arnold’s results at the post-secondary level, it seems unlikely that a positive attitude will translate into actual application and they call for more teacher training. Specialised, custom-made programmes, which will meet specific needs, e.g. the needs that VISL is able to fulfil, are to be developed to bring the field forward and to change positive attitudes into teaching practices. Another approach could be to exploit what are essentially research tools and develop pedagogical environments for their use in a merger between learning and research. Corpora and concordances are obvious examples of this. Corpus work is an old and tried method in linguistics and literary studies, but the electronic age has made it much more accessible and easy to work with. This approach opens new and interesting pedagogical perspectives whose

applications are yet to be explored. The preconditions for success are not only technical and pedagogical but also dependent on curriculum planning.

3.5 Summary

Development in CALL has gone from being technology driven in the 1980s to being more focused on pedagogical issues. The early applications of computer-based learning tools were characterised by features similar to that found in textbooks, i.e. gap-filling and grammar drills. The interactivity of the interfaces was limited, and feedback was characterised as being inflexible, standard solutions which learners found unresponsive to their needs. This was replaced by the more impressive CD-ROMs which could have interactive links as well as contain film clips and authentic material, audio as well as aural. The current generation of CALL material is often net-based, flexible and interactive with learner creativity and autonomy at the centre.

With respect to language learning at the tertiary level there is a scarcity of availability of materials directed at the needs of these advanced learners for specialised tools that can provide curriculum-relevant content. Studies and surveys point towards this shortcoming in that it has been found that instructors at the postgraduate level appreciate the potential of CALL but few find occasion to actually use it in their instruction. There was in the early years of CALL some research into the effectiveness of the medium; results, however, were mixed and the design and reporting of these studies, with a few exceptions, were somewhat incomplete. In recent years, however, the issue of efficiency has receded from the debate, and the interest of researchers has moved towards determining in what way CALL can be meaningfully applied and integrated into the classroom.

The availability of courseware for the tertiary level has been and still is limited. The needs at this level require more specialised material to be appropriate for the content-based curriculum. Metalinguistic tools which can be exploited for learning as well as for research may be the direction which is most promising for the tertiary level.

4 THE QUANTITATIVE STUDY

4.1 Background and aim

English language programmes in Scandinavia are at an advanced level with regard to content and proficiency. The students are regarded as language users more than language learners, and consequently there is an emphasis on the metalinguistic side of language at university level. This is an underdeveloped aspect of linguistic knowledge when students enter university, and they generally regard knowledge about the language as a difficult area to engage in. The study of linguistics from a theoretical approach is a natural and necessary discipline in university language programmes for those students who aim at continued linguistic studies and research but also for those who aim at entering teaching professions at various levels in the educational system and therefore need to be acquainted with pedagogical rules of grammar.

There can be many other good reasons for including the meta level in a language programme. One such involves the linguistic development of the students' own proficiency. This is still a matter of discussion in language acquisition research (see Chapter 2), but the prevention of fossilisation, that is the process which leads to inhibition of further acquisition or even deterioration of already acquired structures, may be connected to linguistic awareness and metalinguistic knowledge especially for the older learners (see Ellis, R., 2006; Ellis, N. & Larsen-Freeman, 2006; Birdsong, D., 2006).

It is in response to this, in addition to the academic syllabus inherent to the level of study, that the formal aspects of the English language have a natural position in the syllabus. However, students find it hard to acquire this meta-knowledge, and this fact led faculty at the Department of English of the then Odense University to pursue ways in which to develop tools to enhance and support interest and learning in this field of linguistics. The result was a computer-based tool called Visual Interactive Syntax Learning (hereafter VISL) which was developed in the nineteen-nineties (it began in

1996) under strong inspiration from Associate Professor John M. Dienhart, but it was a collective effort involving faculty members and students. Today the courseware is available on-line, and it is still being developed to include new features and to comprise a growing number of languages.

VISL is a modern-day version of the traditional sentence diagramming which has been and still is prevalent in for instance the American educational system⁵².

The tradition of sentence diagramming is continued in VISL's tree diagramming courseware. The natural language processing capability is built on a parser whose structure is based on constraint grammar (Karlson et al., 1995; Bick, 2001). A pre-tagged corpus of sentences is available in 24 different languages and can be used for practising and sentence analysis, syntax, and a grammatical metalanguage (for a more detailed description of the courseware see sections 4.3.1 and 4.3.2). In addition, the constraint grammar interface allows for free input in 9 languages including all the major European languages and the Nordic languages. The free-input capacity of the VISL interface is quite unique, and its potential as a pedagogical as well as a linguistic NLP tool is under constant development.

Once the VISL interface had reached a stage where reliability and user-friendliness were acceptable it became a potential supplement in the instruction of grammar. Gradually more and more students became acquainted with VISL and faculty began to refer students to it for self-study. A need arose for knowledge about its value as a learning tool and its pedagogical value compared to traditional classroom instruction.

The quantitative study comprises the first three research questions and research question five (see detailed arguments in Chapter 1):

1. Is VISL as good as traditional classroom instruction?

⁵² A well-known prominent example of this is the Reed-Kellogg diagramming system (see Reed & Kellogg, 1907; Kolln, 1994; Despain, 2006). The University of Central Florida has developed its own computer-based tool for drawing up Reed-Kellogg based diagrams (available at <http://www.sendraw.ucf.edu>). In Europe especially Kempen and associates have been working in this direction (see for instance Kempen, 1999; and <http://www.wi.leidenuniv.nl/home/cvbreuge/bin/spookjes.html>).

2. Is VISL equally good for English students and Cand. Negot. students?
3. Is VISL equally good for all achievement groups, i.e. high, middle, and low achievers?
5. Can the syllabus instruction affect the non-syllabus results?

This chapter opens with the design of the study, including the selection and assignment of the subjects and the test design, and it ends with a detailed description and analysis of the results. The purpose of the quantitative study was to seek answers to these questions and to place results in the context of previous knowledge and research in the field.

The design of quantitative study in this thesis included decisions on selection and assignment of subjects to the experimental groups in a framework of a pre-test, post-test set-up. The quantitative nature of the experiments made it desirable to find as many subjects as possible within similar study programmes and comparable syllabi. The background chapters below will delineate the decisions and the programmes chosen as well as the background of the qualifications of the subjects in these programmes.

4.1.1 The two advanced study programmes selected

It was decided that all first year (first term) students from two study programmes which focus on the English language at the University of Southern Denmark, Odense, should be included. The two study programmes – English and Cand. Negot. in English - are very different in many respects, but not in their first term English grammar tuition. The two programmes include an introduction to and training in grammatical sentence analysis, incorporating function as well as form. Both study programmes base their syllabus on the same textbook, i.e. C. Bache, M. Davenport, J.M. Dienhart & F. Larsen *An Introduction to English Sentence Analysis*. The requirements as set down in the two exam standards⁵³ are identical on this particular discipline. Although the given input in English language instruction does indeed vary in the two study programmes with respect to amount and content over the semesters, the exposure to this discipline, which constitutes the core content of my empirical study, is the same. This is subject to the

⁵³ Subsequently the study programmes have been revised. The current requirements are basically the same, though, and can be retrieved at www.sdu.dk.

reservation that since the classes of the English programme and those of the Cand. Negot. programme were taught by different teachers there might also be variations in the content focus and the quality of the instruction. The teaching plans, which I received from the teachers involved, show, however, that in the 10-week period during which I was conducting my study, the same grammatical structures were included in the instruction in both study programmes.

The study programmes taken as a whole differ widely with respect to purpose and goal, and this, of course, is reflected in the overall content and direction of the curriculum. The Cand. Negot. programme is a combination of language studies (two languages: in this case English (major) and a second language which may be either German, French or Spanish) and economics. The emphasis in the language part of the study programme is partly practical language use and partly metalinguistic language description and comment. Some literature courses in a combination with history and social studies are also central as is cultural studies. The majority of graduates from this programme will be employed by private enterprises. The curriculum-based exposure to instruction with a linguistic content in the first semester, the semester chosen for the experimental sessions, was five weekly lessons of each forty-five minutes (oral proficiency, grammar, translation).

The English programme is a traditional humanities set of courses with an emphasis on academic studies in linguistics, literature, cultural and social studies, mainly pertaining to Britain and the United States. The linguistic courses aim at greater detail, depth and scope than those of the Cand. Negot. study programme. Graduates from the English study programme will typically be employed in education (upper secondary education equivalent to 6th form colleges or senior high schools) or go on to post-graduate studies. The curriculum-based exposure to instruction with a linguistic content in the first semester was 6 weekly lessons of each forty-five minutes (phonetics, grammar, translation).

Grammar and translation together amounted to three lessons a week for students in the English programme and to two lessons a week for students in the Cand. Negot.

programme. From a purely quantitative measure there was a difference in the total amount of time devoted to this linguistic area of one lesson per week; however, as to grammar in general and sentence analysis in particular, the evaluation of the course content, which was based on the formal study plans and demands⁵⁴ as well as the individual instruction plans of their respective teachers and on my own experience with this, was that the two study programmes were equal in time and content when only the syntactical structures investigated in the experiment were taken into account.

4.1.2 University entrance qualifications

The students from the two study programmes who became subjects in the experimental study had in common that they all met university entrance requirements, and in that respect the backgrounds of these students were homogeneous, but they had followed very different routes before enrolling in university. All the participating students filled in a profile questionnaire which related to their previous activities and interests which might have a bearing on their linguistic proficiency. The sections below will contain information on the formal differences between qualifications and individual difference with regard to years of previous instruction and the type of entrance exams held by the subjects.

The Board of each study programme will decide what entrance qualifications to accept. This is primarily a question of deciding whether higher level ('B-niveau') or advanced level ('A-niveau') should be the required minimum. For both study programmes the respective Boards had set level B as the required minimum although quite a number of students in the experimental classes did actually have level A.

There are a number of possible combinations of educational backgrounds which will supply the required standards. At the time it was possible, for instance, to opt for level A as well as level B in both the mathematical and the language lines⁵⁵ at the Danish "Gymnasium" (equivalent to senior high and sixth form colleges, or the European

⁵⁴ Studieordning for Cand. Negot.-uddannelsen med engelsk som hovedsprog 1997 and Studieordning for faget Engelsk. Revideret udgave. Maj 1998.

⁵⁵ This division into a mathematics and a language line has now been discarded. Instead the Levels B and A of each subject can be acquired in combinations depending on the offered programmes at each "Gymnasium" or higher secondary school.

Baccalaureate). The levels A and B are also available at the Higher Preparatory Exam (HF) and the more technical (HTX) and business (HH) oriented forms of secondary education at pre-university level.

In principle there should be no difference between a HHX exam in English at level B or A and an exam from a "Gymnasium" or HTX at level B or A. Universities accept these various exams as being equivalent to each other. However, when the detailed syllabus and exam requirements are looked into, it appears that there are variations on the one hand, but on the other hand it has been difficult to pinpoint the exact difference between levels B and A (see Appendix V. Overview of entrance levels A and B; and Appendix VI. English supplement to "Gymnasiebekendtgørelsen").

4.1.2.1 "Gymnasium" (academically oriented secondary education)

Requirements for the language line and the science/mathematics line do not differ for the same level. So level A is a level A and a level B is a level B regardless of line of study⁵⁶. However a closer examination of the requirements revealed differences whose major characteristics are outlined in the sections below.

Students with a B-level exam in English are required to demonstrate that they understand and master word classes, morphology and syntax⁵⁷. Their prescribed proficiency level in oral as well as written English is an ability to use the language in a correct, fluent, and precise way combined with a level of reading competence and general linguistic awareness which should result in a precise understanding of all linguistic details and the context in which they appear.

⁵⁶ Undervisningsvejledning for gymnasiet. Engelsk. Fagets forskellige niveauer. Undervisningsministeriet. Uddannelsesstyrelsen. Maj 1999. Document retrieved on August 8, 2000 from <http://www.uvm.dk/gymnasie/almen/vejledning/undervisgym/engelsk.htm>. Gymnasiebekendtgørelsen. Kapitel 1. Document retrieved on August 8, 2000 from <http://www.uvm.dk/lov/bek/1999/0000411.htm>. Bekendtgørelse om den erhvervsfaglige uddannelse til højere handelseksamen. Document retrieved on September 18, 2000 from <http://www.uvm.dk/lov/bek/2000/0000249.htm>. Bekendtgørelse om den erhvervsgymnasiale uddannelse til højere teknisk eksamen. Document retrieved on September 18, 2000 from <http://www.uvm.dk/lov/bek/2000/0000524.htm>. Bekendtgørelse om kursus til højere forberedelseksamen og om studieforberevende enkeltfagseksamen. Document retrieved on September 18, 2000 from <http://www.uvm.dk/lov/bek/1995/0000605.htm>.

⁵⁷ I only list requirements which fall directly within the field of this project.

Level A is level B at a higher level of abstraction and increased demands of proficiency in that the students in addition to the requirements for level B must

- attain and be able to demonstrate an assured and nuanced command of the English language, passively as well as actively
- express themselves fluently and variedly with a high degree of idiomatic correctness and precision including an awareness of differences in genres⁵⁸ and levels of formality
- acquire a phonemically correct and distinct pronunciation.

With regard to linguistic knowledge, awareness and general proficiency the major difference between level B and level A seems to lie in the level of expected abstraction and linguistic correctness.

4.1.2.2 “HF” (Higher Preparatory Exam)

The linguistic aim for level B or “Tilvalg” is to develop the students’ ability to use and understand the English language as well as their general linguistic awareness. Their written and oral proficiency is developed through reading, written assignments, and the acquisition of the grammar of the English language.

The students are instructed in how to use dictionaries and other reference works. The text work includes pre-1900 texts. Classes are primarily conducted in English.

The linguistic evaluation focuses on idiomatic expressions, semantics, syntax, morphology, and orthography in which areas the student is expected to demonstrate a fairly good mastery of the English language system.

The specifications and requirements for level A or “Højniveau” are the same as those for ‘Gymnasium’ level A.

4.1.2.3 “HTX ” (Higher Technical Exam)

The aim is to give insight into the English language system so that the students can develop their linguistic knowledge and awareness in order to use the language fluently

⁵⁸ Set texts include a Shakespearean drama and some pre-1900 literary texts in the original.

with some degree of variation and precision in both oral and written language production. The texts which the students are presented with are to a large extent texts whose subject matter is related to technological and scientific aspects even though social and cultural matters are included.

The aim for level B is to give, expand, and systemize linguistic knowledge and proficiency so that the students can

- use the oral language fairly correctly and precisely
- express themselves coherently and with some degree of correctness in written English in a variety of contexts
- apply syntactical structures particular to technical language
- acquire a basic vocabulary of technical and scientific terms
- translate a text from Danish into English and vice versa.

The students are taught grammar, including exercises in structuring linguistic observations. In the evaluation the emphasis is on fluency while exact linguistic correctness is considered of minor importance. Basically, level A is just a further development of level B. Linguistic correctness is evaluated at this level, though.

4.1.2.4 “HHX” (Higher Business Exam)

The general aim is to increase the student’s linguistic knowledge and awareness, and to develop his/her linguistic competence.

The instruction at level B aims at making the students able to understand and express themselves in varied oral and written language. The instruction as a rule is in English, and the set texts are primarily business related although modern literature is incorporated into the syllabus. Through working with the grammatical, semantic and idiomatic aspects of the English language the students are given an understanding of the connection between culture and language, and an awareness of their own language acquisition process.

Level A represents a further development of level B, but for level A students, sociolects

are included in the understanding of the connection between language and culture. Unlike level B students, they are not required to know anything about their own acquisition process.

4.1.2.5 General comments

As is evident from the above outline (and Appendix V), answering the question of how much first year students know about English grammar is not so straightforward. Nor is it clear what their general linguistic proficiency level is. The first thing that strikes one is that it is very different what one finds under the headings of level B and level A, respectively.

Level B *Gymnasium* has very little in common with level B *HHX*, and the same is true for the respective levels A. The above synopsis and the overview in Appendix V represent what is explicitly stated in the government guidelines. Reality in the classroom may prove to be somewhat more complex. For instance, under *HF* level B it is stated that classes are primarily conducted in English. For none of the other streams and/or levels is this stated explicitly, but my experience tells me that this is true at least for all levels and streams at the *Gymnasium* and *HF*. This is just one example, there may be others.

Generally speaking, *Gymnasium* and *HF* seem to value the linguistic side more than *HTX* and *HHX* do. Levels at *Gymnasium* and *HF* appear to have high and precise standards with regard to the traditional grammar syllabus such as word classes, morphology, syntax, and lexis. When one turns to *HF* level B, *HTX* both levels and *HHX* both levels, the emphasis on detail and correctness is not in the foreground. Here the emphasis seems to lie in communication abilities, and where *HTX* and *HHX* are concerned, acquisition and training of English for special purposes (ESP).

It is remarkable that there is no written exam at *HTX* level B, and it is stated explicitly in the *Bekendtgørelse* Section B.3 that linguistic correctness is of less importance:

“*Desuden lægges der vægt på, hvor sammenhængende og naturligt eleven udtrykker sig,*

og i mindre grad på, om sproglige enkeltheder er korrekte”⁵⁹ [It is valued that the student is able to express himself or herself in a coherent and natural manner, and it is considered less important that the linguistic details are correct]. It is also worth noticing that for the *HTX* and *HHX* exams some of the items are in Danish, and this is true for level A as well as level B. The linguistic terms applied in the government guidelines for *HTX* and *HHX* are very broad and general, e.g. linguistic awareness, acquisition of grammar but no details about their manifestations. Consequently, the impression the various *Bekendtgørelser* leaves you is that *HTX* and *HHX* concentrate on describing instruction forms and exam forms and field content but are less concerned with linguistic content, whereas *Gymnasium* and *HF* seem less concerned with form but highlight content in that the degree of detail in the linguistic aspects is very high.

4.2 The subjects

The experimental classes comprise subjects with entrance exams from all the varieties described above (see Tables 4.3 and 4.5). These students have been working with English as an L2 since their fourth or fifth school year, i.e. since they were ten or eleven years old, and must after at least eight or nine years of learning English and passing

Table 4.1
Years of instruction prior to university. By treatment group. Cand. Negot.

Treatment	N	mean	sd	min	p25	p50
BASIS	15	8.40	1.12	6	8	8
NON-VISL	20	8.70	1.22	5	8	9
VISL	18	8.22	1.73	4	8	8
Total	53	8.45	1.38	4	8	8

Treatment	p75	max
BASIS	9	11
NON-VISL	9	11
VISL	8	13
Total	9	13

⁵⁹ URL: <http://www.uvm.dk/lov/bek/20000000524.htm>. Page 3 of retrieved document. Document retrieved on 18 September 2000.

their exams along the way be considered competent language users as well as language learners (for actual years of instruction prior to university see Tables 4.1 and 4.2).

The difference that appears from Tables 4.1 and 4.2 is that the mean number of years of prior English instruction is almost the same, but a closer look at minimum, maximum and the median reveals that the Cand. Negot. students are at a slight disadvantage. The median for Cand. Negot. students is 8 in comparison to 9 for the English students. The median is an expression that half the observations are 8 or below for the Cand. Negot. students but 9 or below for the English students. This indicates that the bottom half of the Cand. Negot. students have fewer years of English instruction than the English lower half. The 75% percentile on the other hand has the same mean, i.e. nine years of instruction.

Table 4.2
Years of instruction prior to university. By treatment group. English.

Treatment	N	mean	sd	min	p25	p50
BASIS	21	9.24	1.67	7	8	9
NON-VISL	15	8.40	.91	6	8	8
VISL	18	8.61	.979	7	8	9
Total	54	8.8	1.31	6	8	9

Treatment	p75	max
BASIS	9	15
NON-VISL	9	10
VISL	9	11
Total	9	15

All of these students know how to form questions, how to conjugate regular verbs and the most common of the irregular verbs, how to put nouns into the plural form, how to place adverbs correctly – although adverbial phrases remain a problem – and they have a fairly large vocabulary and a good pronunciation. On an interactional and a communicative level they are relatively well-functioning. Their knowledge *about* the language is quite another matter.

In conclusion, the evaluation based on the *Bekendtgørelser* alone must be that linguistic knowledge and metalinguistic awareness follow a downward scale from Gymnasium and HF level A, through HF level B to HTX and HHX levels A and B. From a communicative and interactional point of view there is no doubt that all these students are quite proficient, but judged in the light of this project it can be expected that only *Gymnasium* level A and B students and *HF* level A students possess any knowledge of word classes, syntax, and morphology at a metalinguistic level. The entrance level as illuminated by the pre-tests shows that this expectation cannot be said to be fulfilled to any significant degree for the subject group as a whole.

Table 4.3
Type of exam. By treatment group. Cand. Negot.

Exam	Treatment			Total
	BASIS	NON-VISL	VISL	
0	0	0	1	1
1	3	3	6	12
2	2	3	1	6
3	4	4	7	15
4	6	10	3	19
Total	15	20	18	53

Category codes: "Gymnasium", Language line=4; "Gymnasium", Mathematics line=3; "HF"=2; HHX=1; Other(incl. HTX)=0

The number of students with exams at levels B and A, respectively, is hard to ascertain as it was not encoded in the university statistics programme at the time. The registration only pertained to whether the university entrance requirements for their particular study programmes were met.

This practice has since been changed, and this information is now routinely registered. This has meant that the information is now readily available, but as far as the students involved in the present experimental study are concerned this information was not obtainable. In order to obtain some knowledge a survey was carried out among the Cand. Negot. students and this revealed that nearly all students were in possession of level A qualifications.

Table 4.4
Type of exam. By treatment group. English.

Exam	Treatment			Total
	BASIS	NON-VISL	VISL	
	1	0	0	1
0	2	0	0	2
1	1	0	1	2
2	4	0	3	7
3	7	4	2	13
4	6	11	12	29
Total	21	15	18	54

Category codes: "Gymnasium", Language line=4; "Gymnasium", Mathematics line=3; "HF"=2; HHX=1; Other(incl. HTX)=0

The categories of entrance qualification exams were part of the profile questionnaire which means that the students themselves have provided the information on the category of exam they used as entrance to the university. This information is expressed in Tables 4.3 and 4.4. These tables make it clear that the English students to a higher degree than the Cand. Negot. students have traditional "Gymnasium" backgrounds (42 against 34) and the difference is even more pronounced as regards the language line (29 English subjects against 19 Cand. Negot. subjects). The Cand. Negot. students also comprise a large group of "HHX" student (12) whereas no such notable group is to be found among the English students whose exams other than the Gymnasium are more evenly distributed over the category with the second largest after the "Gymnasium" being "HF".

4.3 Experimental groups

The decision to include the whole population of first year students in the Cand. Negot. English programme as well as the English study programme meant that a total number of 161 first year students in these two lines of study were offered to participate, and they all agreed to participate in the pre-test. The participating students were distributed with 86 English study programme students comprising three classes, and 75 Cand. Negot. study programme students comprising three classes. The results of the pre-tests

could not in all cases be matched by the subsequent post-tests as the drop-out rate, not just from the experimental classes but from the study programmes as such, was noticeable. This had, among other things, the consequence that the number of subjects in the various experimental groups in some cases developed unevenly. The final number of subjects in the study came to 107, and despite the original intention to operate with equally-sized experimental groups, the final distribution became uneven (see Tables 4.4 and 4.5 for final group sizes).

4.3.1 Practical constraints

The ideal assignment plan was to divide the three classes from each study programme into each three groups: VISL, NON-VISL, and BASIS⁶⁰ groups, the latter functioning as control groups. The plan could be accomplished for two classes of each study programme. The third class in either study programme could not be accommodated for in this way; in fact, the timetable made it impossible for me to use either of these two classes in my experimental sessions, and they were therefore assigned to a function of extra control groups to supplement the designated basis/control groups of the other two experimental classes.

The availability of the computer rooms was one constraint, the other was the complexity of the timetables, especially the Cand. Negot. one, which proved impossible to rearrange in such a way that all three classes could be subdivided into each three subgroups. The Cand. Negot. timetable is complex because of the complex nature of the study programme, which must fit into timetables covering several centres and departments, e.g. English, Literature, Philosophy, Economics, History, Marketing, Statistics and Demography, Accounting, Finance and Law, so that the students in the three classes of English are assigned in a criss-cross fashion into ten subgroups for subjects other than English. Once the timetable is finished it is practically impossible to change since the ramifications are so far-reaching. The English study programme timetables are not so intricate, but even so changes did prove difficult.

⁶⁰ As noted elsewhere the BASIS groups represent the regular class instruction according to the regular study programmes. In this study they function as extra control groups against which the experimental group results from the VISL and NON-VISL groups can be put into relief and thus evaluated and perspectivised.

4.3.2 Assignment to experimental groups

Students were randomly assigned to the experimental groups. In the English study programme the students were assigned according to date of birth (not including month or year), and the Cand. Negot. students were assigned alphabetically according to the letters of their surname. The group divisions reflected the purpose of the experiment, which is to say that the classes in each study programme were divided into computer-based instruction group groups (called VISL after the Visual Interactive Syntax Learning courseware), non-computer based instruction groups (called NON-VISL) and the basis groups which served as control groups.

For practical reasons (see 4.3 and 4.3.1) the various types of instructional groups were instructed separately from the parallel groups in the other study programme as were the groups from the different classes even though they received the same designated experimental instruction. As mentioned above, the aim was to operate with approximately the same number of students in each group, and the group size was determined by the number of computers (i.e. 12) in the computer room. The final set-up was as follows:

Study programme	Class 1	Class 2	Class 3
ENGLISH	VISL NON-VISL BASIS	BASIS BASIS BASIS	VISL NON-VISL BASIS
CAND. NEGOT.	VISL NON-VISL BASIS	BASIS BASIS BASIS	VISL NON-VISL BASIS

Figure 4.1 Experimental groups by study programme

All three groups in classes 1 and 3 were taught by the same teacher in their ordinary classes whereas class 2 was instructed by a different person which is to say that English classes 1 and 3 were taught by A, and English class 2 by B; Cand. Negot. classes 1 and 3 were taught by C, and Cand. Negot. class 2 was taught by D. The experimental groups

were all given the instruction involved in the experimental study by the same person, i.e. the present researcher.

4.4 Treatment

4.4.1 The amount of input

The two types of experimental groups received the same input (see Appendix I Analysed sentence corpus.) over a period of ten weeks. This input was comprised of 57 English sentences, the length and complexity of which increased progressively over the test period⁶¹. The students also received grammatical information pertaining to the syntactical structures of the training sentences of a given session in order to make sure that all subjects were in no doubt as to which structure was being trained during a given session. The latter part was discontinued after the first seven weeks for the reason that at this point in the experiment all the basic sentence structures had been presented to the students. The remaining sessions were training sessions only, albeit with ever increasing levels of complexity in the training sentences. The increased complexity, however, should be understood to mean a complexity in manifestation and length, not added intricacies, of the same syntactical sentence functions as the students were exposed to during the initial six weeks. Initially, the experimental groups were given pedagogical rules relating to the functional roles of the sentence structures.

Semantic, morphological, and formal functional knowledge in general was reserved to the basic classes which were the same for the experimental groups and the control groups, as was instruction in standard English syntax. The immediate objective of the sessions was for the students to learn the metalinguistic terms as well as the concepts of the syntactic analysis as presented below.

4.4.2 The syntactical input

The sequence of the structural patterns presented to the subjects followed the sequence of their text book (Bache et al., 1999), but all the training sentences were either created

⁶¹ The first sentence was: *Anne wrote a letter*; the last sentence was: *Jeremy told Mathew he knew Sophie, but omitted to mention that she had been called a traitor by several delegates and that she had come to see him.*

with the experiment in mind, or they were sentences derived from novels or previous exams and unknown to all subjects prior to the sessions. During the initial seven weeks the following sentence functions were presented to the subjects: S (subject), Sp⁶² (provisional subject), Sr (real subject)⁶³, P (predicator), Od (direct object), Oi (indirect object), Op (provisional object)⁶⁴, Or (real object)⁶⁵, C (complement) Cs (subject complement), Co (object complement), A (adverbial), H (head), DEP (dependent), SUB (subordinator), CJT (conjoint) and CO (coordinator). Students were introduced to forms simultaneously with the introduction to the various functions. The forms trained were the word classes (n, pro, v, adj, adv, prep, art, infm, num, conj), groups (g), clauses (cl) and compound units (cu)⁶⁶ (see Appendix II Key to VISL symbol set.). No sentences contained interjections, and no distinctions were made between proper nouns and common nouns, nor were classifications required for pronouns and articles and verbs. They followed a progression from the basic sentence function of S, P, Od, Cs, A, as well as H, DEP, moved on to include Sp, Sr, Oi and discontinuous constituents, then Cs and Co. That was followed by a consolidation session containing no new functions or forms (session 4). The next step was to introduce SUB and subclauses followed by CO, CJT and compound units (paratagmas in the new format) and greater variety and complexity in the form manifestations of the various constituents.

The forms were introduced in a sequence progressing from words/word classes (n, pro, v, adj, adv, prep, art, num, infm, conj) and simple groups to complex groups and short, simple non-finite subclauses followed by finite subclauses.

The training sentences contained no Or and Op⁶⁷, but students were introduced to these structures in their textbook, and the training sentences in their basic classes during their time with their regular teacher. The last three sessions were training sessions only, as was session 4. Sentences containing between 17 and 30 words were trained. These sentences were analysed to word level.

⁶² Present interface uses Sf

⁶³ Present interface uses S

⁶⁴ Present interface uses Of

⁶⁵ Present interface uses Od or Oi, respectively. The simple O has disappeared from use.

⁶⁶ Present interface uses par

⁶⁷ Neither pre-test nor post-test contained these structures.

At sentence surface level, the corpus of training sentences contains 24 varieties, the most common being S P Od (10), S P Cs (8), A S P Cs (5), A S P A (5), A S P Od (3), S P Od A (3), CJT CO CJT (3), Sp P C Sr (3), S P Cs A (2), S P Od Co (2), S P- A -P O A (2), S P Oi Od A (2), P- S -P A A (2). The remaining 11 structures are represented in one instance each. Of the 57 sentences 4 are questions and 3 sentences are paratactic at sentence level. There are 41 examples of A, 27 examples of Od, 19 examples of Cs, 4 of Oi, 4 examples of Sp and Sr (but no Op and Or), and 3 examples of Co. Adverbials appear in every conceivable position and form. Cs and Od each appear once in a fronted position, and there are four instances of partial inversion. In other instances the sentence constituents appear in the order which reflects the normal syntactic rules of the English language.

Sessions 1 to 4 concentrated on the introduction and consolidation of the normal word order and the most frequently occurring syntactic structures. In session 3 a special emphasis was placed on differentiating Oi/Od/Co/A. In session 4 the difference between complements, object and adverbials was in the foreground. Less frequent syntactical structures, especially fronted constituents, were trained in session 5, as was the Sp/Sr structure. Complex sentences were the predominant feature in session 6, in which clauses were represented as the constituents S/Sp/Sr, Od, A and in the DEP function. Complex sentences were also predominant in session 7, and in both session 6 and session 7 non-finite as well as finite subclauses were represented. Session 8 is characterized by containing sentences of an increased number of words. Sessions 9 and 10 were old exam sentences and contain approximately 30 words each and are characterized by complex structures representing the whole spectrum of this particular syllabus.

The grammatical information which subjects were given at the beginning of the first seven sessions corresponds to the descriptions given in the paragraph above, i.e. the subjects were made aware of the particular purpose of each training session. They were told which particular syntactic structure was predominant in the training sentences and they were reminded of the grammar pertaining to the particular structure(s).

All sessions were conducted in English. The subjects were not instructed to speak in English, but 99% of the time they all did. This reflected the general practice in these study programmes.

4.4.3 Procedure

Two aspects were of special concern to the experimenter in deciding the procedures of the two types of experimental groups. The first concern was the ecological validity of the experiment. The pervasive idea was to place this as close to the way the regular classes were conducted as possible. The second concern was to establish a framework which would ensure equal preconditions for the two types of experimental groups. In other words, this was modelled on what is the objective of sentence analysis in its own right, and its place in the general grammar instruction in these two study programmes. Consequently, the limited training in sentence analysis which it is possible to give students in ordinary circumstances was transferred and expanded for the purpose of the experimental sessions, but the substance and the form remained the same, which is to say that the sessions and the procedures applied could readily be incorporated in the actual study programme.

Pedagogically, the role of the experimenter was not to give ready answers but to make the subjects find their own answers. When a subject approached the experimenter with a question, he or she was often met with a question in response or asked to remember a mnemonic rule which could guide the subject in constructing the correct answer.

4.4.3.1 The VISL groups

The students worked separately at their individual computers in a room with 12 computers (iMACs in a 10 Megabit LAN network). At the beginning of each session the training sentences were handed out, and the instruction, if any, for that particular session given. The VISL programme would provide the students with immediate response to their actions in that a false choice would give no response from the computer to their input. A correct choice would make VISL fill in the box in question. The VISL interface at the time required the students to mark the unit box by clicking on

it and then click on either function or form line to call forth a scrollable menu from which they could make their choice. Sentence constituents consisting of more than one word had to be grouped first by clicking on a *Group nodes* button in the right hand side of the screen. The VISL programme would then fill in the subsequent structure lines and create new boxes in the next level.

The students also had the possibility of getting the programme to show them the classifications of individual items by pressing a *Reveal node* button. Under Tools in the menu line the students also had the possibility of getting the VISL programme to show them the entire analysis of a given sentence by clicking on *Inspect tree* in the menu.

The experimenter was present and available for questions and, in case of technical problems, for assistance with the software or hardware. Software problems usually turned out to be rooted in inappropriate actions by the students, such as missing box marking or missing groupings/levels. Occasionally, there were problems with the VISL server and with the internet connection. Another infrequently occurring problem was stalled computers which had to be reactivated.

4.4.3.2 The NON-VISL groups

At the beginning of each session the training sentences were handed out, and the instruction, if any, for that particular session given. The students worked individually using pencil and paper, writing on either the handout with the pre-printed sentences or on their own notepads on which they would rewrite the sentences. The NON-VISL groups would build the same tree diagram by hand as the computer built automatically for the VISL groups.

The role of the experimenter was slightly different here than it was in case of the VISL groups. In order to make conditions as similar as possible for the two types of experimental groups, the experimenter tried to give these subjects the same type of feedback as the subjects in the VISL groups would receive from the computer, i.e. the experimenter would oversee the work of the subjects, and whenever an error was detected, the student in question would be told that this particular item needed to be

revised. The experimenter was available for questions in the same way as was the case with respect to the VISL groups.

4.4.3.3 Observable differences between VISL and NON-VISL groups

The VISL tool provided certain enhancement features which it was not possible to re-create in the NON-VISL environment. The most prominent feature is the use of colours in VISL. The form box and the corresponding form line containing the options which the students could select from are kept in the same colour, which is different from the colour of the function box and function line. Colour is also used to tell the students when a word has reached its final level so that when analysing a complex constituent the student will be told by the red colour of the words when they have finished the structure correctly. In other words, the structure as well as the form is supported by colour. The structure building is supported in the sense that the VISL programme constructs the various levels in the sentence analysis once it has been told by the operator to combine a given number of words in a constituent. By means of these visual aids the screen will give the students a nicely ordered overview which could be facilitating perception and understanding. In contrast, the NON-VISL groups had to build the tree diagram themselves while trying to keep order in the levels of the analysis, which was not always so easy for them, and they were not helped along by colouring of any kind.

Subjective observations on the part of the experimenter indicate that there was a difference in the way the students worked. Both types of experimental groups worked on the task individually, but the VISL groups worked more intensely focused on their own screen and their own task. The NON-VISL groups were more inclined to discuss items with their neighbours or to ask questions across the room, just as they were more likely to joke or comment about the tasks. In other words, the VISL groups seemed to work more intensely focused on their own task than the NON-VISL groups did.

4.5 Designing the tests: Problems and considerations

4.5.1 The purpose of the tests

In order to decide what exactly to test the students for, it was necessary to consider the aim of the project, the content of the syllabus, and the content of the experimental tutorials. The test area should be a reflection of the subject matter of the experimental classes which in itself should reflect what the students would be exposed to in their regular classes – what they would read about in their respective grammars, and what they would work with in their sentence analysis book, which for both lines of study programme (i.e. English major/minor and Cand. Negot. English).

The syllabus for the first ten weeks of the first term included, for English students as well as for Cand. Negot. students, all the basic sentence functions and forms (see Appendix I Analysed sentence corpus and Appendix II Key to VISL symbol set.). In the experimental tutorials the exercises were developed in accordance with the syllabus, and the sequence of structures closely followed the sequence given by their text book on sentence analysis.

At the end of the first ten weeks of the term all students were acquainted with the sentence constituents (S P O A C), and their forms (word class, group, clause, compound unit), the relationship between head (H) and dependents (DEP), the difference between subordination (SUB) and co-ordination (CO), the nature of conjoints (CJT), as well as constituent order.

When the students enter university they have at some point during their preparatory education been made familiar with sentence functions and forms but in varying degrees of detail. Consequently, they should be able to demonstrate knowledge of some, if not all, aspects of this limited field. They do not necessarily know the appropriate terms to describe it, nor do they necessarily know the same terms, a problem which will be touched on below. However, it is a reasonable assumption that these students, who supposedly are particularly interested in language, do possess this knowledge at a basic level. The task of testing this knowledge was a matter of formulating question items in

a manner which would allow the students to show that they were in possession of the knowledge and which would allow an assessment and a measurement of that knowledge.

The test questions needed to focus on aspects of grammar which relate to the central areas of sentence constituents and word classes. Finding a relevant form for the test items was an issue since the subjects could not be asked specifically to analyse sentences, as this would be beyond their initial knowledge level. The system of analysis taught during the first term is so different from the content of their previous education that any analysis they might produce would not be of such a character that it could be compared with what they would produce after the first term of tuition. The other problem would be to interpret a given analysis correctly since, as mentioned above, the terms and symbols in their previous education would be incomparable not just with the content of university syllabus but would also vary from student to student. Indeed, it would hardly be possible to construct any objective and reliable correction model; a task which would be problematic even under less diverse conditions.

It is, of course, essential to design the pre-test with a regard for these problems, but one should keep in mind that the pre-test cannot function on its own terms. The pre-test and the post-test make up a unit whose scope must include the ability to measure any potential improvement achieved by the students during the test period. Before deciding on the design of the pre-test, one should therefore have a clear idea of what one wants the post-test to contain since the two tests are interdependent with regard to assessment of test results. Basically, the two tests should cover the same fields of knowledge with the post-test revealing the degree of improvement during the test period, and with the pre-test as a measure of the entrance level competence. The overall difficulty is to combine in one test the way to measure the knowledgeable student as well as the novice, and to make sure that what is being measured is their knowledge of the English language (grammatical structures) and not just their knowledge of the appropriate terms or the system of analysis as such.

4.5.2 The method of testing

The first obstacle was finding a way of phrasing the test questions in such a manner that the uninitiated student would be able to understand what it was he or she was required to do, yet keeping in mind that a co-ordination between pre-test and post-test was necessary right from the start. It is one thing to ask questions which would be understood and which would give the student the opportunity to disclose his or her knowledge of grammatical structures, but it is quite another matter to do it with the object of keeping within the conventions of the field, and it would not do to use imprecise language or use terms which deviate from the ones which they will be required to use later when they are taught the subject in question. This would only lead to confusion and a muddling of issues not to mention the difficulty which would arise with respect to making the pre-test results comparable to the post-test results, should one fall to the temptation of using different terms in the pre- and the post-test, respectively. There had to be a unified approach from the outset.

4.5.3 Sequence of items and the level of difficulty

There is one thing which should be remembered at all times when one conducts experiments and tests of this nature, namely that it all stands and falls with the willingness of the students to co-operate. There is no way of paying the students for participating or of offering any other form of compensation. You have to rely on their goodwill and their "sacrifice" of time and energy. You can only hope that they will think it interesting and find it an additional benefit to participate in the project, and you must be careful not to discourage them unnecessarily.

This was a consideration which played a role in the sequence chosen for the test questions in that it was important that students were not discouraged from completing the test or from taking the test seriously. If from the outset they were to get the feeling that this would be beyond their scope, or if they did not see the relevance of the questions, they might either choose not to participate or they might only make a mock attempt at answering the questions, and test results would be highly unreliable. In consequence the sequence of test questions reflects a progression in difficulty, with the easiest questions first and the difficult questions last. This sequence was based on a

judgement of what Danish students find difficult. It was expected that the majority of the general population of students would be able to give correct answers to the first quarter of the questions whereas only the best students were expected to be able to answer the last questions of the tests. Out of the first 6 questions only two (4 and 5) would require specific content knowledge; the others are questions in which the students are asked to judge the grammaticality of presented English sentences. As far as questions 4 (noun/verb difference) and 5 (complement to subject) are concerned, the knowledge required is equivalent to what Danish children are taught in the fourth form of primary school. From question 7 and onward the questions become more centred on the syllabus knowledge of their first term courses and tutorials, with the exception of questions 9, 10, 21, 25 (grammaticality). In effect, the structure of the tests can be described as easy and giving encouragement to go on, then a long stretch of increasingly difficult items with built-in breathers of an easier type, and the tests end with an item which actually is very difficult, but which they can answer by using their intuitive judgement.

All grammar classes were taught in English so it was only natural that English became the language of the tests. From this followed that the grammar terms should be the English terms (which was what the syllabus stipulated for the discipline). This had the additional advantage of creating a common ground between students of Danish descent and students with another ethnic background in addition to creating coherence between subject matter and form. In order to safeguard understanding of the question items even further, a list of grammar terms was attached to the tests. This list contained all the terms used in the questions, and each term was given in English, Danish and Latinate Danish, e.g. *noun*, *navneord*, *substantiv*.

Berry (2000) investigated the effect of metalanguage and explicit knowledge in order to explore the idea that finds expression in for instance Collins Cobuild publications that a more user-friendly (referred to as ‘youuser-friendly’⁶⁸) linguistic expression might enhance learning in contrast to the exact linguistic description language which could be

⁶⁸ ‘youuser-friendly’ is a term derived from the style of this type of language which “*eschews use of the passive and other impersonal constructions in favour of active constructions starting with you*” (Berry, 2000:195).

presumed to be a stumbling block. The study was limited in scope and thus far from conclusive, yet results seem relevant and could warrant further investigation. Berry found that “Overall the results of the study do not provide any evidence that youser-friendly metalanguage is necessarily more effective with grammar reference materials” (Berry, 2000:205). He also found that a consistent style of language whether ‘*youser-friendly*’ or not was better than a mixed style. The mixed style “may appear easier to read and more appropriate, but this does not guarantee understanding and may actually detract from it” (ibid.). This is an interesting result, however tentative, and it is worth noticing that he argues that in an educational setting the metalanguage approach needs no justification, an argument along the same lines as the ones which underpin the two study programmes from which the subjects of this study were selected (see section 4.1). According to the results from Berry’s study the application of a formal style of language would not in itself be a factor in test results. The issue of what he calls “grammatical jargon” (1995:377) is addressed by Hulstijn (1995) as a pedagogical consideration and he suggests that teachers might use “many well-chosen and well-organised examples in their grammar lessons” (1995:378) rather than explicit explanations, because some grammar points are so complex that explanations cannot be simplified. However, he does refer to Master’s study (1994) which revealed that groups receiving explicit instruction in the English article system were more successful than those who did not. However, the focus in Master’s study was implicit versus explicit instruction rather than the use of metalanguage, albeit these are often coinciding elements.

Level of proficiency as far as it plays its role in the test items also needs to be addressed. Issues such as sequence of acquisition and complexity of structure (see overview in Chapter 2; see also for instance Pienemann, 1999) need to be related to the expected level of attainment in the test subjects. It seems to appear from the sections under 4.1.2 on entrance exams levels that most structures can be assumed to be acquired by the majority of students by the time they enter university. Some structures, such as complex relative clauses (see overview in Chapter 2 and for instance Hulstijn, 1995; Hulstijn & de Graff, 1994) may be acquired in the sense that they are used correctly only sometimes and at other times incorrectly. Hulstijn and de Graff (1994) make a distinction between rules which are difficult to use and those which are difficult to explain

(p.103). In relation to the test items in pre-test as well as post-test, this distinction might apply to subject-predicator concord in test item 9. However, as can be seen from Table 4.5, it appears that item 9⁶⁹ is among the easier ones (it ranks 4th), perhaps due to the fact that the form of the item makes it possible to respond intuitively rather than drawing on explicit knowledge. The position of the same item in the post-test (see Table 4.6) is within the ten easiest (9th) items⁷⁰.

Table 4.5
Easiness of items at pre-score over all groups

(n=107)

	item	category	sum	percent
1.	4	syllabus	97	90.65
2.	1	non-syllabus	95	88.79
3.	21	non-syllabus	94	87.85
4.	9	non-syllabus	88	82.24
5.	3	non-syllabus	86	80.37
6.	11	syllabus	77.86	72.77
7.	10	non-syllabus	66	61.68
8.	8	syllabus	65	60.75
9.	2	non-syllabus	63	58.88
10.	12	syllabus	61	57.01
11.	19	syllabus	53.60	50.09
12.	6	non-syllabus	49	45.79
13.	17	syllabus	48	44.86
14.	13	syllabus	46	42.99
15.	23	non-syllabus	46	42.99
16.	25	non-syllabus	40	37.38
17.	18	syllabus	35.60	33.27
18.	24	syllabus	35	32.71
19.	15	syllabus	33	30.84
20.	5	syllabus	32	29.91
21.	16	syllabus	32	29.91
22.	7	syllabus	15	14.02
23.	20	syllabus	11	10.28
24.	14	syllabus	6.33	5.92
25.	22	syllabus	5.95	5.56

It is also apparent from Tables 4.5 and 4.6 that the non-syllabus items are the easiest

⁶⁹ The item in question was: *The boys, who makes a great team, are all great ball-players* versus, *The boys, who make a great team, are all great ball-players.*

⁷⁰ The item in the post-test might be slightly more difficult: *The two features of life in Britain that gives visitors a bad impression are the weather and the food* versus *The two features of life in Britain that give visitors a bad impression are the weather and the food.*

ones for the subjects. The ranking in the post-test is more diversified. The changed position in ranking of the non-syllabus items is natural in so far as the syllabus items become easier on the background of new knowledge, so for the post-test ranking itself is less open to obvious interpretations.

Table 4.6
Easiness of items at post-score over all groups

(n=107)

	item	category	sum	percent
1.	4	syllabus	98	91.59
2.	21	non-syllabus	97	90.65
3.	19	syllabus	94.05	87.90
4.	12	syllabus	94	87.85
5.	3	non-syllabus	91	85.05
6.	18	syllabus	91	85.05
7.	1	non-syllabus	90	84.11
8.	5	syllabus	88	82.24
9.	9	non-syllabus	87	81.31
10.	11	syllabus	86.44	80.79
11.	15	syllabus	86	80.37
12.	8	syllabus	85	79.44
13.	2	non-syllabus	82	76.64
14.	17	syllabus	78.15	73.04
15.	23	non-syllabus	68	63.55
16.	6	non-syllabus	67	62.62
17.	20	syllabus	64.3	60.09
18.	16	syllabus	61	57.01
19.	10	non-syllabus	56	52.34
20.	13	syllabus	55	51.4
21.	14	syllabus	54.69	51.11
22.	25	non-syllabus	54	50.47
23.	7	syllabus	44	41.12
24.	22	syllabus	31.9	29.81
25.	24	syllabus	28	26.17

There was no doubt that the students would find many of the questions in the pre-test difficult, even very difficult: first of all, because they were asked to demonstrate their knowledge in a way that would be unfamiliar to most of them. In pre-university education it is not common to be given grammar tests so the form would be new, and this in itself would probably add a dimension to the level of difficulty. This fact in itself may be one of the explanations for why some students have voiced the opinion that the post-test seemed easier than the pre-test.

This is contrary to the results of the pre-experiment testing of the test (see Chapter 4.4.5). The pre-experimental testing that was carried out to ascertain the level of difficulty of the two tests in relation to each other concluded that if a difference could be found it would be that the post-test might be slightly more difficult than the pre-test, not the reverse, but the variance was only minute (the pre-experimental mean score for the pre-test was 17.67 points against 17.11 for the post-test; see Table 4.6). Secondly, the difficulty of the various tasks which they were asked to perform may not necessarily be of equal difficulty from a language learning point of view as discussed above, nor might all students find the same tasks difficult. This would depend on the way they had been taught previously, and the way they had been used to working with language. The latter was taken into consideration in the form of the test items, and the fact that the students were asked about the same item of knowledge in a variety of ways was designed to make up for this to some degree. For instance, when the purpose was to find out whether the students knew what an adverb/adverbial was, they were asked about it in three different ways (see Chapter 4.4.4), i.e. the same field of knowledge was explored in three different tasks⁷¹.

4.5.4 Test items

There were 25 test items in both tests (see Appendix III Pre-test and Appendix IV Post-test.). The 25 items of the pre-test were repeated in the post-test, i.e. the tasks of the items were repeated but the actual wording of the items was different. The items could be grouped in a variety of ways according to which angle was in focus to explore.

The first natural category was the category of tasks, which subsequently may be divided into two subcategories, i.e. the technical task category, e.g. multiple choice, underline a given word class, etc., and the ability-based category, e.g. intuitive knowledge versus syllabus knowledge.

As mentioned above, the tests were designed with the main purpose of measuring the students' improvement in the areas of instruction, i.e. items whose content reflected the syllabus. This could be said to be true of items 4, 5, 7, 8, 11, 12, 13, 14, 15, 16, 17, 18,

⁷¹ For example finding adverbials in a text (item 22), classifying specific examples of adverbs (item 18), and an intuitive response in the form of a judgement (item 25).

19, 20, 22, 24 (16 out of 25 items). The remaining nine items (1, 2, 3, 6, 9, 10, 21, 23, 25) comprised a subcategory which could be said to measure the students' intuitive knowledge of the English language in that they were asked to judge the grammaticality/acceptability (see also Chapter 4.4.4.2.1) of the given English sentences. Grammaticality or acceptability judgements are standard measures in tests of implicit knowledge or proficiency tests (cf. Alderson, 1997; Ellis, R., 1991; Mandell, 1999; Odlin, 1993, & N. Ellis, 1993).

The second overall category was a content-based category which comprised six subcategories, namely word classes (items 4, 8, 11, 12, 14, 17, 18, 20, 21, 24, 25) sentence constituents (items 5, 7, 10, 13, 15, 16, 19, 22), tense/aspect (items 6, 23), morphology (items 1, 3, 9, 21, 25), constituent order (items 10, 21), ellipsis (item 2).

To some degree these categories overlapped because any given question could be categorised in a variety of ways depending on what your defined interest in it is. This was intentional in the sense that, as mentioned above, the purpose was to include a variety of items covering the same field in an attempt to avoid letting the task difficulty prevent the student from revealing his/her knowledge on a given matter.

Adverbs/adverbials, for instance, were investigated in three different types of items. In the first type of task they are given examples of adverbs and asked to determine the word class (item 18), in the second type of task they were asked to underline adverbials in a text (item 22), and in the third type of task, it was a question of being able to recognise that neither of the two given sentences was acceptable because an adjective was used where it rightly should have been an adverb (items 21 and 25). Thus items 21 and 25 involved an ability to recognise the difference between adverbs and adjectives and in these two items this was further complicated by the problem of constituent order.

Adjectives as such were asked about in two additional types of tasks, one in which subjects were asked to underline adjectives in a text and one task in which they were given four adjectives, albeit of different origins since one of them is a present participle, and asked whether these words belong to the same word class. Subjects were not asked

about identification of the word class, only about sameness of word class. The difference between adjectives and adverbs is probably one of the most difficult things for a Dane to learn where word classes are concerned, and that was why the issue was included in several instances in the tests.

Adjectives and nouns also have borderline areas, in English as well as in Danish, and this was probed in item 24. In addition, noun recognition played a role in item 11 (underline in text) and in item 4 (noun/verb distinction). Recognition of adjectives was involved in items 8 (identify adjectives of a variety of derivations) and 17 (underline in text).

In contrast, the other (closed) word classes were only asked about in one type of task each: pronouns (recognise in a text, item 20), conjunctions (recognise in a text, item14), prepositions (classify examples given, item12). The reasons for this fell into two categories. On one hand it was a simple logistic matter; the test could not be too long. On the other hand, the expectation was that few students would have worked with classification of these word classes previously, and therefore they could not play too prominent a role in the questionnaire since too many zero points would skew the overall results of the tests, but they needed to be included in the test since the answers to these items would measure potential improvement in areas focused on in the instruction. For the same reasons sentence constituent items were limited to subjects (items 5, 19), direct objects (items 7, 13), indirect object (item15), and adverbials (item 22).

4.5.4.1 Composing the tests

The creation of the pre-test was more complex than that of the post-test in that the format also had to be established in the former process. It was considered desirable to limit the number of different tasks although it was not possible to unify the question format entirely because of the variety in content category. A further consideration was the concern that the students might find it too tedious if all items were exactly alike. However, one might argue that the multiple-choice items, with only one acceptable answer, would not reveal as much information as the items which were text-based and where it was easier to measure the exact improvement. The sequence of items followed

the pattern of increasing difficulty, i.e. the easier test items at the beginning and more difficult ones at the end. The final twenty-five items cover all the word classes, the major sentence constituents, tense/aspect, constituent order, morphology and ellipsis.

As far as the post-test was concerned, the tasks were less complex in that the format was established already and consequently only two drafts were drawn up. Thus the changes were purely concerned with text editing.

4.5.4.2 Underlying principles

It was an essential guiding principle to make the pre-test and the post-test as alike to each other as possible, in format as well as in level of difficulty. As can be seen from the above, this has meant that some texts needed editing, i.e. elements either had to be changed, added to, or removed.

The adjectives asked about in item 8 (i.e. do word 1, word 2, word 3, word 4 belong to the same word class?) could be open to discussion. Since some grammarians could be justified in maintaining that *interesting* (pre-test) as well as *fascinating* (post-test) are present participles, they rightly belong to the word class of verbs. The original argument for including them was first, that the students should be presented with adjectives of diverse roots and still be able to find similarities between them despite their different forms, and second, that the subjects of the study were all first term rather than advanced students of grammar, and therefore the discussion hinted at above would hardly apply.

A further comment could be relevant for items 1 and 9, which present the issue of subject-predicator concord. A simple constellation of a plural/singular noun marked +/-s followed by plural/singular verb was considered too easy based on the argument that all or nearly all students presumably would have no difficulty in finding the correct version; such an item would bring forth very little information. In order to increase the information value of the items, they were composed with a little more complexity. Item 1⁷² has a sentence subject which is a compound unit whose two conjoints are noun

⁷² The item in the pre-test was *The book and its author are well-known by the public* versus *The book and its author is well-known by the public*. The corresponding item in the post-test was *The boy and his father like to eat at Macdonald's* versus *The boy and his father likes to eat at Macdonald's*.

groups, but each group has a singular head noun. The complexity is compounded in the second group whose determiner is the singular possessive pronoun *its*. The issue of number and concord was further complicated in item 9 of the post-test where the subject was the relative pronoun *that*, whose antecedent (*The two features of life in Britain*) is a noun group whose head is a plural noun, but its prepositional complement contains a singular noun whose prepositional complement again contains a singular noun. It would require a highly developed understanding of the principle of S/P concord for students not to be misled by the attraction of two singular nouns and a number neutral relative and to correctly assess the complexity of this structure⁷³. The item therefore has high information value (see also discussion in Chapter 4.4.3).

The same complexity was not present in the equivalent item in the pre-test. Here the subject of the relative clause is *who*, whose antecedent is a plural noun (*boys*). However, the pre-test item as well as the post-test item requires some degree of sophistication with respect to the understanding of the S/P concord issue, albeit of such a character that the students should be able to draw on their intuitive judgement in this type of task to a greater extent than they would be able to do in the syllabus-based items.

The motivation for including test items not exclusively based on syllabus knowledge was founded on the desire to explore the subjects' level of linguistic ability in general at the initial stage, but also on the idea that a test based exclusively on syllabus-based knowledge would have a chilling effect on the students' willingness to participate in the experiment, as explained above. Furthermore, it was my intention to use this group of items to investigate the query about whether gains in syllabus knowledge would have an effect on the general linguistic awareness as expressed in these items (items 1, 2, 3, 6, 9, 10, 21, 23, 25).

4.5.5 Reliability

There are two issues at stake. First, can the tests be relied on to measure the

⁷³ Ruin, who studied errors made by Swedish students of English, points out that subject/verb concord errors “appear when, for various reasons, the plural or singular status of the subject is apt to be misunderstood “ (Ruin, 1996:80), and she further points out, with reference to an example in Svartvik and Sager (1997:345) that “these errors have to do with linguistic sophistication or rather ‘language awareness’” (Ruin, 1996:82).

stipulated issues of investigation? Second, can the tests be said to constitute two equal measurements, the results of which can be the basis of a reliable comparison? In other words, do the measuring instruments constitute a truthful reflexion of the actual effect of the treatment received by the groups in question?

The first question is not an easy one to answer in absolute terms. As explained in the subsection above on how the test was composed, the construction of the test was a result of deliberations on precisely the question of how best to measure the subjects' initial knowledge and the subsequent acquired knowledge in the same test format. The tests did show that substantial gains had been achieved in the treatment groups, but it is difficult, of course, to say whether the improvements measured are expressions of gains achieved through the treatment or whether this gain has been achieved through other means, adeptness at test-taking being one. In this perspective it is not too worrisome that the post-test may be just a fraction more difficult than the pre-test (see Tables 4.5, 4.6, 4.7 and 4.8).

An issue which affects all tests is the question of which knowledge the students draw on to complete the test. The tests used in this experiment were designed to draw on two types of knowledge. The syllabus question items were designed to draw on explicit knowledge, and the non-syllabus item to draw on implicit knowledge. Whether this is what the students did would be hard to ascertain. The non-syllabus item could be completed by drawing on intuitive notions about correct and incorrect, about what would be seen to be in coherence with English language use. This could also be described as intuitive guessing. Whether the students actually behaved in this way, or whether they were trying to apply rules to these items cannot be revealed in a quantitative measure. Reversely, the syllabus items were designed to tap into the students' explicit or conscious knowledge. However, they may well have used a strategy of guessing, especially when unsure of their own knowledge. Dienes and Perner (2002) discuss this issue and suggest that one way to probe into this would be to have students declare their confidence in their choice. For instance if a student declared an expression grammatical or ungrammatical and at the same time declared whether she "knew" or "guessed" this, then a pointer could be given as to the nature of the

knowledge tapped into (Dienes and Perner (2002:83). However, some students will declare they are guessing even when performance curves demonstrate that they perform above chance.

The question concerning the reliability of the two tests in relation to each other with regard to level of difficulty was sought answered through a pre-experimental testing session involving two groups of students: one group who were given the tests in the pre-test then post-test order, and one group who were given the tests in reverse order. The reliability of the assessment of the results of the experimental tutorials would make it paramount that no substantial difference in the degree of difficulty between the pre-test and the post-test could be demonstrated. The pre-experimental testing of the pre-test against the post-test level of difficulty will be described below.

4.5.5.1 Evaluating pre-test against post-test

In an attempt to evaluate the level of difficulty of the pre-test and the post-test, respectively, a group of former students who had received the same instruction a year previously were asked if they would assist in the pre-experimental test, and eight students volunteered as test subjects. Thus it was possible to arrange the testing session with two groups of four students each.

The students were all fourth term students, and all were female. After their second term of studies these students passed an exam in which they were asked to analyse sentences to word level in accordance with the syntactic analysis that all my test subjects were taught during the weeks of experimental tutorials. The set textbook was also the same.

These students were not told about the purpose of the test, nor were the two groups aware that they were given the tests in different orders. The information they received was that it was research related and that the purpose was to measure the time necessary to complete the tests. They were also told that they could work in their own time and that there was no time limit. The students were already familiar with the subject-matter and the terms used in the tests and were an ideal group for whom there should be no apparent reason to react to the tests in a differentiated matter.

Table 4.7**Pre-experimental test of easiness of pre-test and post-test. Total result.**

subject	pre-test	post-test
T1 pre/post	20.97	22.09
T2 pre/post	18.39	16.84
T3 pre/post	15.60	16.72
T4 pre/post	16.93	13.75
T5 post/pre	16.39	13.69
T6 post/pre	17.65	18.46
T7 post/pre	16.90	15.81
T8 post/pre	18.50	19.53
Mean	17.67	17.11

Table 4.8**Pre-experimental test of easiness pre-test and post-test. Detailed results.**

subject	pre-test	post-test	gain points	gain percent	exam mark
T1 pre/post	20.97	22.09	1.12	5.34	9
T2 pre/post	18.39	16.84	-1.55	-8.43	9
T3 pre/post	15.60	16.72	1.12	7.18	6
T4 pre/post	16.93	13.75	-3.18	-18.79	9
Mean	17.97	17.35	-0.62	-3.46	8.25

subject	pre-test	post-test	gain points	gain percent	exam mark
T5 post/pre	16.39	13.69	2.70	19.72	7
T6 post/pre	17.65	18.46	-0.81	-4.39	10
T7 post/pre	16.90	15.81	1.09	6.89	9
T8 post/pre	18.50	19.53	-1.03	-5.27	8
Mean	17.36	16.87	0.49	2.89	8.5

These students had already passed their exam so they would feel secure in that knowledge. The exam results which these students achieved seven months prior to this testing session were distributed with one 10, four 9's, one 8, one 7, one 6, which makes for an average score of 8.37. The average exam score mark of the class they were a part of when they took the exam was 8.19. There were substantial individual differences in overall performances and in gains, the results of which ranged from a gain of 2.70

points to a regression of -3.18 points (see Table 4.8). These individual variations were to be expected.

The students cannot be said to have been randomly selected in the sense that no randomisation techniques were used in the selection process since they volunteered their assistance. However, they can be said to be representative of their class as a whole since their average score at the exam was basically the same. The volunteer students represented a spectrum of abilities from pass (6) to very good (10) so they do represent a wide spread in ability even though the very top and the failed sections were missing. Furthermore, two students were of non-Danish origin, in line with the characteristics of some of the subjects in the later experiment proper on which this thesis is based.

A remarkable feature in the comparison of the pre/post group with the post/pre group was that in both groups half the subjects progressed and the other half regressed. The two groups followed the same pattern here, and irrespective of which test was given first, some subjects found the second test more difficult than the first one, and vice versa. The regression was larger in the pre/post group (-4.73 points in total) than in the post/pre group (-1.84 points in total). The progression in the pre/post group was 2.24 points in total and in the post/pre group 3.79 points in total. This could indicate that the pre-test might have been slightly easier than the post-test. This could be said to be supported by the fact that the mean points achieved for the pre-test (17.67) were higher than the mean points achieved for the post-test (17.11) (see Table 4.7). However, the mean scores of the two groups were indeed close to each other. In the pre/post group there was a difference from pre-test to post-test of -0.62 points. In the post/pre group there was a difference from post-test to pre-test of 0.49 points (see Table 4.8).

The graphs below visualise the score points and illustrate how close the two lines are to becoming one line. Figure 4.3 illustrates all the pre-test scores against all the post-test scores. As can be seen, the post-test appears to be minutely more difficult than the pre-test.

The figures and thus the graphs seem to show that it makes very little difference whether the pre-test comes before the post-test or vice versa (see Figures 4.3 and 4.4). The performance of these eight pre-experimental subjects could indicate that to these subjects the pre-test and the post-test seemed to be equally difficult since the scores showed little variation. The subsequent question was whether it would be justified to assume that these eight subjects could be said to constitute a representative sample of the student population as such.

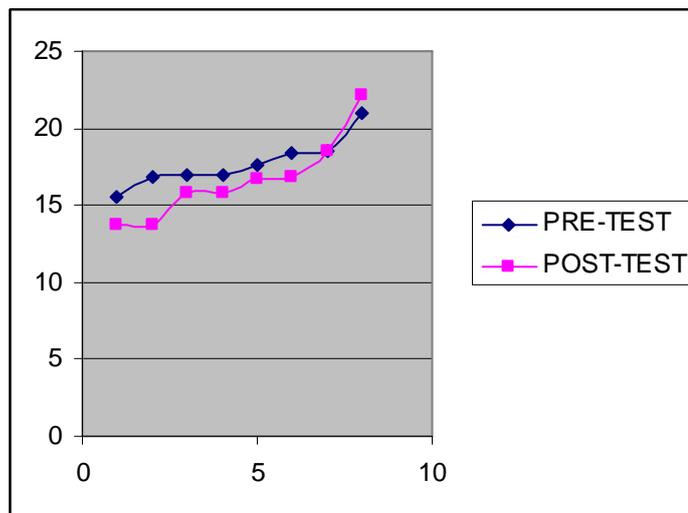


Figure 4.2 Overall comparison pre-test against post-test. Pre-experimental.

In other words, could this result be said to be valid for the students in the experiment proper? Their number was quite small, which statistically makes it difficult to claim generality for the results. In favour of the generalisation of results could be the fact that these eight subjects represent a spectrum of high achievers as well as low achievers. Although the group did not encompass the absolute top nor the absolute bottom (i.e. none of them failed the exam), they did cover a wide range in ability if exam results can be said to represent ability. Despite the individual differences that always persist, it would seem that most subjects would find the two tests of similar difficulty.

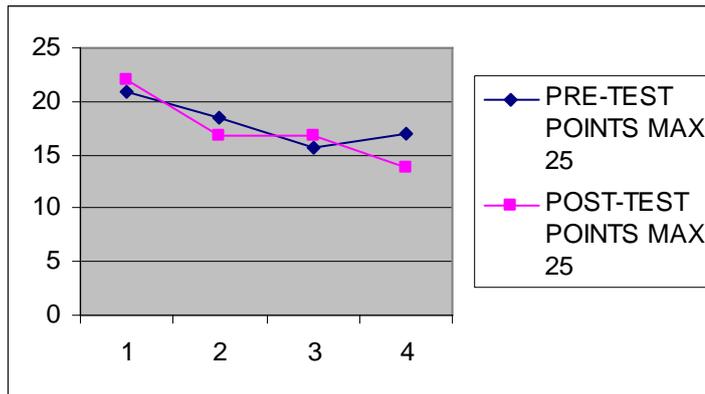


Figure 4.3 Pre-test first. Point score total by subject. Pre-experimental.

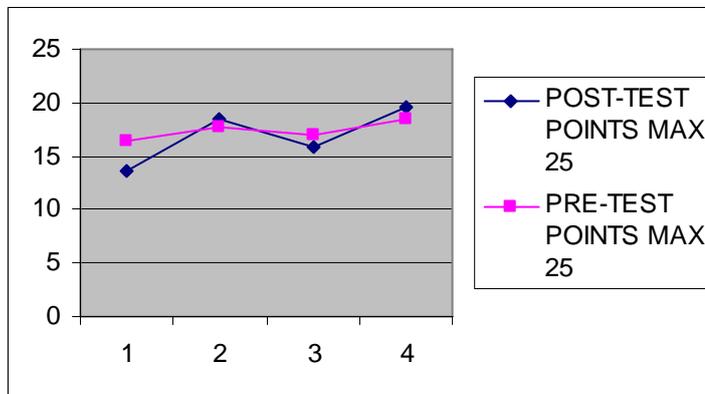


Figure 4.4 Post-test first. Point score total by subject. Pre-experimental.

4.5.5.2 Rating the tests

Test design and assessment are closely correlated. Assessment has its own inherent problems, but these can either be enhanced or alleviated by the test design. The aim in the design of the tests administered in this study was to make the rating of the tests as uncomplicated as possible and to reduce the possibility for interpretation and discussion. One way in which this was sought implemented was to incorporate as many items as possible which could be said to have only one acceptable answer. This applied to all the items whose form is a recognisable multiple choice (items 1, 2, 3, 4, 5, 6, 8, 9, 10, 12, 13, 15, 18, 21, 23, 25), but also other types of items could be said to be disguised forms of multiple choice items (7, 19, 24) in which a range of choices would be open but only one answer would be acceptable, e.g. when the students were asked to look for a given sentence constituent in a sentence containing any number of constituents. These

test items were all evaluated with either one point (acceptable answer) or zero points (unacceptable answer).

The remaining test items were all text-based items of the type: Underline in text (items 11, 14, 17, 20, 22, 24). The number of items varied from question to question, i.e. from thirteen to three. For any correct item in a given case the student would achieve 1/n point, and this would then be multiplied by the number of acceptable items underlined. No matter how many items a fully answered question contained, it was only possible to achieve one point per question. Another way of doing it would have been to give each item one point; then the points for these items would have varied from thirteen to three. This would have meant an imbalance in the weighting of the items which would not necessarily be founded in the level of difficulty of the structure in question, and in order to avoid skewing the overall result of the experiment the former procedure was chosen. Another method sometimes applied in recognised professional tests is to decide on a breaking point above which one point is given and achievement below which will be rated at zero. The breaking point is often set at 50 % (Kim, 2007). This could have been an acceptable method for this type of item, but the more precise measure of fractions was adopted for reason of information value.

In hindsight there ought perhaps to have been more items of the text-based type and fewer of the multiple-choice type. The text-based type has the advantage that any improvement achieved by the students can be monitored in detail, and therefore this type of test item is better from an informative point of view. But it is not an entirely unproblematic type as it poses other problems, primarily because the rater occasionally has to make difficult decisions with regard to accepting/not accepting underlinings that are dubious or ambiguous, and therefore open to interpretation. This might affect the reliability of the assessment. In item 17 of the post-test (underline all adjectives), for instance, there is the following noun group: "a gay, dashing sort of fellow with dark, romantic-looking eyes and black curly hair", which faced the rater with two problems. The first was fairly simple and concerned the question of whether an unbroken underlining of "gay, dashing" should count as one or two underlined adjectives. The rater decided to count them as two. The second problem concerned "romantic-looking".

Some students underlined "romantic" only, and is that a valid answer? The rater decided to judge this an invalid answer.

The same type of problems arose in item 19 of the post-test (underline subjects). The third subject is "the friends and family of the bride and groom". How much should be underlined for it to count as a valid answer? The rater decided that an underlining of only "the friends and family" would be an invalid answer, but accepted an underlining of "the friends and family of the bride" as a valid answer. In the equivalent item in the pre-test there are two complex subjects: "Dickens' extreme energy" and "His weekly journalism". Underlinings of only "energy" and "journalism" were not accepted as valid answers.

Adverbials pose a particular problem. Many adverbials are prepositional groups, but not all prepositional groups are adverbials in that they are actually dependents in a noun group. In item 22 of the pre-test the following noun groups: "there are gardens with trees in them" (no adverbial), "the people of New Street" (no adverbial), and "talk to the people in the houses next to theirs" (no adverbial) might lead to subjects perceiving the prepositional dependents as adverbials.

The issue in item 14 of the pre-test (underline all co-ordinating conjunctions) is of a principle character pertaining not just to this item but to all items. What is special about this item is that the text contains only three co-ordinating conjunctions, but most students underline more than three words, some of them subordinating conjunctions rather than co-ordinating conjuncts; also other word classes were underlined. The principle question then is whether only acceptable underlinings should be counted and the rest ignored, or should unacceptable underlinings in some way distract from the validity of the answer? The rater opted for the first strategy in order to avoid the paradoxical situation in which a student would get negative points even though all three co-ordinating conjunctions were underlined. The same principle was adhered to in all text-based items.

4.5.5.3 Inter-rater reliability

As may have become apparent from above, the concern for reliability was present from the outset, but as may also have become apparent, the format of the tests also left open the possibility that raters might make decisions from different decisions on the interpretation of acceptability, which again may be motivated by different perceptions of the grammatical issues involved.

Table 4.9
Inter-rater reliability. Pre-test and post-test.

Pre-test			
Rater	R1	R2	R3
R1		0.99	0.99
R2	0.99		1.00
R3	0.99	1.00	

Post-test			
Rater	R1	R2	R3
R1		1.00	1.00
R2	1.00		0.99
R3	1.00	0.99	

In order to measure this potential deviation in rating points a limited inter-rater reliability test was conducted between three raters (see table 4.9). The inter-rater reliability proved very high for the pre-test as well as for the post-test. This should be related to the aim of the test development which was to make the rater's task involve as little interpretation as possible.

4.6 Results of the quantitative study: statistical description

The experiment contained three groups which received different treatments: BASIS groups, NON-VISL groups, and VISL groups. The basis groups and the NON-VISL

groups received a traditional 'chalk and talk' type of instruction, and the difference between the basis groups and the NON-VISL groups was one of degree rather than one of content or method in that the NON-VISL groups received one hour more per week than the basis groups. The difference in treatment from the NON-VISL groups to the VISL groups was of a qualitative nature in that the VISL groups were given computer-assisted instruction using the VISL programme (see Figure 4.5).

Instructional groups	BASIS	NON-VISL	VISL
Traditional	X hours	X+1 hours	
Computer			X+1 hours

Figure 4.5 Hours of training per experimental group

At the outset, i.e. at the planning stage of the assignment to the various groups, the groups were identical in size, and the reason why the number of subjects varies in the ensuing statistics is the simple fact that students dropped out as the experimental period unfolded, in other words the statistical material encompasses only the subjects who were present at both the pre-test and the post-test. Obviously, dropouts were not included in the post-test results, but they were not included in the pre-test results either.

**Table 4.10
Number of subjects per treatment per cohort**

Student	BASIS	NON-VISL	VISL	Total
Cand.Negot.	15	20	18	53
English	21	15	18	54
Total	36	35	36	107

This information is a necessary pre-condition for understanding the reported difference in the number of subjects in the various groups and for evaluating any reported variation in pre-test levels within the study programmes, and the subsequent differences from pre-test to post-test scores as well as differences in success rates are based on this underlying condition.

4.6.1 Total test item results

4.6.1.1 Total pre-test levels

4.6.1.1.1 The Cand. Negot. cohort. Totals pre-test.

The Cand. Negot. group contains 53 subjects in all, distributed in the three experimental groups with 15 in the basis group, 20 in the NON-VISL group, and 18 in the VISL group (see Table 4.10). The total pre-test results for these three experimental groups ranges from 10.5 points (NON-VISL) over 10.8 points (BASIS) to 11.1 points (VISL). In other words none of the groups achieves a mean above half of the 25-point maximum (see Table 4.11).

The proximity of the mean scores is clear from the standard deviations which are 3.18, 3.21, and 3.49, respectively. The standard deviation is a measurement of how many observations are at variance with the mean value. On average the mean pre-test value for the Cand. Negot. group is 10.8 points the corresponding average standard deviation is 3.25. This means that with a confidence level of 95% the mean spread of observations lies within $(-3.25 + 10.8)$ and $(10.8 + 3.25)$, that is between 7.55 and 14.05 points out of a maximum of 25 points.

Table 4.11
Cand. Negot. cohort. Pre-test levels. Point scores. Total (max. 25).

Treatment	N	mean	sd	min	p25	p50
BASIS	15	10.80	3.18	6.67	8.17	10.00
NON-VISL	20	10.50	3.21	4.62	8.60	9.91
VISL	18	11.10	3.49	5.00	8.45	11.20
Total	53	10.80	3.25	4.62	8.45	10.40

Treatment	p75	max
BASIS	13.60	16.70
NON-VISL	12.60	18.90
VISL	12.90	18.50
Total	12.90	18.90

If we assume that the sample follows the normal distribution, then 68% of all observations should fall between mean value -1 sd and mean value $+1$ sd., in casu between 7.55 and 14.05 points. This means that the bulk of observations fall within this range, but there will be some observations higher or lower, which can be read from the indications of minimum and maximum (see Table 4.11).

Interestingly, the BASIS group has the highest minimum score (6.67), but the lowest maximum score (16.70) of all the groups. The NON-VISL group has the middle ranking minimum score (4.62) but the highest maximum score (18.9); the VISL group has the middle ranking minimum score (5.0) and the middle ranking maximum score (18.50), i.e. the spread is smallest in the BASIS group in which 75% of all observations are between 6.67 and 13.60 points or below. The corresponding range for the NON-VISL group is between 4.62 and 12.60 points, and for the VISL group the range for 75 % of the observations is between 5.00 and 12.90 points. The standard deviations for all three treatment groups are close, and so are the means. It may be worth noticing, however, that the mean minimum score of 4.62 for the Cand. Negot. as a group is very low and that the mean score for 75 % of the observations is 12.90 points, which means that only 25 % of the subjects in the Cand. Negot. group are able to score above chance level at the outset of the experiment.

4.6.1.1.2 The English cohort. Totals pre-test

The total number of English subjects was almost the same as that for the Cand. Negot. cohort, namely 54 compared to 53 (see Table 4.10). The distribution was 21 in the BASIS group, 15 in the NON-VISL group and 18 in the VISL group.

The mean pre-score was 13.10 points, but the means of the three groups were unequally distributed, ranging from 14.20 in the BASIS group, 11.90 points in the NON-VISL group and 13.00 in the VISL group. This diversity was also reflected in the standard deviations. The mean standard deviation was 2.75, which is an expression of a variation from 3.07 in the BASIS group, 1.73 in the NON-VISL group and 2.75 in the VISL group. The largest spread was in the BASIS group where, following the normal distribution, 68% of the observations lie between $(-3.07 + 14.20)$ and $(14.20 + 3.07)$,

that is between 11.13 and 17.27 points. The NON-VISL group equivalent range is $(-1.73 + 11.90)$ and $(11.90 + 1.73)$, or from 10.17 to 13.63 points. The comparable VISL spread is $(-2.75 + 13.00)$ and $(13.00 + 2.75)$, or from 10.25 to 15.75 points.

Table 4.12
English cohort. Pre-test levels. Point scores. Total (max. 25).

Treatment	N	mean	sd	min	p25	p50
BASIS	21	14.20	3.07	8.09	12.30	13.90
NON-VISL	15	11.90	1.73	8.38	10.50	12.30
VISL	18	13.00	2.75	7.53	11.40	12.30
Total	54	13.10	2.75	7.53	11.30	12.90

Treatment	p75	max
BASIS	16.30	19.50
NON-VISL	13.40	14.70
VISL	14.80	19.50
Total	15.00	19.50

A closer look at these figures reveals that the spread in the BASIS group takes place at a high level; the minimum score is 8.09 and the maximum score is 19.50 points, which is very high, and the same maximum score as for the VISL group, but the general level is higher in the BASIS group, which can be read from the fact that 75% of all subjects in the BASIS group had a score of 16.30 or below where the figure for the NON-VISL group is only 13.40 points. The spread in the NON-VISL group is small in comparison, although the minimum score (8.38 points) is slightly higher than that in the BASIS group; the maximum score is only 14.70 points, which gives the NON-VISL group a low mean value of 11.90 points. In graphic terms the curve for the BASIS group would be relatively flat and broad where the NON-VISL curve would be higher and narrower. Somewhere between these two groups is the VISL group with 7.53 as the lowest score of all English observations but also a maximum score of 19.50 points, which equals the highest score in the BASIS group. The figures for the various quartiles reveal that the VISL and the NON-VISL groups are similar to each other in that for both groups 50% of all observations are below and 50% are above 12.30 points, with the

difference being that for the upper quartile the score is higher for the VISL group observations.

4.6.1.1.3 VISL, NON-VISL and BASIS (Total pre-test) differences

An analysis of the comparative pre-test levels of the two cohorts is best illustrated through percentage rates (see Tables 4.13 and 4.14). These rates are the percentage expression of the raw scores in Tables 4.11 and 4.12 and will be referred to as the pre-test success rates. When comparing the pre-test success rates of the two participant cohorts, it becomes clear that the English cohort has a higher initial mean level (52.60 %) than the Cand. Negot. cohort (43.10 %).

Table 4.13
Cand. Negot. cohort. Pre-test levels (percentage success rates). Total (%).

Treatment	N	mean	sd	min	p25	p50
BASIS	15	43.20	12.70	26.70	32.70	40.00
NON-VISL	20	41.90	12.80	18.50	34.40	39.60
VISL	18	44.50	14.00	20.00	33.80	44.80
Total	53	43.10	13.00	18.50	33.80	41.70

Treatment	p75	max
BASIS	54.40	66.80
NON-VISL	50.30	75.70
VISL	51.40	74.20
Total	51.40	75.70

The intra-cohort spread tendencies, which were registered in the raw scores, become obvious when comparing the percentage expressions of the standard deviations. The Cand. Negot. cohort is fairly homogeneous. The BASIS and the NON-VISL groups are almost identical with 12.70 and 12.80, respectively. There is a minimally higher spread in the VISL group (14), but all three groups are close to a mean of 13, i.e. Cand. Negot. observations on average fall between $(-13 + 43.10)$ and $(43.10+13)$, which following the normal distribution, means that approximately 68% of all observations have pre-test total success rates ranging from 30.10% to 56.10%.

Table 4.14
English cohort. Pre-test levels (percentage success rates). Total (%).

Treatment	N	mean	sd	min	p25	p50
BASIS	21	56.60	12.30	32.40	49.30	55.40
NON-VISL	15	47.80	6.94	33.50	42.10	49.20
VISL	18	51.80	11.00	30.10	45.70	49.40
Total	54	52.6	11.00	30.10	45.30	51.60

Treatment	p75	max
BASIS	65.30	78.10
NON-VISL	53.80	58.60
VISL	59.00	78.20
Total	59.90	78.20

The Cand. Negot. rates are considerably lower than the corresponding English rates which range from $(-11 + 52.60)$ and $(52.60+11)$, i.e. from 41.60% to 63.60%. The general pre-test levels indicate that the Cand. Negot. cohort starts off at a lower initial level of knowledge and proficiency than the English cohort. The English cohort is characterised by noticeable differences in standard deviations between the cohort groups, which was distinguishable in the raw scores, but it becomes very apparent in the transformation into pre-test total success rates; especially the much lower standard deviation in the NON-VISL group whose sd is 6.94 compared to 11.00 in the VISL group and 12.3 in the BASIS group. The English NON-VISL group has a higher minimum and a lower maximum success rate than the two other English groups and thus a smaller span within which the observations lie.

The higher level of the English cohort is also expressed in the pre-test total success rates for the various quartiles which clearly show that the success rates of approximately 50% in the Cand. Negot. cohort lie in the 75% quartile but in the 50% quartile for the English cohort, which is to say that the bulk of observations lies a quartile lower for the Cand. Negot. observations than for the English observations.

The BASIS groups of both cohorts have relatively high minimum rates, and the English

BASIS group also has high maximum, i.e. the concentration of observations fall at the high end of the scale (see Figure 4.6). The Cand. Negot. BASIS group has a low maximum, i.e. little spread and a concentration at the low end. The NON-VISL English group has a concentration of observations at the low to medium band of the scale, and little spread. The Cand. Negot. NON-VISL group observations spread from low to high, but the group has a low mean value. The VISL English group has observations at the medium to high end of the scale whereas the Cand. Negot. VISL group observations are concentrated at the low to medium band of the scale but with a low maximum.

4.6.1.2 Total post-test levels

4.6.1.2.1 The Cand. Negot. cohort. Totals post-test

After ten weeks of treatment, the Cand. Negot. cohort shows a clear ranking in the post-test results of the three groups with the means gradually ascending from the basis group mean score of 14.30 points, over the 16.70 points in the NON-VISL group, to the VISL group's score of 18.30 points (see Table 4.15).

These results are underscored by the very similar standard deviations in the three groups – 3.44, 3.23 and 3.47, respectively. The different treatments seem to have affected the three experimental groups in different ways, and the differential effect seems to indicate that the most effective treatment for Cand. Negot. students is the VISL treatment. Not only is the mean score in this group the highest, it is also in this group that we find the highest maximum score of 23, which is a very high score in absolute terms (max. possible 25). The average mean score, 16.50, and the median as expressed in the 50% percentile, 16.30, are almost identical, which is a good indication that the observations are normally distributed.

It is interesting to see that the ranking as expressed in the mean values is reflected in the maximum score of the three groups, 19.50 points, 22.20 points, and 23.00 points, respectively. The ranking is also clearly detectable in the quartiles, and the minimum and maximum scores. These score figures reveal that the 50% score level is reached by even the lowest scoring subjects the VISL group (minimum score is 12.70 points in this group), for the NON-VISL group the 50% score mark is reached in the 25% quartile, and for the BASIS group it is achieved somewhere between the 25% and the 50%

quartile, but even the 25% quartile is close to the 50% level with a score of 11.80 points.

Table 4.15
Cand. Negot. cohort. Post-test levels. Point scores. Total (max. 25)

Treatment	N	mean	sd	min	p25	p50
basis	15	14.30	3.44	7.55	11.80	14.00
NON-VISL	20	16.70	3.23	8.38	14.90	16.30
VISL	18	18.30	3.47	12.70	15.00	19.10
Total	53	16.50	3.67	7.55	14.30	16.30

Treatment	p75	max
BASIS	17.00	19.50
NON-VISL	18.60	22.20
VISL	20.90	23.00
Total	19.50	23.00

There is hardly any difference between the BASIS group score of 18.00 points and the VISL group score of 17.80 points, and the NON-VISL group score is only slightly lower at 17.10 points.

4.6.1.2.2 The English cohort. Totals post-test

The mean scores of the three groups are very close in the post-test (see Table 4.16). The equality of the groups is supported by standard deviations which are close to the mean of 3.01 points, which indicates that all three groups have similar spreads in the observations. The differences between the three groups manifest themselves in the minimum and maximum scores. The VISL groups have by far the highest minimum score with 12.50 points which equals the 50% achievement level, which is very high as a minimum score. The NON-VISL group trails behind with 9.92 points followed closely by the BASIS group whose minimum score is 9.66 points. The VISL group also scores the highest maximum, 23.00 points, with the BASIS group at 22.50 and the NON-VISL group at 21.70 points. These scores show a clear advantage to the VISL group over the

NON-VISL group but hardly any difference from the BASIS group, and this distribution

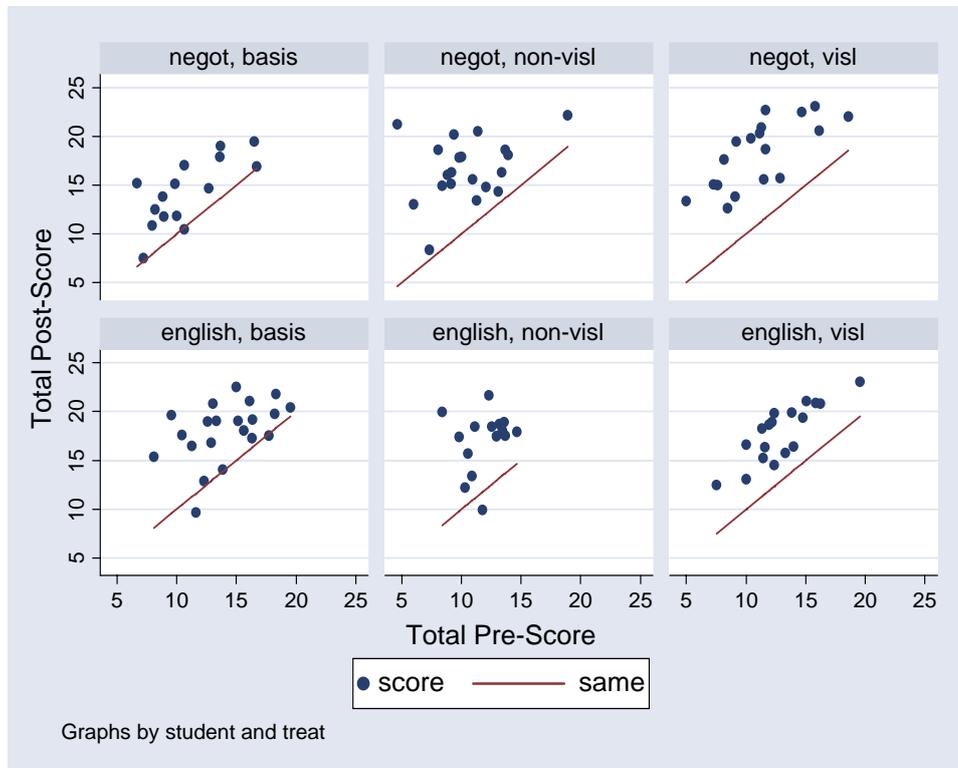


Figure 4.6 Scatterplot. Total scores pre-test by post-test per experimental groups.

also holds for the 75% quartile, but although the scores show the same ranking, they are so close that the difference between the VISL group, 19.90 points, and the BASIS group, 19.80, is negligible, but distance to the NON-VISL group, 18.70 points, is more substantial.

The BASIS group's score is high all the way through, but a higher number of the VISL group subjects are brought to the highest level as expressed by the figures for the upper quartile and the maximum scores.

As mentioned earlier the mean figures are similar for all three groups, but the quartiles reveal that the NON-VISL group has difficulties in keeping pace with the VISL group.

Table 4.16
English cohort. Post-test levels. Point scores. Total (max. 25).

Treatment	N	mean	sd	min	p25	p50
BASIS	21	18.00	3.09	9.66	16.80	19.00
NON-VISL	15	17.10	3.06	9.92	15.70	17.90
VISL	18	17.80	2.95	12.50	15.70	18.40
Total	54	17.70	3.01	9.66	16.30	18.20

Treatment	p75	max
BASIS	19.80	22.50
NON-VISL	18.70	21.70
VISL	19.90	23.00
Total	19.80	23.00

4.6.1.2.3 VISL, NON-VISL and BASIS (Total post-test) differences

The English subjects achieve a higher rate of success in the post-test than the Cand. Negot. subjects do with a mean English success rate of 70.8% (see Table 4.18) versus 66.10% for the Cand. Negot. subjects (see Table 4.17).

The treatment shows a clear differential effect in the Cand. Negot. cohort, i.e. the VISL group with a success rate of 73.10% clearly outperforms the NON-VISL group which with its success rate of 66.70% . The NON-VISL Cand. Negot. group clearly outperforms the BASIS group which has a success rate of only 57.10%, only slightly above chance level. Equally differential effects are not detected in the English cohort although here, too, the VISL group with its 71.40% success rate outperforms the NON-VISL group which only achieves a success rate of 68.20%, but the BASIS group outperforms them both, albeit only minimally with a success rate of 72.00%. The reason for the high success rate in the BASIS group may lie in the fact that the BASIS group subjects were high-achievers from the outset.

The VISL groups in the English cohort as well as the Cand. Negot. cohort score the highest maximum success rates, which in both cases is 92.20%. For both cohorts it is

also clear that the VISL groups significantly outperform the NON-VISL groups. It is notable that the VISL groups of the Cand. Negot. cohort as well as the English cohort have all of their observations above the 50% mark. It seems that the VISL treatment is able to raise the bottom sections to a higher level than the other treatments. The NON-VISL groups are similar in their success rate levels with regard to the top 25 % of

Table 4.17
Cand. Negot. cohort. Post-test levels. Success rates. Total (%).

Treatment	N	mean	sd	min	p25	p50
BASIS	15	57.10	13.80	30.20	47.10	58.80
NON-VISL	20	66.70	12.90	33.50	59.50	65.20
VISL	18	73.10	13.90	50.60	60.20	76.20
Total	53	66.10	14.70	30.20	57.40	65.20

Treatment	p75	max
BASIS	68.20	77.80
NON-VISL	74.40	88.60
VISL	83.60	92.20
Total	77.80	92.20

Table 4.18
English cohort. Post-test levels. Success rates. Total (%).

Treatment	N	mean	sd	min	p25	p50
basis	21	72.00	12.40	38.60	67.30	76.00
NON-VISL	15	68.20	12.30	39.70	62.90	71.80
VISL	18	71.40	11.80	50.00	63.00	73.80
Total	54	70.80	12.00	38.60	65.30	72.70

Treatment	p75	max
basis	79.20	90.00
NON-VISL	74.80	86.70
VISL	79.60	92.20
Total	79.20	92.20

observations, but the English group has a higher average for the bottom end of observations. The BASIS group for the Cand. Negot. cohort remains lowest in the hierarchy whereas the BASIS English group observations has a top similar to the VISL group, but the size of the English BASIS group may also mean that the comparison of the means of these two groups may be slightly skewed. Indeed, the English BASIS group does have lower observations than the VISL group and a larger spread (see Figure 4.6).

4.6.1.3 Total gain levels

Gain levels can be measured by two different methods, namely in relative differences (Tables 4.19 and 4.21) and in percentage point differences (Tables 4.20 and 4.22). The relative difference is tabulated by subtracting the pre-test total score from the post-test total score; subsequently the rest sum is divided by the pre-test total score. The percent difference is the pre-test percent success rate minus the post-test per cent success rate. Both methods take the pre-test levels into account, albeit that the relative difference method measures the gain in a directly readable scale against the starting point of the pre-test level whereas the percent gain method can be said to be a simple measure of absolute increases in success rate. The percentage point differences (=gains) are thus more immediately comparable.

For example, subject 1, who increases from a success rate from 5% to a success rate of 10%, has an increase of 5% (percentage points). Subject 2, who has an increase from 10% to 15 %, also has an increase in success rate of 5% (percentage points). If the relative difference/gain of the subjects was measured it would be post-test score minus pre-test score, divided by pre-test score, i.e. in the case of subject 1 an increase of 100% and in the case of subject 2 an increase of 50 %. For illustrative and comparative reasons both sets of gains tables are included and described here, but this is not the case for the syllabus and the non-syllabus gains descriptions. As far as the syllabus and the non-syllabus sections are concerned only the percentage point gains are included, as the comparison below did not prove any substantial differences in relations or clarity of explanation that would make it advantageous or paramount for it to be included.

4.6.1.3.1 The Cand. Negot. cohort. Total gains

4.6.1.3.1.1 Relative gains. Total

When looking at Table 4.19, one notices that the gap between the average mean, .631, and the average 50% quartile increase, .53, is so large that the observations may not be normally distributed, which could indicate that there is a difference between the groups. A closer look will reveal a marked difference in means between the BASIS group and the two other groups. The BASIS group increase of .365 is approximately only half of the NON-VISL increase of .741 and the VISL increase of .731. The reason for this could be that the starting point of the BASIS group was very high and, since this is a relative difference, the figure might be an expression of the effect demonstrated in Chapter 4.5.1.3. However, as Table 4.11 demonstrates, the pre-test level of the Cand. Negot. BASIS group is at about the same level as the Cand. Negot. NON-VISL group and lower than the VISL group.

Table 4.19
Cand. Negot. cohort. Relative gain levels. Total.

Treatment	N	mean	sd	min	p25	p50
BASIS	15	.365	.324	-.0169	.158	.319
non-visl	20	.741	.768	.0972	.222	.718
visl	18	.731	.392	.188	.458	.712
Total	53	.631	.568	-.0169	.227	.530

Treatment	p75	max
BASIS	.536	1.28
non-visl	.818	3.59
visl	.969	1.67
Total	.822	3.59

The low achievement of the BASIS group must therefore be due to other factors. One such factor could be the extremely low minimum gain, which is negative, i.e. the score is lower in the post-test than in the pre-test, a result which has a very averse effect on the mean. However, the NON-VISL group minimum gain is also very low, .0972, especially compared to the VISL minimum gain of .188, and yet the NON-VISL mean

gain, .741, is approximately the same as the VISL gain, .731, which, to be exact, is actually lower than the NON-VISL gain. The 75% quartile figures reflect the same ranking as the pre-test and the post-test success rates did, with the VISL group gain at .969, the NON-VISL group gain at .818 and the BASIS group gain at .536. The reason for the equality between the NON-VISL and the VISL gain may lie in the extremely high maximum gain demonstrated in the NON-VISL group which towers at 3.59. The standard deviations tell us that the NON-VISL group has a larger spread than the other groups so the NON-VISL group observations range from very low to extremely high gains, which results in a high mean equivalent to that of the VISL group. The equality of the two groups is further segmented by the 50% quartile gain figures of .718 for the NON-VISL group and .712 for the VISL group.

4.6.1.3.1.2 Percentage points gains. Total

Table 4.20, which shows the success rate gain, supports the marked difference between the BASIS group and the other two groups in that the BASIS group mean gain is only 13.90%, followed by the NON-VISL group with a gain of 24.80%, and the VISL group on top with 28.50%.

Table 4.20
Cand. Negot. cohort. Percentage point gain levels. Total.

Treatment	N	mean	sd	min	p25	p50
BASIS	15	13.90	9.80	-.72	7.28	11.80
NON-VISL	20	24.80	15.20	4.20	12.30	25.10
VISL	18	28.50	10.20	11.60	17.80	30.30
Total	53	23.00	13.40	-.72	11.80	21.10

Treatment	p75	max
BASIS	21.10	34.10
NON-VISL	31.80	66.40
VISL	37.50	44.20
Total	31.60	66.40

This ranking is supported by all the quartiles. The standard deviations show that the BASIS, sd 9.8, and the VISL, sd 10.2, groups have similar spreads whereas the NON-

VISL, sd 15.2, group has a much larger spread than the other two, ranging from a minimum gain of 4.20% to a maximum gain of 66.40%. With this method of measuring gain the distribution is, in fact, closer to the normal distribution in that the average mean value of 23% is close to the average 50% quartile value of 21.10%, but the NON-VISL group contains some extremes in the lower end and the upper end of the distribution of the observations, and the maximum gain of 66.40% in this group is a clear example of this in that the maximum gain in the VISL group is well over 22% lower, and the BASIS group gain almost 32% lower.

4.6.1.3.2 The English cohort. Total gains

4.6.1.3.2.1 Relative gains. Total

Relative gains (see Table 4.21) in the English cohort show that the NON-VISL group has higher gains, .454, than the VISL group, .399, which in turn has higher gains than the BASIS group, .315. Part of the explanation for the low relative gain in the BASIS groups could be that this group initially had a very high score (see Table 4.12 and 4.14) compounded by the negative gain (loss) listed as the minimum of -.167. However, this is contravened by the high maximum gain of 1.05 which indeed is lower than the maximum gain of the NON-VISL group which has the highest maximum gain, 1.38, of all three groups.

Table 4.21
English cohort. Relative gain levels. Total.

Treatment	N	mean	sd	min	p25	p50
BASIS	21	.315	.314	-.167	.06	.257
NON-VISL	15	.454	.349	-.156	.228	.395
VISL	18	.399	.174	.173	.282	.369
Total	54	.381	.287	-.167	.185	.341

Treatment	p75	max
BASIS	.504	1.05
NON-VISL	.657	1.38
VISL	.563	.665
Total	.555	1.38

The low gain in the BASIS group is further illustrated by the fact that the gains of the upper 25% of the observations lie between .504 and 1.05 with the equivalent span for the NON-VISL group being .657 to 1.38.

The span, from .563 to .665, for the VISL group is even narrower than for the NON-VISL and the BASIS groups, which means that the gain of the upper quartile in the VISL group is low compared to the two other groups, but unlike the BASIS and the NON-VISL groups, the VISL group has no negative gain, i.e. the spread is much lower in the VISL group than in the other two groups, which also can be read from the standard deviations. The standard deviation of the VISL group is, at .174, noticeably lower than the NON-VISL group's of .349 and the BASIS group's .314.

4.6.1.3.2.2 Percentage points gains. Total

The English BASIS and the NON-VISL group has the exact same standard deviation, 12.8 (see Table 4.22), which is almost double that of the English VISL group so this method gives the same general picture as the relative difference, i.e. the VISL group observations are centred in a narrower band than the other two groups, mainly due to the negative gains (loss) in the English BASIS and the NON-VISL group.

Table 4.22
English cohort. Percentage points gain levels. Total.

Treatment	N	mean	sd	min	p25	p50
BASIS	21	15.40	12.80	-7.76	3.92	15.60
NON-VISL	15	20.40	12.80	-7.36	13.10	20.80
VISL	18	19.60	6.82	8.72	14.00	19.40
Total	54	18.20	11.20	-7.76	9.96	18.70

Treatment	p75	max
BASIS	25.60	40.20
NON-VISL	29.20	46.30
VISL	26.60	29.90
Total	26.60	46.30

That said, it must also be pointed out that the English cohort seems to be normally distributed since the mean (18.20%) is very similar to the 50% quartile figure (18.70). The quartiles show us that the lower quartiles of the BASIS group are very low, the minimum being a negative gain (loss) of -7.76% , with 50% of observations at or below 15.60%, and the upper quartiles are not high enough to make up for this, unlike achievements in the NON-VISL group, which also has a negative minimum gain (loss), -7.36 , almost equivalent to the BASIS group, but the NON-VISL 50% quartile proffers a gain of 20.80% or below, 5% higher than the BASIS group; add to this that the upper 25% of observations are higher than the BASIS group, i.e. between the 20.80% of the 75% quartile and the maximum of 46.30%.

The significant difference between the three groups becomes very noticeable in the upper quartiles where the English VISL group is distinguishable for its low gains and the English NON-VISL group by its high gains. It must be noted that the VISL group's standard deviation is only half that of the BASIS and the NON-VISL groups. So the English VISL observations have less spread than observations in the other two groups. The ranking in the percent gains is the same as the ranking which appears in the relative gains, with the highest gain in the English NON-VISL group, 20.40%, closely followed by the VISL group's 19.60%, which is remarkably high in comparison to the English BASIS group gain of 15.40%, and this despite the low maximum, which is a modest 29.90%, and upper quartile gains.

4.6.1.3.3 VISL, NON-VISL and BASIS (Total gain) differences

The Cand. Negot. cohort percentage point gains follow the same ranking as we saw in the post-test scores: Cand. Negot. VISL outperforms the Cand. Negot. NON-VISL group with the BASIS group straggling behind both of the others at approximately half the gain of the VISL group. This ranking is different in the relative gains in that the VISL and the NON-VISL groups have reversed their positions with the NON-VISL group achieving higher gains than the VISL group. The results of the two groups are close, though. The Cand. Negot. NON-VISL group is higher in the upper quartile but also lower in the bottom quartile indicating the extent of the spread of observations as expressed in the high standard deviation, which is so much higher than those of the

other two groups. Both methods of measuring gain find that the BASIS group gain is approximately half of the Cand. Negot. VISL group gain.

Regardless of tabulation method, the BASIS group gain in the English cohort has the lowest mean value of the three groups. The difference from the other two groups is not as marked as it is in the Cand. Negot. cohort, though, but the difference is notable. In the English cohort the NON-VISL group has the highest gain, both relatively and in per cent difference. The starting point of the English BASIS group was a relatively high pre-test level, and thus one might say that the room for improvement was less than for the other groups. However, the per cent difference to some extent takes this into account, and the English BASIS group gain is still much lower than the English VISL group's gain. The English VISL pre-test level was also high and yet the BASIS group gain is so much lower than that of the English VISL group, in other words the difference in gains between the English BASIS and the VISL group results need to find other explanations.

What emerges is that the VISL group floor of achievement for both cohorts appears to be lifted to a higher level than that of the other two groups in as much as the gains in the lower quartiles of the two VISL groups are higher than in the other groups. This and the low standard deviations for the VISL groups of both cohorts indicate that the VISL treatment is able to lift the lowest subjects to a higher extent than is the case in the other groups. The VISL groups also has high gains overall. For the Cand. Negot. group the VISL gain is higher than those of the NON-VISL and the BASIS groups; for the English VISL group the gain is higher than for the BASIS and about the same as that of the NON-VISL group.

The middle group of students seem to be doing equally well in the VISL and the NON-VISL groups of both cohorts. However, the English VISL group has a low upper quartile and a low maximum, which again is an expression of the low spread and the concentration in the English VISL group. The VISL treatment seems to be working better for the Cand. Negot. cohort than the English cohort. Reversely, the English NON-VISL group has higher gains than the Cand. Negot. NON-VISL group. The NON-VISL

English group and the BASIS groups of the English and the Cand. Negot. cohorts experience negative gains (losses) as their minimum scores. For the upper quartiles the BASIS groups again achieve less than the other two groups. The BASIS groups of both cohorts and the English NON-VISL group have large spreads of observations and larger spreads than the VISL groups. The BASIS groups are not able to achieve the same gains for the middle group of students as the other two treatments are.

In order to investigate the substantiality of the observed differences in gains, a one-way analysis of variance (Barlett) was carried out for the Cand. Negot. results as well as the English cohort results. The hypothesis of equal means was not proved to hold for the Cand. Negot. cohort ($p=0.0038$) at a significance level of 95%, but did hold for the English cohort ($p=0.3376$). Subsequently, a pairwise (Scheffe) comparison of the per cent differences between the treatment groups was carried out (see Table 4.23 for Cand. Negot. and Table 4.24 for English).

Table 4.23
Cand. Negot. cohort. Comparison of per cent difference (Scheffe). Gain (per cent). Total.

Row Mean- Col Mean	BASIS	NON-VISL
NON-VISL	10.9367 0.040	
VISL	14.6436 0.005	3.70689 0.649

Table 4.24
English cohort. Comparison of per cent difference (Scheffe). Gain (per cent). Total.

Row Mean- Col Mean	BASIS	NON-VISL
NON-VISL	5.05181 0.413	
VISL	4.18603 0.509	-.865777 0.976

The conclusion is that there is a statistical difference between Cand. Negot. BASIS and the two experimental groups NON-VISL and VISL. The gains in these two treatment groups are higher than the gains in the basis group. However, the difference in gains between the Cand. Negot. NON-VISL and the Cand. Negot. VISL groups is not large enough to be statistically significant. For the English cohort the ANOVA did not reveal a significant difference in gains, in other words the differences in gains were not substantial enough to hold statistically at the 95% significance level for the total test item figures. The figures in Table 4.24 do indicate, though, that similar to the Cand. Negot. cohort the largest difference is between the BASIS group and the VISL and NON-VISL groups whereas the statistical difference between the NON-VISL and the VISL groups is smaller.

4.6.2 Syllabus test levels

The overview of levels and gains given in the description of the total figures is an indication of the effects of the change which takes place in the students' knowledge from the initial stage of the experiment to the fulfilment of the ten-week treatment period, but it gives no indication of what type of knowledge the students possessed to begin with, nor does it in any way indicate which type of knowledge, if any, is affected by the treatment.

A break-down of the test items into a syllabus section and a non-syllabus section is needed to give a more informed idea of this process. The syllabus section consists of 16 test items and the non-syllabus section consists of 9 test items. The hypothesis is that the entrance level will be low in the syllabus section (for a detailed description of these two test sections and the differences between them see Chapter 4.5.4). It is further assumed that the entrance level will be equal for the two cohorts as far as the syllabus section is concerned. The students in two cohorts have various entrance exams (see sections 4.1 and 4.2), but their proficiency levels in general should be equal to each other in that all exams are graded in the same way.

4.6.2.1 Syllabus pre-test levels

4.6.2.1.1 The Cand. Negot. cohort. Syllabus pre-test

The three Cand. Negot. experimental groups are very similar with regard to mean values and standard deviations (see Table 4.25). The BASIS and the VISL groups have exactly the same mean score, 5.47 points, and the NON-VISL group's score is only slightly lower at 5.26 points. The Cand. Negot. cohort is a homogenous but low-scoring cohort.

Table 4.25
Cand. Negot. cohort. Pre-test levels. Point scores. Syllabus (max. 16).

Treatment	N	mean	sd	min	p25	p50
BASIS	15	5.47	2.40	2.20	3.00	5.65
NON-VISL	20	5.26	2.09	2.00	3.99	5.04
VISL	18	5.47	2.92	1.00	3.09	5.54
Total	53	5.39	2.44	1.00	3.82	5.12

Treatment	p75	max
BASIS	7.69	9.51
NON-VISL	6.06	10.90
VISL	6.50	12.60
Total	6.42	12.60

The minimum for the Cand. Negot. VISL group is 1 point, 2 points for the NON-VISL group and 2.2. for the BASIS group; in other words this cohort comprises subjects with barely any knowledge of the subject matter. The low level is supported by the fact that even for the 75% quartile the mean is only 6.42 points, and all groups are below 8 points, which is the chance level. The maximum score is highest for the VISL group, 12.60 points, followed by 10.90 points in the NON-VISL group, and the lowest maximum score, 9.51, is achieved in the BASIS group. Despite the high maximum score for the VISL group, it is clear that the VISL subjects are concentrated in the low end of the scale as indicated by the minimum and 75% figures.

The statistical differences between the groups were examined through two-tailed t-tests: VISL against NON-VISL ($t=0.2367$, $df=36$, $p=2.0395$), VISL against BASIS ($t=-$

0.0062, df=31, p=0.4997), NON-VISL against BASIS (t=-0.2605, df=33, p=0.7963). None of these tests shows statistical significance for a difference between the pre-test levels of the three Cand. Negot. groups, which leads to the conclusion that these groups have initial levels of knowledge which are comparable with other.

4.6.2.1.2 The English cohort. Syllabus pre-test

The English cohort's syllabus pre-test level is not quite as a homogenous as the Cand. Cand. Negot. cohort (see Table 4.26), but the means still show that the groups are close. The bottom level is low with a minimum score of 1.81 points. The pattern with the high-scoring basis group from the total figures re-emerges here in the mean scores, in that the basis group scores 7.78 points, the NON-VISL group 5.68 points, and the VISL group 6.68 points. The standard deviations are similar for the basis and the VISL groups but lower for the NON-VISL group, which is also the group with the smallest number of subjects.

Table 4.26
English cohort. Pre-test levels. Point scores. Syllabus (max. 16).

Treatment	N	mean	sd	min	p25	p50
BASIS	21	7.78	2.63	3.09	5.61	7.45
NON-VISL	15	5.68	1.88	1.81	4.38	5.58
VISL	18	6.68	2.62	2.53	4.33	6.75
Total	54	6.83	2.55	1.81	5.00	6.74

Treatment	p75	max
BASIS	9.32	12.30
NON-VISL	7.66	8.45
VISL	8.75	12.50
Total	8.64	12.50

The general level in the English NON-VISL group is low compared to the other two groups. Not only does the group have the lowest mean, it also has the lowest minimum score, the lowest maximum score, and the lowest scores in the upper quartile. Add to this the lowest standard deviation, 1.88, and it is clear that the observations in the NON-

VISL group are concentrated in the low end of the scale with little variance, an illustration of which is that the highest maximum score in the group is only 8.45 points so basically the subjects here score only at chance level or below. The VISL and the BASIS groups do much better and the maximum scores are similar, i.e. the VISL maximum score of 12.50 points is only minimally higher than the 12.30 points in the BASIS groups. The same similarity is true in reverse for the upper quartile. The difference between the BASIS group and the VISL group comes through in the higher mean of the BASIS group, and the fact that the lower quartile scores are higher than in the VISL group. The statistical differences between the groups were examined through two-tailed t-tests: VISL against NON-VISL ($t=1.3031$, $df=31$, $p=0.2024$), VISL against BASIS ($t=-1.2879$, $df=37$, $p=0.2059$), NON-VISL against BASIS ($t=-2.7886$, $df=34$, $p=0.0086$). None of these tests shows statistical significance for a difference between the pre-test levels of the English VISL and the NON-VISL groups, nor between the VISL and the BASIS groups. There is a statistically significant difference, however, between the English NON-VISL group and the BASIS group, which leads to the conclusion that the English NON-VISL and the BASIS have initial levels of knowledge which it is difficult to compare. With the score points in mind this result points to the interpretation that the English BASIS group has a pre-test level which is so high that its function as a control group is questionable, at least as far as relating it to the NON-VISL group results.

4.6.2.1.3 VISL, NON-VISL and BASIS results (syllabus pre-test) differences

Tables 4.27 and 4.28 are syllabus pre-test level percentage rates of the Cand. Negot. and English cohorts respectively. They demonstrate that the syllabus level of the Cand. Negot. cohort is considerably lower than the level of the English cohort, i.e. the mean Cand. Negot. syllabus success rate of 33.7% versus the mean English syllabus success rate of 42.7%.

The lowest English mean, 35.50% in the NON-VISL group, is higher than the highest in the Cand. Negot. groups, the 34.20% achieved by both the BASIS and the VISL groups. Both, however, are below chance level. The English VISL group and the Cand. Negot. VISL group have the highest maximum score of all, and they both

have high scores in the upper quartiles, but for both cohorts it is remarkable how close the VISL groups are to the NON-VISL groups. The NON-VISL groups are the ones with the lowest mean syllabus level in both cohorts. In the Cand. Negot. cohort the

Table 4.27
Cand. Negot. cohort. Pre-test levels (percentage success rates). Syllabus (%).

Treatment	N	mean	sd	min	p25	p50
BASIS	15	34.20	15.00	13.80	18.80	35.30
NON-VISL	20	32.90	13.00	12.50	24.90	31.50
VISL	18	34.20	18.30	6.25	19.30	34.60
Total	53	33.70	15.20	6.25	23.90	32.00

Treatment	p75	max
BASIS	48.10	59.40
NON-VISL	37.90	68.30
VISL	40.60	78.40
Total	40.10	78.40

Table 4.28
English cohort. Pre-test levels (percentage success rates). Syllabus (%).

Treatment	N	mean	sd	min	p25	p50
BASIS	21	48.60	16.4	19.30	35.10	46.60
NON-VISL	15	35.50	11.8	11.30	27.40	34.90
VISL	18	41.70	16.3	15.80	27.10	42.20
Total	54	42.70	15.9	11.30	31.30	42.10

Treatment	p75	max
BASIS	58.30	77.00
NON-VISL	47.90	52.80
VISL	54.70	78.40
Total	54.00	78.40

difference from the other groups is moderate unlike the case in the English cohort in which the difference from the other two groups is higher, and especially the difference

between the English NON-VISL and the English BASIS is notable. The assumption that the syllabus level would be low appears to be verified.

Despite the fact that the English cohort has a higher syllabus level in comparison to the Cand. Negot. cohort, it is in both cases clear that the knowledge in this field is sparse: not even in the 75% quartile do any of the Cand. Negot. groups succeed above chance level. The same situation is true of the English NON-VISL group, but for the VISL and BASIS groups the equivalent level is achieved in the 50% quartile which means that the Cand. Negot. cohort and the English NON-VISL perform at a comparatively similar level.

4.6.2.2 Syllabus post-test levels

4.6.2.2.1 The Cand. Negot. cohort. Syllabus post-test

The uniformity of results which characterised the pre-test levels is not evident in the post-test levels, on the contrary. The post-test results proffer a clear indication of progression from the BASIS group mean of 8.14 points, the NON-VISL group's 10.70 points to the VISL group's result of 11.70 points (see Table 4.29). The standard

Table 4.29
Cand. Negot. cohort. Post-test levels. Point scores. Syllabus (max. 16).

Treatment	N	mean	sd	min	p25	p50
BASIS	15	8.14	2.82	3.55	5.78	7.79
NON-VISL	20	10.70	2.40	4.38	9.24	11.00
VISL	18	11.70	2.02	7.75	10.00	11.60
Total	53	10.30	2.77	3.55	8.42	10.90

Treatment	p75	max
BASIS	10.90	12.00
NON-VISL	12.40	14.20
VISL	13.30	15.00
Total	12.10	15.00

deviations vary very little, with 2.02 points in the VISL group as the lowest, the NON-VISL marginally higher at 2.40, and 2.82 points for the BASIS group indicating that the variety of the mean results are genuine differences and not ascribable to a few deviating observations of a nature which might skew results.

The rising scale of results in the cohort, with the BASIS group at the bottom, the NON-VISL group in the middle, and the VISL group as the highest achievement is a general feature which is reflected in minimum-maximum scores as well as all the quartiles.

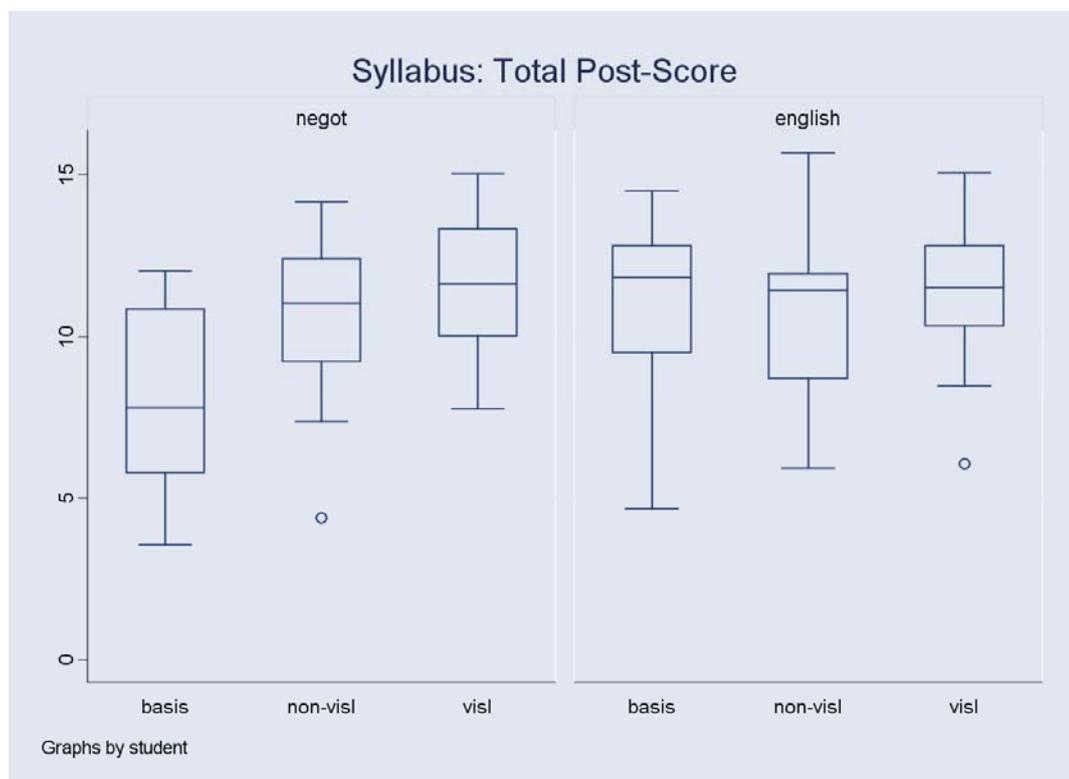


Figure 4.7 Boxplot. Syllabus post-test results by experimental groups.

It is interesting to see that the floor and the lower quartile of the Cand. Negot. VISL group is almost twice as high as that of the BASIS group. The NON-VISL group result comes close to that of the VISL group for the middle group of subjects but fails to do so for the bottom and the top observations. The boxplot in Figure 4.7 visualises the higher concentration of results for the Cand. Negot. NON-VISL and VISL groups compared to the BASIS groups.

4.6.2.2.2 The English cohort. Syllabus post-test

The mean score of the English cohort is 11 points, and the scores of all three groups are close to this mean (see Table 4.30). In the pre-test the BASIS group had the highest score mean, but now the VISL group manages to outperform the BASIS group with 11.30 points against 11.10 points. The scores are extremely close but the relation is interesting as it signifies a changed ranking from pre-test to post-test. However, it noteworthy that the BASIS group has managed to maintain such a high score. An investigation of the standard deviations support the equality of the three groups and so do the figures for the 50% quartile, the 75% quartile and the maximum score.

Table 4.30
English cohort . Post-test levels. Point scores. Syllabus (max. 16).

Treatment	N	mean	sd	min	p25	p50
BASIS	21	11.10	2.51	4.66	9.51	11.80
NON-VISL	15	10.50	2.36	5.92	8.72	11.40
VISL	18	11.30	2.16	6.05	10.30	11.50
Total	54	11.00	2.33	4.66	9.52	11.50

Treatment	p75	max
BASIS	12.80	14.50
NON-VISL	12.00	15.70
VISL	12.80	15.10
Total	12.40	15.70

It is at the bottom of the scale that a difference can be detected in as much as the minimum scores are hierarchical, ranging from 4.66 points in the basis group, over 5.92 in the NON-VISL group, to 6.05 points in the VISL group. In the span between the minimum and the 25% quartile the VISL group lead is still evident (10.3 points), and the VISL observations are clearly higher that the NON-VISL mark of 8.72 points. In contrast to this, the distance to the BASIS group, 9.51 points, is narrower, and for the 50% quartile the BASIS group figure is higher than the other two group scores, albeit only modestly so.

4.6.2.2.3 VISL, NON-VISL, and BASIS (syllabus post-test) differences

The tendency in the success rates of the post-test (see Tables 4.31 and 4.32) relating to the 16 test items comprising the syllabus section is that the Cand. Negot. observations (mean 64.40%) and English observations (mean 68.70%) do not deviate greatly when the means are weighed against each other. The variation is in the detail. The standard deviations tell us that 68% of Cand. Negot. observations have success rates between $(-17.30 + 64.40)$ and $(64.4+17.3)$, i.e. between 47.10% and 81.70%.

Table 4.31
Cand. Negot. cohort. Post-test levels. Success rate. Syllabus (%).

Treatment	N	mean	sd	min	p25	p50
BASIS	15	50.90	17.6	22.20	36.10	48.70
NON-VISL	20	66.70	15.0	27.40	57.70	69.00
VISL	18	73.20	12.6	48.40	62.80	72.70
Total	53	64.40	17.3	22.20	52.60	67.90

Treatment	p75	max
BASIS	67.90	75.30
NON-VISL	77.50	88.50
VISL	83.30	94.00
Total	75.40	94.00

The equivalent English post-test success rates are $(-14.60 + 68.70)$ and $(68.70+14.60)$, i.e. between 54.10% and 83.30%. The English cohort may not achieve significantly higher success rates, but the cohort proffers a lower spread, which is evident from the fact that the lower end observations are higher than those of the Cand. Negot. cohort, in fact the English cohort can be characterised as having a markedly higher floor than the Cand. Negot. cohort. In Table 4.27 we saw that the Cand. Negot. group was very homogenous and yet this homogeneity is not evident in the post-test results, which seem to show a differential effect for the respective treatments of the three Cand. Negot. experimental groups.

The same differential effect is hard to demonstrate in the English cohort though the VISL group, with a 70.60% mean success rate, has the advantage over the other groups

in overall terms. When looking at the minimum scores and the 25% quartile, the numbers indicate that the VISL group observations at this end of the scale lie at a similar level to the NON-VISL group, and when considering the 50% percentile, the BASIS group is the highest English group. The illustration provided by Figure 4.8 makes it clear that the

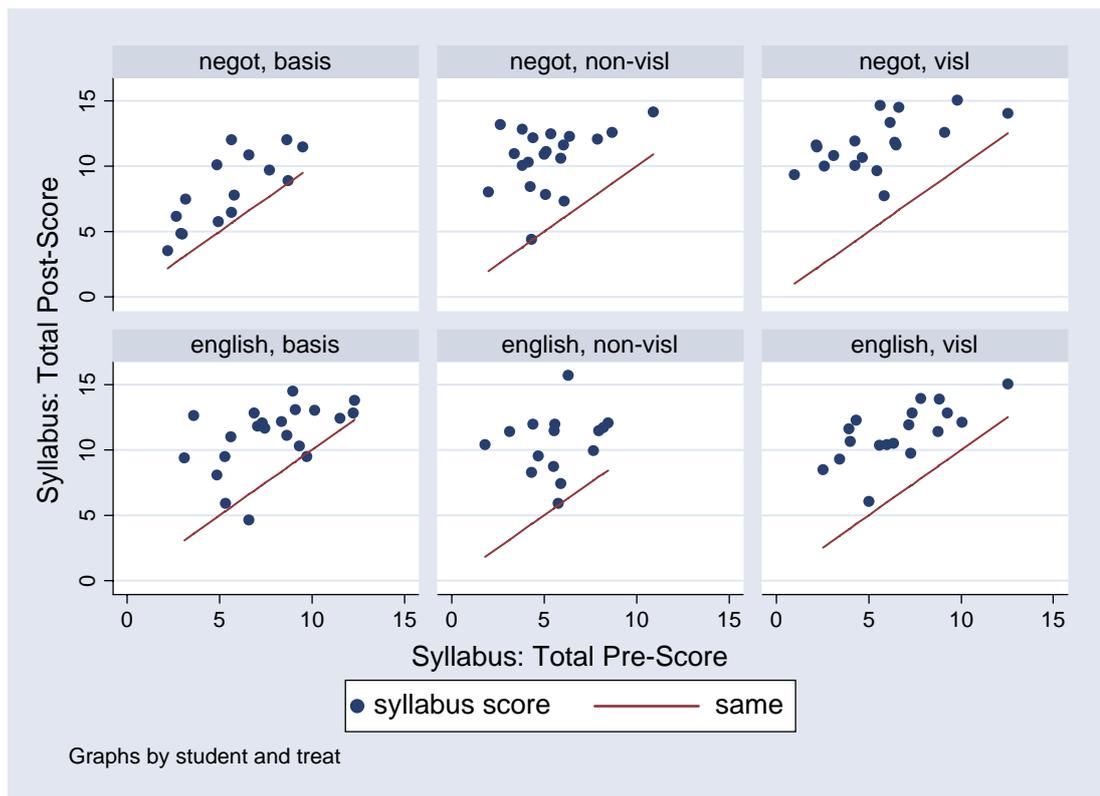


Figure 4.8 Scatterplot. Syllabus scores pre-test by post-test per experimental groups.

English syllabus post-test results have similar median levels for all three groups but the VISL group has lower spread and has observations concentrated at the upper end of the scale.

What is remarkable is that the English BASIS group, with a mean success rate of 69.1%, has managed to stay ahead of the NON-VISL group, mean 65.8%, especially in the light of the fact that the NON-VISL group contains the highest maximum success rate, which is 98% or almost complete success. However, this cannot outweigh the solid

results of the BASIS group in the middle field of observations (cf. the 25% and 50% quartiles).

Table 4.32
English cohort. Post-test levels. Success rates. Syllabus (%).

Treatment	N	mean	sd	min	p25	p50
BASIS	21	69.10	15.7	29.10	59.40	73.90
NON-VISL	15	65.80	14.7	37.00	54.50	71.40
VISL	18	70.60	13.5	37.80	64.60	71.90
Total	54	68.70	14.6	29.10	59.50	71.70

Treatment	p75	max
BASIS	80.00	90.70
NON-VISL	74.70	98.00
VISL	80.10	94.10
Total	77.50	98.00

The VISL treatment groups of both the Cand. Negot. and the English cohorts are highly successful and have a concentration of observations at the high percentiles. Especially for the Cand. Negot. VISL group this is clear as even the lowest observation is 48.40% and for the English VISL the 25% quartile is found at 64.60%. The NON-VISL groups are nearly as successful as the VISL groups but their spread is larger. The English BASIS mean group remains exceptionally high, but the spread is large. The Cand. Negot. BASIS is the lowest achievement group and in addition, it has a large spread in observations.

The evaluation of the various group results builds on the acceptance of the hypothesis of equal variances, i.e. a normal distribution of the observations. The homogeneity of variances was tested in a Kolmogorov-Smirnov test, and the p-values ranged from 0.571 for English BASIS group's post-test syllabus distributions to 0.956 for Cand. Negot. BASIS post-test syllabus observations. These results indicate that the results are normally distributed, and this means that the measured result differences of experimental groups can be compared with each other.

4.6.2.3 Syllabus gain levels

4.6.2.3.1 The Cand. Negot. cohort. Syllabus gains

It is quite clear that the Cand. Negot. VISL group gains most of all three groups (see Table 4.33). The VISL group gain of 39% is more than double that of the BASIS group, whose gain is a modest 16.7%.

Table 4.33
Cand. Negot. cohort. Gain levels. Syllabus (%).

Treatment	N	mean	sd	min	p25	p50
BASIS	15	16.70	11.1	1.31	8.44	12.40
NON-VISL	20	33.80	15.6	.313	25.20	36.90
VISL	18	39.00	14.9	9.31	31.90	41.20
Total	53	30.70	16.7	.313	17.10	32.80

Treatment	p75	max
BASIS	26.60	39.90
NON-VISL	41.70	66.20
VISL	49.00	59.10
Total	44.50	66.20

The NON-VISL group achieves a gain of 33.80% which is also double that of the BASIS group which has the lowest spread of all three groups, sd 11.1, which should indicate that the result of the group is fairly stable and concentrated at the low end of achievements. The NON-VISL group has the lowest minimum gain (.313) which is one reason why the spread in this group, sd 15.6, is higher than the spread in any of the other groups, but the spread in the group is compounded by the fact that it also has the highest maximum gain score at 66.20%.

Notwithstanding these results the observations are evenly distributed over the quartiles. The VISL gains have a lower spread, sd 14.9, than the NON-VISL group but higher than the BASIS group. The VISL observations are characterised by having a high minimum gain, and the gain in all quartiles is higher than in those of the other two

groups, which is an indication that the high mean gain is an expression of a genuine change in this group.

4.6.2.3.2 The English cohort. Syllabus gains

Expectedly, the BASIS group has the lowest gain (see Table 4.34) with 20.50% whereas the NON-VISL group with 30.30% and the VISL group with 28.90% clearly do much better. Unexpected is the negative minimum gain (loss) of -12.10% which is due to two negative observations. Generally, however, the observations at this end of the scale are low, e.g. the 25% quartile figure is only 6.12%. The figure for the 75% quartile is also low at 29.90%. Observations as expressed by the 50% quartile indicate that the BASIS group and the NON-VISL gains approximate each other while the VISL gain here is higher, just as the minimum gain for the VISL group is extremely high at 6.56% in contrast to 1% for the NON-VISL group, and ,as mentioned, the negative gain in the BASIS group.

Table 4.34
English cohort. Gain levels. Syllabus (%)

Treatment	N	mean	sd	min	p25	p50
BASIS	21	20.50	16.3	-12.10	6.12	24.00
NON-VISL	15	30.30	17.2	1.00	20.00	24.60
VISL	18	28.90	12.2	6.56	16.50	29.70
Total	54	26.00	15.7	-12.10	15.40	26.20

Treatment	p75	max
BASIS	29.90	56.60
NON-VISL	47.40	58.70
VISL	37.30	49.60
Total	37.00	58.70

The mean gains of the NON-VISL and the VISL groups are not significantly different although the NON-VISL gain is slightly higher than the VISL gain. Essentially, it seems that the VISL achievements lie in raising the floor and eliminating the low gains that can be seen in the other two groups. However, the VISL subjects do not reach the height achieved by NON-VISL subjects. The 75% quartile figure for the VISL group is 37.3%,

but the NON-VISL figure is 47.40%. There is an equal difference in the maximum gains, i.e. the VISL group maximum is 49.60% against the NON-VISL maximum of 58.70%. In essence, the NON-VISL group has a large spread, sd 17.2, with observations clustering at the bottom and the top end of the scale. In contrast, the VISL spread is limited, sd 12.2, and observations cluster round the middle with a high minimum.

4.6.2.3.3 VISL, NON-VISL, and BASIS (syllabus gain) differences

There seems to be a very distinct differential effect in the Cand. Negot. cohort which is not present in the English cohort to the same extent. The Cand. Negot. cohort shows a clear hierarchy, i.e. the basis group has the lowest gain, the NON-VISL group comes next, and the gains in the VISL group are highest of all.

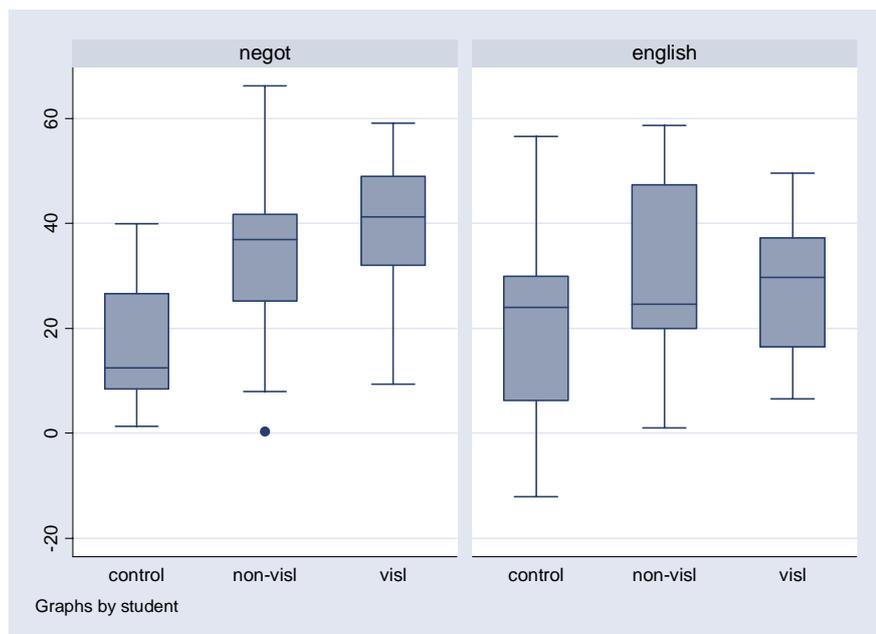


Figure 4.9⁷⁴ Boxplot. Syllabus gains percentage points by treatment groups.

The mean Cand. Negot. figures for all the measuring points are higher than for the VISL group, but it must be kept in mind that the English pre-test levels were higher than the Cand. Negot. equivalents; potentially there was less room for improvement or gains. For the VISL treatment there appears to be a common feature in the effect on the two cohorts in that the lowest achieving subjects are lifted more by the VISL treatment than

⁷⁴ The BASIS group is here listed as 'control'.

is the case in the other two experimental groups. This also becomes evident from the visualisation of gain level in Figure 4.9. The VISL median (50% quartile) is higher than in any of the other groups. The NON-VISL treatment gives a large spread, but it appears to have a good effect on the middle and high achievers. The BASIS groups do not serve the lowest segment of subjects well. The English BASIS group subjects even experience losses. In the BASIS groups of both the Cand. Negot. and the English cohorts the 25% quartile gains are low (8.44% and 6.12%, respectively) compared to the other experimental groups which see gains ranging from 16.50% (for English VISL) to 31.9% (for Cand. Negot. VISL).

In order to investigate the substantiality of the observed differences in gains a one-way analysis of variance⁷⁵ was carried out for the Cand. Negot. results as well as the English cohort results. The hypothesis of equal means was not proved to hold for the Cand. Negot. cohort ($p=0.0001$; $F(2, 50) = 10.75$) at a significance level of 95%, but was rejected for the English cohort ($p=0.1142$; $F(2, 51) = 2.26$). Subsequently, a pairwise (Scheffe) comparison of the per cent differences in means between the treatment groups was carried out (see Table 4.35 for Cand. Negot. and Table 4.36 for English).

Table 4.35
Cand. Negot. cohort. Comparison of per cent difference (Scheffe). Gain (%).
Syllabus.

Row Mean- Col Mean	BASIS	NON-VISL
NON-VISL	17.09 0.004	
VISL	22.33 0.000	5.24 0.533

The test showed that there is a significant difference between the Cand. Negot. VISL and BASIS groups ($p=0.0001$), and between the Cand. Negot. NON-VISL and BASIS groups ($p=0.004$). The difference between the VISL and NON-VISL groups is not statistically significant at an alpha (α) level of 0.05. The overall result for the English

⁷⁵ Barlett's test for homogeneity of variances was applied.

cohort showed no difference between the groups that could be said to be statistically significant at the 95% confidence level.

Table 4.36
English cohort. Comparison of per cent difference (Scheffe). Gain (%).
Syllabus.

Row Mean- Col Mean	BASIS	NON-VISL
NON-VISL	9.80 0.177	
VISL	8.38 0.244	-1.42 0.965

The English BASIS group has from the start mustered atypically high levels, which makes it less useful as a comparison and a control group. However, the statistical test could not establish any difference in overall gains between the English VISL and the NON-VISL groups that would be statistically significant.

4.6.3 Non-syllabus test levels

4.6.3.1 Non-syllabus pre-test levels

This section comprises 9 test items (described in Chapter 4.5), all of which can be said to fall into the category of grammaticality judgements, i.e. no concrete explicit knowledge was asked for in so far as respondents could react to the items intuitively. All items are in the multiple choice format and for each item five choices are possible, but in essence it is only a choice of choosing one sentence over another. The choices are that the subjects can choose sentence A, sentence B, both sentence A and sentence B, neither sentence A, not sentence B, and finally there is a 'Don't know' option.

The issue here is formulated in research question 5: *Can syllabus instruction affect the non-syllabus results?*

The assumption was that the level of correctness would be high and certainly beyond the chance level. The level of performance was expected to be equal in all of the groups

since these items call upon procedural knowledge rather than declarative knowledge. It was hypothesised that these items would be immune to the various treatments of the experiment (see Chapter 1).

4.6.3.1.1 The Cand. Negot. cohort. Non-syllabus pre-test

Table 4.37 shows that the subjects in the various groups constitute a very homogenous cohort. The mean values range from 5.20 points (NON-VISL) to 5.33 (BASIS) to 5.67 (VISL). The standard deviations are also almost identical. This trend is repeated in all the quartiles and the maximum scores; all groups have 8 points as the maximum. All groups score at or beyond chance level in the 25% quartile which is remarkable.

Table 4.37
Cand. Negot. cohort. Pre-test levels. Point scores. Non-syllabus (max. 9)

Treatment	N	mean	sd	min	p25	p50
BASIS	15	5.33	1.35	3	5.00	5
NON-VISL	20	5.20	1.64	2	4.50	5
VISL	18	5.67	1.46	3	5.00	6
Total	53	5.40	1.49	2	5.00	5

Treatment	p75	max
BASIS	7.00	8
NON-VISL	6.50	8
VISL	7.00	8
Total	7.00	8

The weakest result is found in the NON-VISL groups; there is no substantial difference between the three groups as far as their pre-test- levels are concerned, which indicates that subsequent post-test results should be reliably comparable. The range of 68% of observations as indicated by the standard deviations is for the VISL group $(-1.46+5.33)$ and $(5.33+1.46)$, i.e. between 3.87 and 6.79 points. The NON-VISL range is $(-1.64+5.20)$ and $(5.20+1.64)$, i.e. between 3.56 and 6.84 points. The BASIS range is $(-1.35+5.33)$ and $(5.33+1.35)$, i.e. between 3.98 and 6.68 points. In order to ascertain whether the perceived similarities and differences were statistically significant two-

tailed t-tests were carried out: VISL against NON-VISL ($t=0.9289$, $df=36$, $p=0.3591$), VISL against BASIS ($t=0.6828$, $df=31$, $p=0.4997$), NON-VISL against BASIS ($t=-0.2638$, $df=33$, $p=0.7935$). The tests confirmed the uniform pre-test level of proficiency of all the Cand. Negot. groups.

4.6.3.1.2 The English cohort. Non-syllabus pre-test

Homogeneity is a characteristic feature also in the English cohort (see Table 4.38). The BASIS group scores the highest mean of 6.38 points, the NON-VISL group has the lowest score of 6.27, which is almost identical to the VISL group's score of 6.28. The standard deviation is highest in the NON-VISL group, $sd\ 1.49$, and lowest in the VISL group, $sd\ 1.07$, both of which are close to the BASIS sd of 1.32. The median of 6 points (= the 50% percentile) is almost the same as the mean value of 6.31 points, which is a sign that the cohort observations are normally distributed.

Table 4.38
English cohort. Pre-test levels. Point scores Non-syllabus (max. 9).

Treatment	N	mean	sd	min	p25	p50
BASIS	21	6.38	1.32	3	6	6
NON-VISL	15	6.27	1.49	4	5	6
VISL	18	6.28	1.07	5	5	6
Total	54	6.31	1.27	3	5	6

Treatment	p75	max
BASIS	7	9
NON-VISL	8	9
VISL	7	8
Total	7	9

It means that 68% of scores in the English VISL group are contained in the band ($-1.07 + 6.28$) and $(6.28+1.07)$, i.e. 68% of all scores fall between 5.86 and 7.35 points; for NON-VISL the band is 4.20-7.76 points, for the BASIS group the band is 5.06-7.700 points. This illustrates that the spread is higher in the NON-VISL group than in the other groups, especially higher than the spread of the VISL group whose scores lack the

low end which is evident in the NON-VISL group. Despite these variations the quartile figures tell us unequivocally that the groups are very homogeneous and that the level is very high. The maximum score for the BASIS and the VISL groups have reached the absolute maximum of 9 points (one observation in either group), and the NON-VISL group maximum score is 8 points.

In order to ascertain whether the perceived similarities and differences were statistically Significant, two-tailed t-tests were carried out: VISL against NON-VISL ($t=0.0241$, $df=31$, $p=0.9809$), VISL against BASIS ($t=-0.2688$, $df=37$, $p=0.7895$), NON-VISL against BASIS ($t=-0.2380$, $df=34$, $p=0.8135$). The tests confirmed the uniform statistical level of proficiency of all the English groups.

4.6.3.1.3 VISL, NON-VISL, and BASIS (non-syllabus pre-test) differences

The Cand. Negot. cohort mean values (see Table 4.39) exceed the chance level in that it has a 60% average success rate, and so do the mean values of the English cohort (see Table 4.40), whose mean success rate is 70.2%. Even the 25% quartile figures of either cohort exceed the chance level.

Table 4.39
Cand. Negot. cohort. Pre-test levels percentage. Non-syllabus (%).

Treatment	N	mean	sd	min	p25	p50
BASIS	15	59.30	14.9	33.30	55.60	55.60
NON-VISL	20	57.80	18.2	22.20	50.00	55.60
VISL	18	63.00	16.2	33.30	55.60	66.70
Total	53	60.00	16.5	22.20	55.60	55.60

Treatment	p75	max
BASIS	77.80	88.90
NON-VISL	72.20	88.90
VISL	77.80	88.90
Total	77.80	88.90

Table 4.40
English cohort. Pre-test levels percentage. Non-syllabus (%).

Treatment	N	mean	sd	min	p25	p50
BASIS	21	70.90	14.7	33.30	66.70	66.70
NON-VISL	15	69.60	16.5	44.40	55.60	66.70
VISL	18	69.80	11.9	55.60	55.60	66.70
Total	54	70.20	14.1	33.30	55.60	66.70

Treatment	p75	max
BASIS	77.80	100.00
NON-VISL	88.90	100.00
VISL	77.80	88.90
Total	77.80	100.00

Generally speaking, the English cohort seems to have a slightly higher level of performance than the Cand. Negot. cohort. The English cohort alone comprises results which are one hundred per cent correct. However, the overall conclusion is that the VISL, the NON-VISL and the BASIS groups are very homogenous with regard to inter-cohort as well intra-cohort levels.

4.6.3.2 Non-syllabus post-test levels

4.6.3.2.1 The Cand. Negot. cohort. Non-syllabus post-test

The post-test results (see Table 4.41) show a level field between the three groups. The highest score is found in the VISL group, with a mean of 6.56 points. VISL is the group with the largest spread, sd 2.01 (see also Figure 4.12 for individual observations), mainly due to the low minimum score of three (two observations) at one end combined with the maximum score of 9 at the high end of the scale (one observation). Especially the NON-VISL group is very homogenous, but the differences between the groups are only minor (see Figure 4.10) and this is also supported by the quartile figures.

The post-test Cand. Negot. observations are concentrated at the upper end of achievements. The chance level is surpassed at the 25% quartile level, in which all groups show a mean score of five points.

Table 4.41
Cand. Negot. cohort. Post-test levels. Point scores. Non-syllabus (max. 9).

Treatment	N	mean	sd	min	p25	p50
BASIS	15	6.13	1.51	4	5	6
NON-VISL	20	6.00	1.45	4	5	6
VISL	18	6.56	2.01	3	5	8
Total	53	6.23	1.66	3	5	6

Treatment	p75	max
BASIS	7	9
NON-VISL	7	8
VISL	8	9
Total	8	9

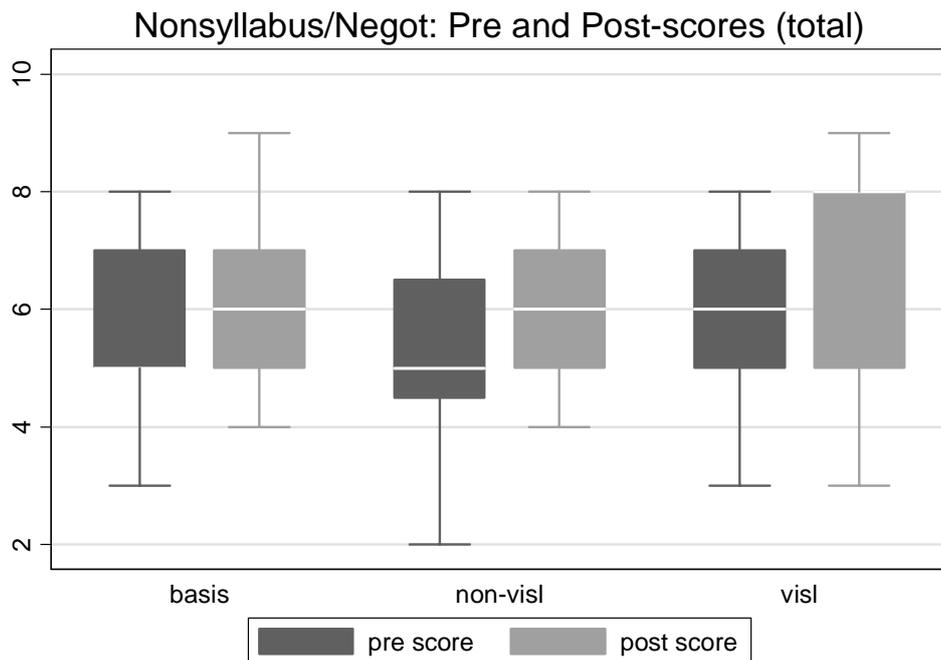


Figure 4.10 Boxplot. Cand. Negot. cohort. Non-syllabus. Pre-test by post-test by experimental groups.

Like the VISL group, the BASIS group encompasses one observation of nine points, the absolute maximum possible, whereas eight is the highest score in the NON-VISL group.

The ranking between the groups is the same as in the pre-test. VISL has the highest mean and the highest median level. This is visualised in Figure 4.10.

4.6.3.2.2 The English cohort. Non-syllabus post-test

The English post-test results (see Table 4.42) show the same ranking as the pre-test results did, i.e. the BASIS group has the highest score, 6.95 points, followed by the VISL group, 6.53 points, and lowest is the NON-VISL score of 6.53 points.

Table 4.42
English cohort. Post-test levels. Point scores. Non-syllabus (max. 9).

Treatment	N	mean	sd	min	p25	p50
BASIS	21	6.95	1.16	4	6	7.00
NON-VISL	15	6.53	1.25	4	6	7.00
VISL	18	6.56	1.29	4	6	6.50
Total	54	6.70	1.22	4	6	7.00

Treatment	p75	max
BASIS	8	9
NON-VISL	7	8
VISL	7	9
Total	8	9

As these figures show, the between-group variation is little and so is the within-group variation, i.e. the spread as expressed by the standard deviation, which is very similar for the various groups. The quartile figures support this, as do the minimum and maximum figures, which are almost identical for all the English experimental groups. The chance level is surpassed already in the 25% quartile for all groups, which score the same mean (6 points).

The visualisation of the English cohort results in Figure 4.11 illustrates the concentration of results at the upper end of the scale.

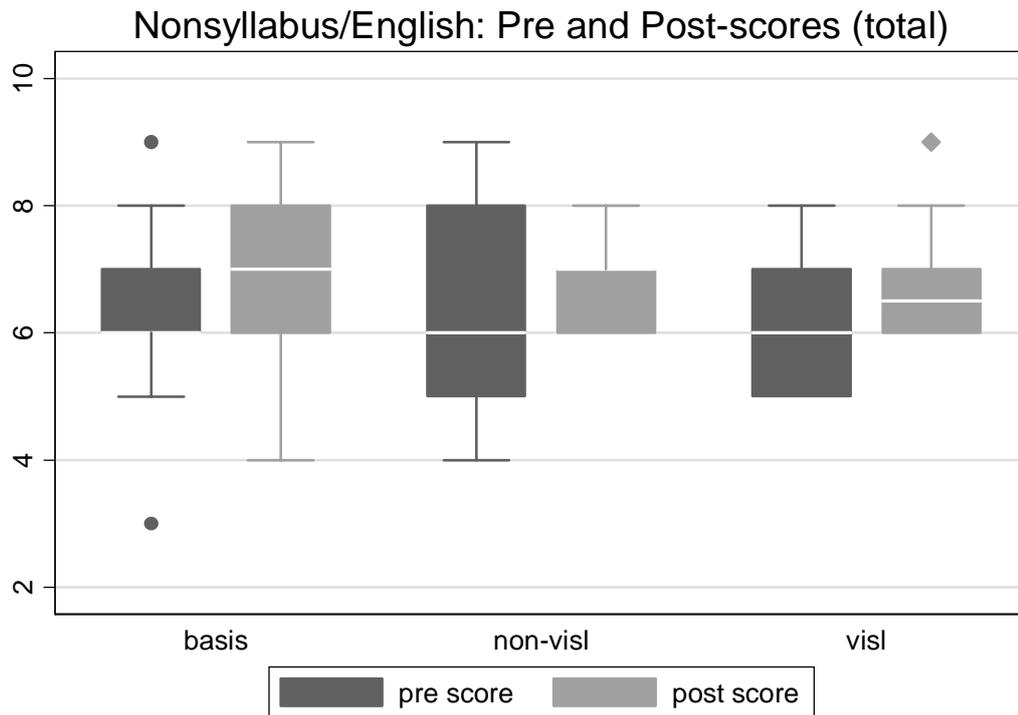


Figure 4.11 Boxplot. English cohort. Non-syllabus. Pre-test by post-test by experimental groups.

4.6.3.2.3 VISL, NON-VISL, and BASIS (non-syllabus post-test) differences

The inter-cohort as well as the intra-cohort non-syllabus post-test results (see Tables 4.43 and 4.44) appear to be very close. It must be taken into consideration that, given the low maximum of 9 point, one point's score difference can translate into substantial percentage differences. It is the closeness of results rather than differences which attracts attention, especially for the lowest 25 % of observations. Things change slightly for the middle and the upper level of observations.

The percentage figures reveal that the VISL groups of the two cohorts have exactly the same success rates, but the percentage figures also reveal that there are differences between the groups which were hardly noticeable in the point score due to the low point sums in items, e.g. all six groups had mean scores of six points with the variation only evident in the decimals.

Table 4.43
Cand. Negot. cohort. Post-test levels. Success rates. Non-syllabus (%).

Treatment	N	mean	sd	min	p25	p50
BASIS	15	68.10	16.7	44.40	55.60	66.70
NON-VISL	20	66.70	16.1	44.40	55.60	66.70
VISL	18	72.80	22.3	33.30	55.60	88.90
Total	53	69.20	18.4	33.30	55.60	66.70

Treatment	p75	max
BASIS	77.80	100
NON-VISL	77.80	88.90
VISL	88.90	100
Total	88.90	100

The mean percentage figures make clear that the English BASIS group (77.20%) outperforms the other English groups. The Cand. Negot. VISL group (72.80%) outperforms the other Cand. Negot. groups and achieves the same result as the English VISL group mean, in other words the two VISL groups are higher than those of all the other groups, but differences are very small. For both cohorts it is true that the NON-VISL groups are the lowest performing groups.

Table 4.44
English cohort. Post-test levels. Success rate. Non-syllabus (%).

Treatment	N	mean	sd	min	p25	p50
BASIS	21	77.20	12.9	44.40	66.70	77.80
NON-VISL	15	72.60	13.8	44.40	66.70	77.80
VISL	18	72.80	14.4	44.40	66.70	72.20
Total	54	74.50	13.6	44.40	66.70	77.80

Treatment	p75	max
BASIS	88.90	100
NON-VISL	77.80	88.90
VISL	77.80	100
Total	88.90	100

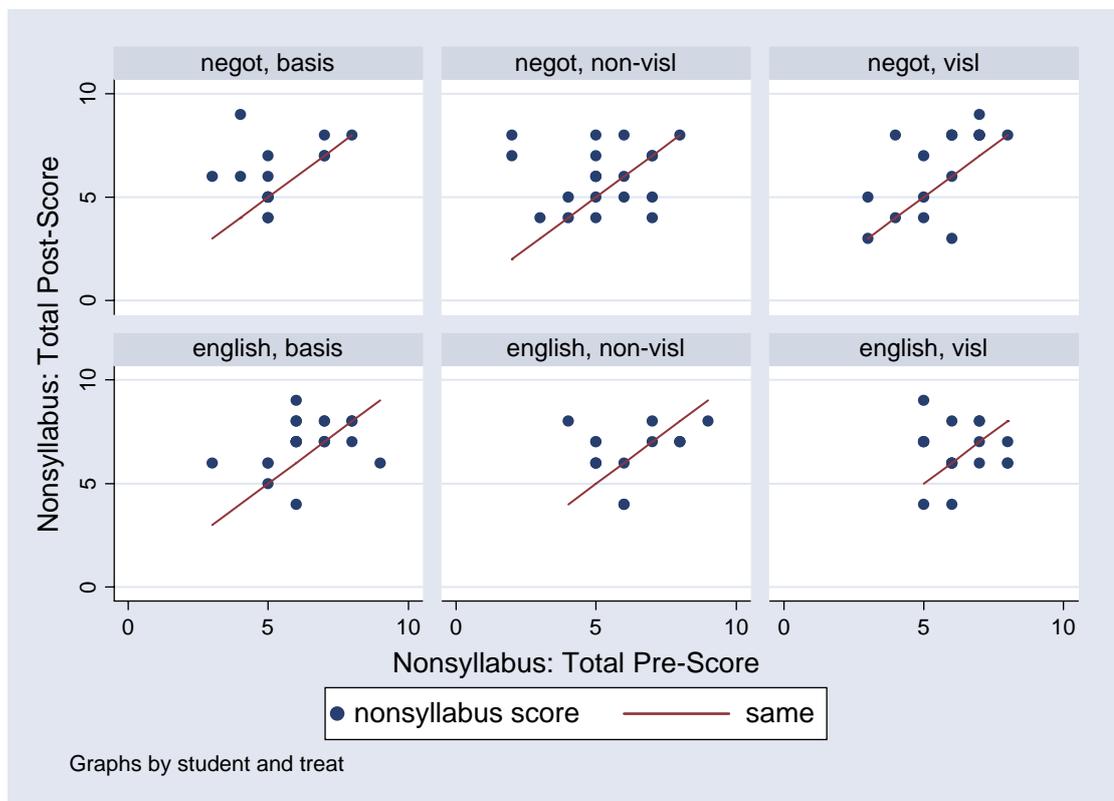


Figure 4.12 Scatterplot. Non-syllabus scores pre-test by post-test per experimental groups.

When looking at the top quartiles it appears that the Cand. Negot. VISL group and the English BASIS levels are similar and higher than the others. The conclusion is, however, that uniformity is more prominent than differences. One feature which must be noted is the very high general success rate in that chance level is reached by the 25% quartile.

4.6.3.3 Non-syllabus gain levels

At this point it is important to note that the items in question measure general proficiency and not the metalinguistic knowledge targeted by the syllabus instruction of the experimental classes.

Table 4.45
All experimental groups overview. Non-syllabus mean gains (points) and standard deviations.

Student	BASIS	NON-VISL	VISL	Total
Cand.Negot	0.80	0.80	0.89	0.83
	1.61	2.12	1.53	1.76
English	0.57	0.27	0.28	0.39
	1.43	1.67	1.67	1.56
Total	0.67	0.57	0.58	0.61
	1.49	1.93	1.61	1.67

4.6.3.3.1 The Cand. Negot. cohort. Non-syllabus gains

The overview of point gains in Table 4.45 makes it clear that the non-syllabus gains for all experimental groups in both cohorts are unimpressive. We are dealing with small numbers in absolute terms, and what may appear to be significant gains percentage-wise (see Table 4.46) amounts to less than one score point.

The most striking feature apparent from Table 4.46 is the negative minimum gain (loss) which happens in all groups, less in the BASIS group than in the NON-VISL and the VISL groups. Figures 4.10 and 4.12 visualise the distribution of observations and the relation between pre- and post-test results. The BASIS group is the only group which has more observations of gain than of unchanged or negative results.

As mentioned above, only the Cand. Negot. VISL group has an interesting gain figure, and this seems to be due not so much to the number of gain observations but rather to the fact the gains which do occur are high and that a large number of observations are unchanged with the number of actual negative gains limited to two observations.

The NON-VISL group has a higher number of negative observations but also a high number of unchanged observations and some high positive observations. This group has the highest maximum gain of all (66.70%).

Table 4.46
Cand. Negot. cohort. Gain levels. Non-syllabus (%).

Treatment	N	mean	sd	min	p25	p50
BASIS	15	8.89	17.9	-11.10	0.00	0.00
NON-VISL	20	8.89	23.5	-33.30	0.00	5.56
VISL	18	9.88	17	-33.30	0.00	11.10
Total	53	9.22	19.6	-33.30	0.00	11.10

Treatment	p75	max
BASIS	22.20	55.60
NON-VISL	16.70	66.70
VISL	22.20	44.40
Total	22.20	66.70

A further unusual aspect is that the median in the BASIS group, despite an overweight of positive gain observations, is zero per cent, which combined with a mean value of 8.89% amounts to an unusual distribution. This is evident from Figure 4.10, which illustrates that the mean value of scores is above the median; indeed this group, like the NON-VISL group, has a high maximum gain (55.60%) combined with negative gains.

4.6.3.3.2 The English Cohort. Non-syllabus gains

The gains in the English cohort are modest (see Table 4.47). The unusual aspect is the negative gains (loss) registered in all groups, and in the VISL and the NON-VISL groups to such an extent that the median value gains in both these groups is zero.

Figure 4.12 illustrates the distribution of positive and negative movements from pre-test to post-test, and only the English BASIS group has more positive gain observations than unchanged and negative gains. The English VISL group on the other hand has a clear majority of negative and unchanged gain observations. The VISL group as well as the NON-VISL group 50% quartile figure is zero %. The maximum gains in all English groups are low. For the BASIS group, the case is that the negative minimum gain equals the positive maximum gain (-/+ 33.3%).

Table 4.47
English cohort. Gain levels. Non-syllabus (%).

Treatment	N	mean	sd	min	p25	p50
BASIS	21	6.35	15.9	-33.30	0.00	11.10
NON-VISL	15	2.96	18.5	-22.20	-11.10	0.00
VISL	18	3.09	18.6	-22.20	-11.10	0.00
Total	54	4.32	17.3	-33.30	-11.10	0.00

Treatment	p75	max
BASIS	11.10	33.30
NON-VISL	11.10	44.40
VISL	22.20	44.40
Total	11.10	44.40

4.6.3.3.3 VISL, NON-VISL, and BASIS (non-syllabus gain) differences

The two experimental groups VISL and NON-VISL as well as the BASIS group have hardly any demonstrable progress from pre-test to post-test. Further, the results are characterised by negative (i.e. loss) observations. The figures appear to invite the conclusion that the groups are unaffected positively by the instruction. It must be noted, however, that there are large individual differences. Some subjects have high gains whereas a number of individuals have high negative gains (losses), which in itself is remarkable. This happens in both cohorts. Both cohorts are characterised by an overrepresentation of negative or unchanged observations in the VISL and NON-VISL groups whereas the BASIS groups seem to be more resistant to loss.

The pre-test results were quite high compared to the syllabus items. This can give rise to at least two comments or interpretations. First, that the high success rate in the pre-test would make room for only limited improvement and second, that the high pre-test levels might have been the result of guesswork which the individuals concerned have not been able to keep up in the post-test, with a resulting negative gain score as the

consequence. It is in the nature of judgement tests that responses do not so much build on specific knowledge but rather on intuitive responses. It is possible that the nine items included do indeed belong to a different category of knowledge than the items comprised by the syllabus section. This type of knowledge would then be processed and stored differently than the knowledge contained in the syllabus instruction, and naturally the non-syllabus proficiency would be left unaffected were that the case (see Chapter 5 for an overview of processes of learning).

Table 4.48 Cand. Negot. cohort. Comparison of per cent difference (Scheffe). Gains. Non-syllabus.

Row Mean- Col Mean	control	NON-VISL
NON-VISL	0.00 1.000	
VISL	0.99 0.990	0.99 0.988

Table 4.49 English cohort. Comparison of per cent difference (Scheffe). Gains. Non-syllabus.

Row Mean- Col Mean	control	NON-VISL
NON-VISL	-3.39 0.851	
VISL	-3.26 0.847	0.12 1.000

It is evident, no matter which interpretation one chooses, that the effect of the treatment is different with regard to the non-syllabus items than is the case for the syllabus items. In order to examine whether the observed differences were of significant size, a one-way analysis of variance⁷⁶ was carried out for both the Cand. Negot. cohort and the English cohort. The hypothesis of equal means was rejected for the Cand. Negot. cohort ($p=0.9856$; $F(2, 50) = 0.01$) and the English cohort ($p=0.7962$; $F(2, 51) = 0.23$). The

⁷⁶ Barlett's test for homogeneity of variances was applied.

relations between the experimental groups were examined in pairwise (Scheffe) comparisons and the results for the differences and p-values are given in Table 4.48 for the Cand. Negot. cohort and in Table 4.49 for the English cohort.

The p-values for all group comparisons are very high, that is, close to or at 1, which is an indication that the means values are almost identical. This means that the differences observed between the groups were not strong enough to stand up to statistical testing.

4.6.4 Success rate (post-test) differences in syllabus and non-syllabus results

The differences and similarities in success rates will be described separate from gain results, which follows in Chapter 4.6.5. The reason why success rates as well as gain rates are given separate attention is that the gain results do not necessarily reveal the information which would be most relevant when judging the instruction and its effects in relation to the external context of education planning. The success rates illustrate aspects of relevance to the subject matter in general whereas the gain rates illustrate the efficacy of the various methods measured against each other. From the point of view of the external setting, for instance with regard to the study programmes involved, it would be of interest to know the effect the instruction has in the particular educational framework, which has an interest in information pertaining to the level of knowledge attainable from a particular instructional input.

4.6.4.1 VISL, NON-VISL, and BASIS syllabus success rate (post-test) differences

In experimental designs it is a point of discussion as to which level of attainment can be considered meaningful as a breaking point for distinguishing successful students from less successful students. By extrapolation the question is which level of attainment should be set as a mark for the distinction between what is considered satisfactory and what is less so. For this experiment and with this method of measuring it would be natural to say that post-test success rates should exceed a correctness level of 50% for the syllabus section items.

The set success criterion is only just achieved for the Cand. Negot. BASIS group (see Table 4.31), which only achieves a mean success rate 50.86% in the syllabus section,

and even the 25% quartile level is only 48.70%; this means that half of all Cand. Negot. BASIS subjects do not meet the set criterion for success, which makes it the lowest performing of all six groups in the study. For all the other groups results are beyond chance level. The English BASIS had a high score already at the entrance level when this group was the highest scoring group of all six (see Table 4.14), and the score was so high that it was significantly different from the English NON-VISL group ($t = -2.27886$, $df = 34$, $p = 0.0086$), and therefore to some extent not comparable to these groups as far as the post-test success rates are concerned.

The two NON-VISL groups are very similar with their intermediary level syllabus results, English 65.78%, Cand. Negot. 66.69%. The syllabus NON-VISL groups are not affected by their treatment to the same extent as the VISL groups, but like the results from the VISL groups, it is the Cand. Negot. group which shows the higher success rate.

The clearly best results are achieved by the VISL groups; especially the Cand. Negot. group result of 73.18% is outstanding and the highest of all six groups. The Cand. Negot. VISL group is also the group with the lowest spread, which indicates that the result is reliable and solid. The English VISL group result of 70.59% is only slightly higher than the English BASIS group's 69.09%. The boxplot of the post-test results (see Figures 4.8) illustrates how the floor (i.e. the lower percentile) in the VISL groups was lifted, creating a large group of equal results in the middle to high area. To some degree this also happens in the NON-VISL groups but not in the BASIS groups. This gives an effect which results in the VISL groups being more unified (leaving few students behind). The VISL treatment appears to be a treatment which works for low-achievers and high achievers alike, but compared to the other treatments it seems especially effectual for the weaker students.

The overall means of the two cohorts are not very far apart – the English syllabus post-test mean is higher at 68.67% against the Cand. Negot. syllabus post-test mean of 64.41%. Neither are impressive results in the light of ten weeks of treatment in a subject matter they were all supposed to have some knowledge of before entering university. Apparently, the subject matter is a difficult discipline to master. Especially the result of

the Cand. Negot. BASIS group is discouraging because this group represents the case of the regular scenario of the study programme for these students and thus this result gives occasion for concern. It must also be noted that even in the Cand. Negot. VISL group, which is the highest performing syllabus group, the mean post-test success rate is only 73.18%, and consequently a quarter of the syllabus content has not been learned on average. The decision to set the 50% level as the criterion for success is very modest. It therefore needs to be perspectivised with the fact that very few students master the syllabus content completely⁷⁷ after ten weeks of instruction.

The syllabus section results represent a measure of the metalinguistic knowledge learned by the subjects in the experiment and thus a measure of the effect of the treatments. There is a differential effect of the treatments, with VISL groups showing an advantage over the others, a difference which for the Cand. Negot. VISL groups is especially high.

Table 4.50
Overview mean values syllabus (in %) ⁷⁸

Student	Treatment			Total
	BASIS	NON-VISL	VISL	
Cand. Negot	34.16	32.90	34.16	33.69
	50.86	66.69	73.18	64.41
	16.70	33.78	39.02	30.73
English	48.60	35.49	41.73	42.67
	69.09	65.78	70.59	68.67
	20.49	30.28	28.86	26.00
Total	42.59	34.01	37.95	38.22
	61.49	66.30	71.89	66.56
	18.91	32.28	33.94	28.34

1. number: pre, 2. number: post, 3. number: post - pre

The analysis of the quartile percentages pointed to an interpretation of the VISL treatment results that, as far as the syllabus content is concerned, the weaker students

⁷⁷ The highest individual scores in the syllabus section are 94% in the Cand. Negot. VISL group (one observation) and 94.10% in the English VISL group (one observation)

⁷⁸ The figures in this table give the exact two decimal figures whereas the original tables in previous chapters, including the percentiles, have been rounded up or down in the second decimal.

appear to be supported better by the VISL treatment than the treatment in the other groups.

4.6.4.2 VISL, NON-VISL, and BASIS non-syllabus success rate (post-test) differences

The non-syllabus difference between pre-test and post-test levels was investigated to establish where the perceived differences were statistically significant. A two-tailed paired t-test (with Pearson correlation for related means) reveals that the hypothesis of equal means is accepted for the English VISL group ($t=0.7042$, $df=17$, $p=0.4908$), the English NON-VISL group ($t=0.6193$, $df=14$, $p=0.5456$), the English BASIS group ($t=1.8257$, $df=20$, $p=0.0828$), the Cand. Negot. NON-VISL ($t=-.6895$, $df=19$, $p=0.1074$), the Cand. Negot. BASIS ($t=1.9215$, $df=14$, $p=0.0752$), but not for the Cand. Negot. VISL group, which makes it the only group which has a significant change from pre-test to post-test ($t=2.4654$, $df=17$, $p=0.0246$). The success rates of all the other groups are not significantly different from the pre-test results, including the English BASIS group results. In fact, the characterising feature is uniformity of results across groups and across cohorts with a visible but statistically insignificant advantage to the English cohort. This uniformity allows for a fairly certain assumption that the results are valid and that the conclusion of no significant change holds true with the one exception of the Cand. Negot. VISL group although the p-values of the English BASIS and the Cand. Negot. BASIS are close to the 0.05 significance level.

The t-test results build on the assumption that the observations are normally distributed. Whether that is the case was investigated through a one-sample Kolmogorov-Smirnov test. The post-test non-syllabus results were normally distributed with the exception of post-test observations of the Cand. Negot. VISL groups ($p=0.039$). For the other experimental groups the hypothesis of the normal distribution of observations was formally accepted, but the value for the Cand. Negot. BASIS is so small that a supplementary non-parametric (Wilcoxon) test was carried out for the non-syllabus results of all experimental groups. The Wilcoxon signed rank test, however, gave basically the same results as the paired t-test in the respect that the Cand. Negot. VISL group is the only group which has a statistically significant change from pre-test to post-

test ($p=0.33$, two-tailed). If the one-tailed values were included, the English BASIS group ($p=0.042$, one-tailed) and the Cand. Negot. NON-VISL ($p=0.067$, one-tailed) were very close to being significant.

The non-syllabus success rate for the English BASIS group (77.25%) is the highest of all six experimental groups. The remaining five groups achieve fairly similar success rates (from 66.67% to 72.84%). Disregarding the English BASIS group, the two VISL groups achieve the best results, but the other groups are not far behind. However, actual score point figures here are quite small and therefore a slight difference in score might look larger in per cent. In that light – and compared to the syllabus section – uniformity of achievement levels seems to be a general feature which is present not only in means but indeed across all quartiles, signifying equality of distribution across the board.

The content of the experimental treatment was aimed towards improving the metalinguistic knowledge and metalanguage of the students. The non-syllabus criterion for success must therefore be different from the one applying to the syllabus section items. Research question 5 queried the connection between the instruction and the non-syllabus proficiency. The non-syllabus pre-test level was high compared to the syllabus pre-test level, and the criterion of complying with the 50% success rate mark cannot apply to the non-syllabus section as this was surpassed by all experimental groups before the experimental instruction. Further, following from the fact that the instruction was in the syllabus content, the criterion for success in the non-syllabus section needs to be a status quo from pre-test and post-test levels.

The Cand. Negot. VISL group in the non-syllabus section distinguishes itself from the other groups notwithstanding the general characteristics of the results of this section, which is the absence of a statistically demonstrated differential effect with the above-mentioned exception.

The issue raised in research question 5 of the influence of the treatment on the non-syllabus results can only be answered tentatively. The paired t-test for related variables

and the Wilcoxon signed rank test show that only the Cand. Negot. VISL group had a post-test result which could be said to be significantly different from the pre-test result for that group. There does seem to be a modest influence, but this might just be coincidence, except for the Cand. Negot. VISL group. There are several observations of a negative relation, unlike the syllabus section results. The modestly positive results combined with the negative results may point to influences which are less than systematic and may be derived from other sources than the experimental treatment.

Table 4.51
Overview mean values non-syllabus (in %).⁷⁹

Student	Treatment			Total
	BASIS	NON-VISL	VISL	
Cand. Negot	59.26	57.78	62.96	59.96
	68.15	66.67	72.84	69.18
	8.89	8.89	9.88	9.22
English	70.90	69.63	69.75	70.16
	77.25	72.59	72.84	74.49
	6.35	2.96	3.09	4.32
Total	66.05	62.86	66.36	65.11
	73.46	69.21	72.84	71.86
	7.41	6.35	6.48	6.75

1. number: pre, 2. number: post, 3. number: post - pre

The subjects are exposed to the English language in many hours outside the experimental classes, and any improvement in general proficiency, of which the non-syllabus items are an expression, may be the result of an influence from this environment and interaction in and with the English language. Indeed, should there be an influence from the syllabus instruction on the non-syllabus results, it seems that the VISL treatment could prove more effective than the other experimental treatments.

A further comment is that the Cand. Negot. VISL and NON-VISL, and the English BASIS groups are the three groups which have the highest non-syllabus scores. This is

⁷⁹ The figures in this table give the exact two decimal figures whereas the original tables in chapters 4.6.2-4.6.3 including the percentiles have been rounded up or down in the second decimal.

highly interesting and might occasion interest in investigating further if there might be some level of proficiency which would be a threshold for this effect to be possible or what other factors might be at work here.

When the syllabus success rates are compared with the non-syllabus success rates, the non-syllabus success rates appear to express a internalised proficiency level in contrast to which the syllabus success rates seem less impressive (see Tables 4.50 and 4.51). The total mean values for the syllabus section of test items are lower than the mean values for the non-syllabus section success rates. Further, the treatments have not managed to change the pre-treatment position of the non-syllabus section having a general advantage over the syllabus section.

4.6.5 Gain rate differences in syllabus and non-syllabus results

Gain rates are different from success rates in that the pre-test level is a factor. High pre-test levels may mean that there is little room for improvement and thus a gain result may be low, and vice versa low-pre-test results. One might also argue, however, that the group or subject who from a low starting point is able to overtake groups or other subjects who started at a low level and reach a high post-test level signifies a true improvement. Yet, as a supplement to the post-test results, gains are a measure for testing the efficacy of the various experimental treatments.

4.6.5.1 VISL, NON-VISL, and BASIS syllabus gain rate differences

The syllabus gain rates of the Cand. Negot. cohort are hierarchical in the sense that the Cand. Negot. VISL group has the highest gain, the NON-VISL result is intermediate, and the Cand. Negot. BASIS has the lowest gain. This might indicate a differential effect of the treatments (see Tables 4.33 and 4.50). The gain in the VISL group (39%) is more than double that of the BASIS group (16.70%), with the NON-VISL gain close to the VISL gain (33.8%).

The English syllabus gains did not vary as much between groups as the Cand. Negot. gains did. The English NON-VISL and the VISL gains were close, but the NON-VISL

gain was higher than the VISL gain. The English BASIS group gain was the highest of all (see Tables 4.34 and 4.50).

The result of an ANOVA test of significance of the between group differences is given in Tables 4.35 (Cand. Negot. cohort) and 4.36 (English cohort). These tables demonstrate that the difference between Cand. Negot. VISL and NON-VISL is non-significant ($p=0.533$), but the difference is significant between VISL and BASIS groups ($p=0.000$) and the NON-VISL and BASIS groups ($p=0.004$). The between group gain differences in the English cohort are not significant ($p=0.1142$).

4.6.5.2 VISL, NON-VISL, and BASIS non-syllabus gain rate differences

A quite remarkable feature, and a contrast to the syllabus section, is the negative gains (losses) that we witness in the minimum scores of both cohorts and even in the 25% quartiles of the English cohort (see Table 4.47). Equally noteworthy is the fact that the Cand. Negot. 25% quartile for all experimental non-syllabus groups are all zero gains (see Table 4.46).

The pattern of uniformity recurs in the between group differences of the two cohorts (see Tables 4.48 and 4.49) in that neither the Cand. Negot. between group gain differences ($p=0.9856$) nor the English between group gain differences ($p=0.7962$) are statistically significant. In fact, the characterising feature is uniformity of results across groups and across cohorts with a visible but statistically insignificant advantage to the English cohort. This uniformity allows for a fairly certain assumption that the results are valid and that the conclusion of no significant difference in non-syllabus gains between the experimental groups and cohort.

4.6.5.3 Correlating syllabus and non-syllabus gains

The gain results were non-significant for the non-syllabus test items for both cohorts in contrast to the syllabus test item results, where the general picture was one of significance but with marked variation between group variation. If the difference in gain results of the two sections of test items is tabulated (see Tables 4.52 and 4.53) we get a

measure of how much more the treatments appear to be beneficial with regard to the syllabus items than the non-syllabus items.

Table 4.52
Cand. Negot. cohort. Difference of differences: syllabus %difference (post-pre) minus non-syllabus %difference (post-pre).

treat	N	mean	sd	min	p25	p50
BASIS	15	7.81	20.7	-33.60	-1.22	11.40
NON-VISL	20	24.89	23.4	-20.70	14.00	24.10
VISL	18	29.14	24.9	-12.90	10.50	26.40
Total	53	21.51	24.40	-33.60	7.94	22.30

treat	p75	max
BASIS	26.60	39.90
NON-VISL	33.40	70.30
VISL	46.70	81.50
Total	34.30	81.50

Table 4.53
English cohort. Difference of differences: syllabus %difference (post-pre) minus non-syllabus %difference.

Treatment	N	mean	sd	min	p25	p50
BASIS	21	14.14	20.6	-13.10	-1.31	12.40
NON-VISL	15	27.32	25.5	-2.22	2.93	23.20
VISL	18	25.77	27.4	-31.60	7.53	30.70
Total	54	21.68	24.7	-31.60	3.14	15.30

Treatment	p75	max
BASIS	22.60	59.40
NON-VISL	50.90	64.70
VISL	48.40	60.70
Total	45.50	64.70

The mean effect is approximately the same for the Cand. Negot. cohort (21.51%) as for

the English cohort (21.68%). Also similar for the two cohorts is the fact that the gain per cent differences of the two BASIS groups are distinctly different from the gain differences of the other groups with the English gain difference of 14.14% almost double that of the Cand. Negot. 7.81% difference.

The differences of the gains in the other groups are fairly similar with the familiar pattern of the Cand. Negot. VISL group having an advantage over the NON-VISL group and the English NON-VISL group having an advantage over the VISL group.

In the same fashion, it is possible to pool the two BASIS groups, the two NON-VISL groups, and the two VISL groups and the percentage point gain differences for the syllabus and the non-syllabus test item section (see Table 4.54). This shows the familiar difference and progression from the BASIS, to the NON-VISL, and the VISL experimental group.

Table 4.54
Total difference by gains (%) syllabus and non-syllabus.

Student	Treatment		
	BASIS	non-visl	visl
syllabus	18.91	32.28	33.94
non-syllabus	7.41	6.35	6.48
Total	11.50	25.93	27.46

This relationship between syllabus and non-syllabus results can be demonstrated through a correlation test (Pearson), which tests whether those who achieve high gain results in the syllabus section will also score high gain results in the non-syllabus section. The observations are pooled for the two cohorts. For the Cand. Negot. cohort the Pearson correlation was positive ($r = .101$), but non-significant ($p = 0.474$, two-tailed). This indicates a certain relationship between the gains in syllabus and non-syllabus.

For the English cohort the correlation was negative ($r = -.119$), but insignificant ($p = 0.465$). This indicates the reverse relationship between the two test sections in that there is a weak tendency for those who do well in the syllabus section to not do well in the non-syllabus section.

The answer to research question 5, which queried the relationship between the syllabus instruction and the non-syllabus results, can be further informed by this result. It was noted in Chapter 4.6.3.3 that the test items build on the assumption that the non-syllabus test items tap into a different kind of knowledge than the syllabus test items (see Chapters 2 and 5). It was further noted in Chapter 4.6.4.1.2 that only the groups which had high syllabus levels, e.g. the Cand. Negot. VISL group, seemed to have gains in the non-syllabus section. Pearson's correlation test in one respect supports this as the Cand. Negot. cohort has a positive relation between syllabus and non-syllabus level. On the other hand, it seems to contradict the assumption, as the English cohort had a high initial level which the cohort was able to sustain. However, as far as gains are concerned the Cand. Negot. mean gains are higher than those of the English cohort. Tables 4.50 and 4.51 may contain some of the explanation in that these two tables show that for the lower end of observations, the non-syllabus gains are larger than the syllabus gains (for the BASIS groups of both cohort the 25% quartile is negative) and for the upper quartile the Cand. Negot. outperforms the English cohort.

4.6.6 Results of discrete test item results

The analysis of the overall quantitative test results demonstrated that the two sections of test items evidenced differential results with regard to the efficacy of the treatments, in that the VISL treatment seemed to have a modest advantage over the other two treatments. Further, the analysis illustrated a difference between the progress in the syllabus section of items and the non-syllabus section in terms of gain achievements. In either section the results of individual items gave rise to a need for a more detailed analysis.

The syllabus and the non-syllabus items are manifestations of two heterogeneous or complementary fields of learning and knowledge. The syllabus section can be described

as an expression of metalinguistic knowledge which is related to explicit or declarative types of knowledge. The non-syllabus section items on the other hand can be described as an expression of linguistic proficiency or procedural knowledge (see Chapter 5 for a detailed discussion). The treatments have as their main objective a modification or manipulation of the explicit or declarative sphere, and the assessment instrument in the form of the pre-test and the post-test was designed to measure any subsequent change in this area (for a detailed description of the tests see Chapter 4.5). Equally, the tests were designed to be a sensitive instrument which to some degree would be able to measure any potential influence of the treatments, as a co-effect, in an area which was not being targeted by the treatments, i.e. procedural knowledge.

An assessment of the results of the experiments has several dimensions: first, the efficacy of the discipline over the three treatments, which is classified as the effect of the instrument; second, there is the measurement of knowledge in particular items of grammar, some of which were present at the outset of the experiment, and how this knowledge is subject to change over the course of instruction (see Appendix VII for pooled item results of pre-test and post-test and Appendix VIII for pooled results of syllabus and non-syllabus item results of pre-test and post-test).

In the following the overall test results will be used as an indicator as to whether this knowledge was affected in the course of the experiment and specifically whether any method(s) or student type are particularly susceptible to modification. This entails a description and discussion of the teachability and learnability of the subject matter.

4.6.6.1 Initial knowledge

4.6.6.1.1 Initial knowledge: syllabus

Of the sixteen syllabus items only five, items 4 (verb-noun distinction), 11 (noun), 8 (adjective), 12 (preposition), and 19 (subject), can be said to represent a solid knowledge base in overall terms (see Table 4.55).

The overall pre-test result (see Table 4.5, 4.55 og 4.56) makes it clear that the non-syllabus items have a dominant position at the top of the rank. When the syllabus and

the non-syllabus items are separated in their own tables the ranking in each section and the interrelationship between the items become clearer and for this purpose the tables are separated below.

The subject matter of the top five syllabus items is fundamental to any language instruction, namely knowledge of the characteristics of nouns, verbs, adjectives, prepositions and the sentence constituent of subject (see Tables 4.5 og 4.55 or Appendix VII). The knowledge of these elements of grammar is evidenced in a descending scale related to the above-mentioned items in pre-test results varying from 90% to 50% of all students, a proportion which allows for the assumption that this is stable and solid knowledge.

The middle group of pre-test result items to which between 30% and 50% of subjects give the correct answer contains six items, namely items 17 (adjective), 13 (direct object), 18 (adverb), 24 (substantival use of noun), 15 (indirect object), and 5 (complement to subject).

Table 4.55
Easiness of syllabus items at pre-test.

Pre-score syllabus (n=107)

	item	syllabus	sum	percent
1.	4	syllabus	97	90.65
2.	11	syllabus	77.86	72.77
3.	8	syllabus	65	60.75
4.	12	syllabus	61	57.01
5.	19	syllabus	53.6	50.09
6.	17	syllabus	48	44.86
7.	13	syllabus	46	42.99
8.	18	syllabus	35.6	33.27
9.	24	syllabus	35	32.71
10.	15	syllabus	33	30.84
11.	5	syllabus	32	29.91
12.	16	syllabus	32	29.91
13.	7	syllabus	15	14.02
14.	20	syllabus	11	10.28
15.	14	syllabus	6.33	5.92
16.	22	syllabus	5.95	5.56

Fewer than half of the subjects can recognise the sentence constituents Od, Oi, Cs. The subjects are not asked to identify the constituents, only to choose a sentence in which they recognise the function. Direct object is recognised by 46 subjects, but the complement to the subject is recognised by only 32 subjects out of a total of 107. Doubtless, the term 'complement to the subject' is unknown to many respondents, and the situation is likely to be the same with regard to 'indirect object', yet the wordlist provided did give the Latin-based and Danish equivalents as well as the English term. This would be a help to those familiar with the concept but to whom the term was unfamiliar. However, it could be assumed that these concepts were unfamiliar to some respondents. In the case of the word classes the successful items (over 50%) have demonstrated that generally speaking, subjects are able to recognise/identify nouns and adjectives, but the borderline case of a substantival use of an adjective poses problems to most students, which is evident in item 24 since only 35 subjects, or 29.91%, are able to identify the correct answer. The items in this middle group represent a knowledge which is precarious and limited.

Items 7, 20, 14 and 22 reveal areas which few students have any knowledge of at all. These are direct object, pronouns, co-ordinating conjunctions, and adverbials. The means range from 14.2 % for direct object to 5.56% for adverbials (item 22). The knowledge of co-ordinating conjunctions is equally low at 5.92% (item 14) with that of pronouns slightly better at 10.28% (item 20). The low figure for pronouns is a surprise, since they include common ones such as you, they, my, his, etc., and the respondents were asked to identify pronouns not categorise them. A closer look at the results reveals that out of the 107 subjects 77, or 71.96%, had not identified one single pronoun. In the light of the high error rate, one tentative explanation could be that the concept, not just the term, is unknown to the majority of students. It is more understandable that the co-ordinating conjunction is an unknown concept since this represents a higher level of linguistic specialisation. For both categories it is important to keep in mind that the knowledge measured is a metalinguistic knowledge, and in this context unrelated to the use of these items. The coordinating conjunctions the students were asked to identify were three instances of the conjunction 'and', which is probably one of the first words a learner of English as a foreign language comes to master.

As far as adverbials are concerned, it is appropriate to compare the results of item 22 (identification of the sentence constituent adverbial) with item 18 (categorisation of discrete adverbs). Adverbials pose a genuine problem area whereas adverbs appear less problematic. Adverbials are also tested for in the non-syllabus section (items 10, 21, 25), and results here are laid out in the section below, but success rates are much higher in the latter items.

The sentence constituent of direct object as tested in item 7 has a low success rate of 14.02 %, which is much lower than the success rate for direct object as tested in item 13, which showed a success rate of 42.99%. From this result it may be ascertained that the concept is generally familiar, and that the explanation for the difference in success rate must be sought in other aspects of the two items. First of all there is a difference in the format of the two items in that item 13 requires less precise identification (the respondents were asked to select the sentence containing the direct object) than item 7 (the respondents were asked to identify by underlining the direct object in a given sentence). Secondly, there is a difference in the form of the direct object (clause versus noun), i.e. a complex structure on the one hand and a simple structure on the other hand.

4.6.6.1.2 Initial knowledge: non-syllabus

The non-syllabus section comprises items to which subjects could respond intuitively drawing on knowledge which is procedural rather than declarative⁸⁰ when asked to judge the acceptability/grammaticality⁸¹ of the language presented to them in the sentences in question. All of these items had success rates in the pre-test above the 30% mark, and only three (items 6, 23, and 25) below the 50% mark (see Table 4.56).

These nine items represent the students' entrance level of general proficiency or competence in the two respective areas of grammar as tested in the pre-test (see Table 4.59 for the syntax item group, Table 4.60 for the morphology item group, and Table 4.61 for the tense/aspect group).

⁸⁰ For a theoretical exposition of the connection between declarative, procedural and implicit, explicit knowledge see Dienes and Perner, 1999. Chapter 5 contains a discussion of Dienes and Perner and other references to the subject.

⁸¹ See R.Ellis, 1991, and Odlin, 1993.

The most fundamental aspect, viz. that of subject-verb concord, turned out to be well-established in that item 1 has a pre-test rate of 88.79%, and item 9 has a success rate of 82.24%. The third item in the morphology category, item 3, which deals with adjective comparison, has a pre-test rate of 80.37%.

Table 4.56
Easiness of non-syllabus items at pre-test.

Pre-score non-syllabus (n=107):

	item	non-syllabus	sum	percent
1.	1	non-syllabus	95	88.79
2.	21	non-syllabus	94	87.85
3.	9	non-syllabus	88	82.24
4.	3	non-syllabus	86	80.37
5.	10	non-syllabus	66	61.68
6.	2	non-syllabus	63	58.88
7.	6	non-syllabus	49	45.79
8.	23	non-syllabus	46	42.99
9.	25	non-syllabus	40	37.38

There are three items which are related to adverbials, items 21, 10 and 25 (see Table 4.60). The lowest score is found in item 25 (37.38%). In contrast, item 21 has a success rate of 87.85 and item 10 one of 61.88%. The issue in items 21 and 10 is that of placement⁸², i.e. a syntactical problem, whereas item 25 is a morphological or word class problem dealing with the distinction between adverbs and adjectives. Item 2 (58.88%) is syntactical in nature (*that*-ellipsis). Items 6 (45.79%) and 23 (42.99%) deal with tense/aspect problems, i.e. simple past versus present perfect (item 6), and continuous versus simple tense forms (item 23), respectively.

4.6.6.1.3 Initial differences in syllabus and non-syllabus knowledge

The syllabus category comprises word-class items (nine items) and syntactical items (seven items). The pre-test success rate on the nine word-class items is satisfactory (see Table 4.57). Only pronouns (item 20) and co-ordinating conjunctions (item 14) have exceptionally low rates.

The pre-test rate for the six syllabus items comprising syntactical issues (see Table

⁸² See White, 1991a.

4.58) is somewhat lower. Among these items, the rates for direct object in clause form (item 7) and adverbials (item 22) fall under the 20% mark.

Table 4.57
Word class (syllabus) pre-test (%).

Item	Category word class syllabus	Pre(%)
4	noun/verb	90.65
8	adjectives	60.75
11	nouns	72.77
12	prepositions	57.01
14	conjunctions	5.92
17	adjectives	44.86
18	adverbs	33.27
20	pronouns	10.28
24	noun/adjective	32.71
Mean		51.03

In the non-syllabus category three types of issues are tested, viz. tense/aspect (items 6 and 23), morphology (items 1, 9, 3), and syntax (items 21, 10, 2, 25). The results of this group of syntactical items are relatively high, with the lowest value being 37.38% (item 25, adverb placement).

Table 4.58
Syntax items (syllabus) pre-test (%).

(n=107)

Item	Category syntax syllabus	Pre(%)
5	complement to S	29.91
7	direct object	14.02
13	direct object	42.99
15	indirect object	30.84
16	S of main clause	29.91
19	subject	50.09
22	adverbial	5.56
Mean		29.05

It is particularly interesting to compare the results of the declarative knowledge of adverbials in the syllabus section with the procedural knowledge of adverbials in the

non-syllabus section (see Table 4.59).

Table 4.59
Syntax (non-syllabus) pre-test (%).

(n=107)

Item	Category syntax non-syllabus	Pre(%)
2	that-ellipsis	58.88
10	adverbial	61.68
21	adverbial	87.85
25	adverbial	37.38
Mean		61.45

Even though adverb placement⁸³ (especially items 10 and 21 would be typical) is generally considered a difficult aspect of second language acquisition, the success rates in these two items are well beyond the chance level already at pre-test level (see Table 4.59). In contrast, the declarative (syllabus) knowledge of adverb placement (see Table 4.58) as exemplified in item 22 is extremely low - in fact almost non-existent (83 subjects have not identified one single adverbial); in other words, with regard to adverbials there is a clear distinction between procedural and declarative knowledge.

Table 4.60
Morphology (non-syllabus) pre-test (%).

(n=107)

Item	Category morphology non-syllabus	Pre(%)
1	S-P concord	88.79
3	S-P concord	80.37
9	comp. adjectives	82.24
Mean		83.80

The morphology items achieve the highest score of the genres. The procedural knowledge (as tested in the non-syllabus items) is high, especially where subject-verb concord is concerned. This is somewhat surprising since a certain consensus seems to

⁸³ For a discussion on adverb placement and adverbials see White, 1991a.

exist among SLA teachers that this is a problem area⁸⁴. The subject-verb concord test items require complex steps in analysis in that the issues of relative clauses and attraction might make choices harder. The items did not in any of the cases contain a subject consisting of just one noun (see Chapter 4.5.4).

Table 4.61
Tense/aspect (non-syllabus) pre-test (%).

(n=107)

Item	Category		Pre (%)
	syntax	non-syllabus	
6		tense/aspect	45.79
23		tense/aspect	42.99
Mean			44.39

Tense/aspect issues constitute an area where Danish (which was the L1 of nearly all of the subjects) and English have contrasting applications and expressions, and which Danish students therefore tend to find difficult or problematic. In this light the pre-test scores cannot be considered low although slightly more than half of the subjects still experience difficulties with this issue.

The general features as described above hold true for all the experimental treatment groups with the already mentioned exception of the English BASIS group, which is statistically different (better) from the other groups at pre-test level (see Chapter 4.5.2.1). Within the English cohort it is also noteworthy that the English NON-VISL subjects score much higher in items 6 and 23 (both tense/aspect) than do English VISL subjects (see Appendix VII). In the Cand. Negot. cohort the reverse pattern is seen for these two items in the Cand. Negot. NON-VISL group, whose subjects score a lower rate than the Cand. Negot. VISL and BASIS groups, but the difference is much smaller than it is in the English cohort.

⁸⁴ Færch, Haastруп & Phillipson (1984) investigated the errors in written language production of Danish learners of English in a comparison of error types at three different stages in the learning process: 8th grade, 1st year of “Gymnasium” language line, and 3rd year of “Gymnasium”, language line. At the 8th grade level there is no specific data for concord, but at the 1st year of “Gymnasium” the error rate for concord was 12.4% of grammatical errors, which made this error type top of the list; at 3rd year of Gymnasium the error rate had been reduced to 6.6% and the ranking position among the grammatical errors listed concord in 6th place of fourteen (p.106).

4.6.6.2 The efficacy of the treatment on post-test results of discrete items

The pooled post-test item results give an indication of the easiness or difficulty of each item and the area of grammar to which it belongs. From an overall point of view Table 4.6 (or see Appendix VII) makes it clear that the non-syllabus items continue to be represented with high success rates when compared to the syllabus success rates.

4.6.6.2.1 Syllabus post-test item results across treatment groups

The findings illustrated in the post-test result in Table 4.6 (see also Table 4.62 or Appendix VII) are measures of the efficacy of the treatments in the sense that the sum of knowledge is ranked according to the treatment effect on each item of grammar, and thus is not an expression of the richness of knowledge held by the individual learners with regard to the subject matter (see Chapters 4.6.1-4.6.4).

Table 4.62
Easiness of syllabus items at post-test.

(n=107)

	item	syllabus	sum	percent
1.	4	syllabus	98	91.59
2.	19	syllabus	94.05	87.90
4.	12	syllabus	94	87.85
3.	18	syllabus	91	85.05
5.	5	syllabus	88	82.24
6.	11	syllabus	86.44	80.79
7.	15	syllabus	86	80.37
8.	8	syllabus	85	79.44
9.	17	syllabus	78.15	73.04
10.	20	syllabus	64.3	60.09
11.	16	syllabus	61	57.01
12.	13	syllabus	55	51.40
13.	14	syllabus	54.69	51.11
14.	7	syllabus	44	41.12
15.	22	syllabus	31.9	29.81
16.	24	syllabus	28	26.17

The syllabus (declarative) section items consist of syntax items (5, 7, 13, 15, 16, 19, 22) and word class items (4, 8, 11, 12, 14, 17, 18, 20, 24). From an overall perspective the syllabus post-test results appear satisfactory in that only three items remain below the

50% mark (see Tables 4.6 and 4.62). The items are from different categories of grammar, i.e. two syntax items (item 7, which pertains to direct (clause) object, and item 22, which pertains to adverbials), and one word class item (item 24, which pertains to substantival use of an adjective). All three items contain some complexity.

The items in the syntax group have a lower mean value (61.41%) than the word class group, but it is well above the critical fifty per cent mark. However, there are two sub-themes which fall short of that, namely direct object (item 7) at 41.12%, and adverbial identification (item 22) which is particularly low at 29.81%.

Direct object, which is in the form of a single noun, as in item 13, has successfully been identified by 55 subjects (=51.11%), but the added complexity of the clause form makes item 7 more difficult. Again the Cand. Negot. BASIS groups have the most problems in both cases, i.e. the lowest scores and much lower than the other groups, but these two items represent a general problem for all groups, English, Cand. Negot. VISL, and NON-VISL alike. The direct object is also tested in item 13 and here the problems are less pronounced. The Cand. Negot. BASIS group, the English BASIS group and the Cand. Negot. NON-VISL group all have scores below 50% in contrast to the English VISL (72.22%) and NON-VISL (66.67%) groups which have high scores. The Cand. Negot. VISL group (55.56%) has somewhat lower scores, but still much higher than the three low-scoring groups. There is a marked difference in the two test items in that the direct object tested for in item 7 has the form of a clause whereas the direct object tested in item 13 has the form of a noun. This might lead to the difference in ability to identify the object simply because of the higher complexity and added number of processing steps. Another tentative explanation might be an uncertainty among the students about the clause as a form.

The problems which clearly exist for all experimental groups (see Appendix VII) with regard to adverbials as sentence constituents was obvious in the pre-test as well. Item 22, however, is one of the items with the highest gain in the course of the experimental period, though not enough for this knowledge to be really established in the minds of the students.

The constituent of subject is tested for in items 16 and 19. The treatment is less successful (57.01%) for item 16, seeing that the English NON-VISL group score is as low as 33.33%, lower than the 46.77% gain for the Cand. Negot. BASIS group. This item is more complex than item 19 in that the students are asked about the subject of the main clause. Since the score on the simple subject in item 19 (87.90%) is so high, the natural explanation might be that the problem in item 16 lies in identifying the main clause rather than in identifying the subject.

Table 4.63
Syntax items (syllabus) post-success rate (%).

(n=107)

Item	Category		Post (%)
	syntax	syllabus	
5	complement to S		82.24
7	direct object		41.12
13	direct object		51.40
15	indirect object		80.37
16	S of main clause		57.01
19	subject		87.90
22	adverbial		29.81
Mean			61.41

The genuine surprises are the very high score (80.37%) in item 15 (indirect object) and the equally high score (82.34) in item 5 (complement to the subject). Both of these items were among the top items with the highest gains (see Table 4.62) and the starting point for both items was approximately 30% in the pre-test. Remarkably, all experimental groups are successful in these two items and subjects (see Appendix VII), and even more remarkable is the fact that the success surmounts that of the direct object in item 13. No explanation readily offers itself as to the causes of this phenomenon⁸⁵. The VISL groups clearly outperform the NON-VISL groups in the syntax group of items (see Appendix VII).

⁸⁵ The think-aloud protocols (see chapter 7) are in contrast to this finding as the transcripts clearly show a general difficulty to recognise complements to the subject.

Item 24 deals with a borderline case between noun and adjective (adjectival use of a noun). The borderline nature is apparently too complex for some learners. Adjectives (items 8, 17) as such do not pose the same problem, and nouns (items 4, 11) are the category with the highest success rate (see Table 4.64). The fuzziness of word-class boundaries contained in item 24 is and remains a problem, which is demonstrated in the post-test success rates which for all experimental treatment groups are below 30%. In general the other items and categories of grammar have satisfactory success rate – also the pronouns (item 20) and co-ordinating conjunctions (item 14), which were a problem in the pre-test.

Table 4.64
Word class (syllabus) post success rate (%).

Item	word class	Category syllabus	post(%)
4		noun/verb	91.59
8		adjectives	79.44
11		nouns	80.79
12		prepositions	87.85
14		conjunctions	51.11
17		adjectives	73.04
18		adverbs	85.05
20		pronouns	60.09
24		noun/adjective	26.17
Mean			79.39

The conclusion with regard to word class items is that this is an area which is well-served by the instruction, and one that all experimental groups seem to be equally successful in mastering with the Cand. Negot. BASIS group as the least successful group.

4.6.6.2.2 Non-syllabus post-test item results across treatment groups

The change from pre-test to post-test in the non-syllabus section items is not as noticeable as in the syllabus section items. The ranking is basically the same in that items 23, 6, and 25 are still among the lowest ranking (see Tables 4.51 and 4.65). However, all non-syllabus items are above the 50% mark in success rates which items 23 and 25 were not in the pre-test (see Tables 4.51 and 4.56). The bottom four non-syllabus items

pertain to adverb tense/aspect (items 23 and 6), and adverb placement (items 10 and 25). Adverbials/adverbs are tested in a variety of ways, and the results do not give a unitary answer, perhaps especially as far as the non-syllabus section is concerned. It is interesting to see that the highest item score in the non-syllabus section (item 21) and the lowest (item 25) both pertain to adverb placement⁸⁶.

Table 4.65
Easiness of non-syllabus items at post-test.

(n=107)

	item	non-syllabus	sum	percent
1.	21	non-syllabus	97	90.65
2.	3	non-syllabus	91	85.05
3.	1	non-syllabus	90	84.11
4.	9	non-syllabus	87	81.31
5.	2	non-syllabus	82	76.64
6.	23	non-syllabus	68	63.55
7.	6	non-syllabus	67	62.62
8.	10	non-syllabus	56	52.34
9.	25	non-syllabus	54	50.47

The non-syllabus section consists of three categories of items, namely syntax (see Table 4.66), morphology (see Table 4.67), and tense/aspect (see Table 4.68),⁸⁷ of which only one has a direct counterpart in the syllabus section, viz. syntax.

The syntax category comprises four items, three of which deal with the sentence constituent adverbial. The item results cover a large span of values; especially items 10 and 25 have low scores in that for these two items it is true that only three of the six experimental groups have post-test scores above fifty per cent (see Appendix VII).

⁸⁶ In reality the two cases concern the sentence constituent of adverbial. In item 21 the issue is the placement of '*extremely badly*' and in item 25 the issue is the placement of '*hardly*'. The grammatical problem is referred to as adverb placement with reference to White's work on the matter (1991a).

⁸⁷ Some grammarians would include tense/aspect in the morphology category, but the pedagogical issues are clearly different, and the two tense/aspect items are therefore listed in a separate category here.

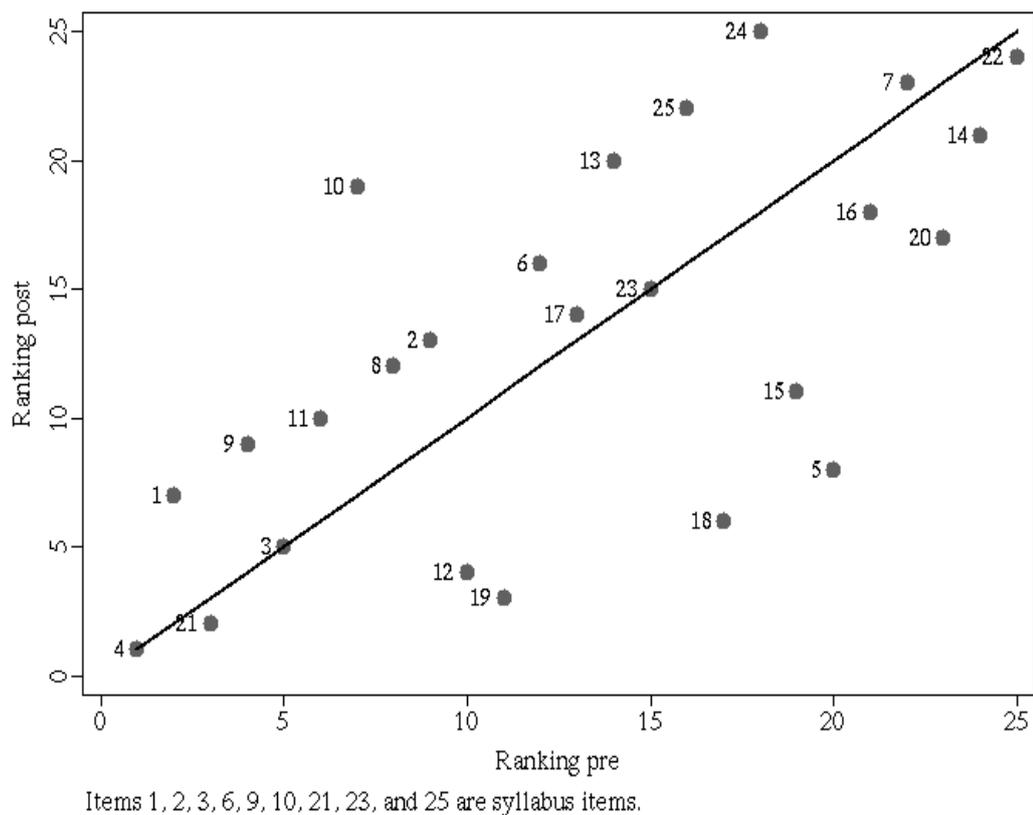


Figure 4.13 Item results pre-test ranking by post-test ranking

Table 4.66
Syntax (non-syllabus) post success rate (%).

(n=107)

Item	Category syntax non-syllabus	post (%)
2	that-ellipsis	77.64
10	adverbial	52.34
21	adverbial	90.65
25	adverbial	50.47
Mean		67.53

Despite the fact that item 10 has a higher mean value than item 25, there are various reasons why item 10 is especially problematic. One is that this item has experienced a negative gain (loss) of almost ten per cent from pre-test to post-test. The second is that the number of students who did not judge this item correctly increased by ten people; all

in all 51 of 107 subjects did not judge the item correctly. Adverbial placement is the issue in item 10 as it is in items 21 and 25.

Table 4.67
Morphology (non-syllabus) post success rate (%).

(n=107)

Item	Category		post (%)
	morphology	non-syllabus	
1	S-P concord		84.11
3	S-P concord		85.05
9	comp. adjectives		81.31
Mean			83.49

The morphology group of items has the highest success rate of all the categories, which is parallel to the situation at the entrance level, indicating that very little has happened during the experimental period. Item 1 shows a negative gain (loss) equivalent to the positive gain in item 3. However, these changes are so small that they might be mere fluctuations due to chance. In contrast, the constant high performance in all of these items must be significant since the results are so stable.

Table 4.68
Tense/aspect (non-syllabus) post success rate (%).

(n=107)

Item	Category		post (%)
	syntax	non-syllabus	
6	tense/aspect		62.62
23	tense/aspect		63.55
Mean			63.09

The tense/aspect category is represented by item 6 (mean 62.62%), in which the issue is the past-present perfect dichotomy, and item 23 (mean 63.55%) which contains the simple versus continuous tense form dichotomy⁸⁸. The success rates in the two items

⁸⁸ See Salaberry and Shirai, 2002 for an overview of acquisitional issues of tense-aspect.

are almost the same but the between group difference reveals that there is one group which is clearly different from the others, namely Cand. Negot. BASIS, which in both items has markedly lower success rates than the other groups (see Appendix VII). It appears that there might be an unintended or accidental effect of the experimental instruction on this category, unlike what was the case in the other procedural themes in the non-syllabus section. The English cohort does better than the Cand. Negot. cohort. All the English groups do well on this item and the Cand. Negot. VISL and NON-VISL groups have similar levels. It is only the Cand. Negot. BASIS groups which falls behind. There is no obvious explanation for the high success rates in this group of items.

4.6.6.3 The efficacy of the treatment on gain results of discrete items

The gain rates per item express the efficacy of the treatments per item rather than the effect on the individual subjects. The two angles of efficacy need to complement each other in order to examine the differences in treatment effects. The item gains, or the linguistic gains, were not the ultimate objects of the experiment, whose ultimate object is the learning potential of the experimental treatments. The linguistic items will therefore be scrutinised only as far as they inform this object. The relationship between pre-test, post-test and gain results is visualised in Figure 4.13.

Table 4.69 shows how the treatment affects the ranking of the test item with respect to their gains. This indicates that the syllabus items are more affected by the treatment than the non-syllabus items, which cluster at the bottom of the table when it comes to gain rates rather than success rates. Three non-syllabus items suffer losses after the treatment, and this is only the case with one syllabus item.

If the average gain results of all groups are used as a benchmark, we can see that the syllabus items have robust gain rates (see Table 4.69). The eleven highest-ranking gains feature in the syllabus section, and they all have two-digit gain percentage points. These eleven high-gain items comprise word classes (co-ordinating conjunctions, pronouns, adverbs, adjectives, and prepositions) as well as sentence constituents (adverbial, direct

Also Bardovi-Harlig (2002) has an extensive account of tense/aspect issues.

object, subject complement). Of the remaining fourteen items, five belong to the syllabus section and only one of these (item 8) has a two-digit gain (18.69%), while three have small gains from 0.93% to 8.41% (items 4, 1, 13). These items pertain to adjectives, direct object and nouns. The unsuccessful and last syllabus item is 24, which has a loss of -6.54%; it pertains to adjective/noun recognition or distinction.

Table 4.69
Item gain all groups.

(n=107)

	item	non_syl	points	percent
1.	5	syllabus	56	52.34
2.	18	syllabus	55.4	51.78
3.	20	syllabus	53.3	49.81
4.	15	syllabus	53	49.53
5.	14	syllabus	48.36	45.20
6.	19	syllabus	40.45	37.80
7.	12	syllabus	33	30.84
8.	17	syllabus	30.15	28.18
9.	16	syllabus	29	27.10
10.	7	syllabus	29	27.10
11.	22	syllabus	25.95	24.25
12.	23	non-syllabus	22	20.56
13.	8	syllabus	20	18.69
14.	2	non-syllabus	19	17.76
15.	6	non-syllabus	18	16.82
16.	25	non-syllabus	14	13.08
17.	13	syllabus	9	8.41
18.	11	syllabus	8.58	8.02
19.	3	non-syllabus	5	4.67
20.	21	non-syllabus	3	2.80
21.	4	syllabus	1	0.93
22.	9	non-syllabus	-1	-0.93
23.	1	non-syllabus	-5	-4.67
24.	24	syllabus	-7	-6.54
25.	10	non-syllabus	-10	-9.35

In spite of the variation in the discrete items there is a clear distinction between the results pertaining to the syllabus category (declarative or explicit metalinguistic knowledge) (see Table 4.70) and the results pertaining to the non-syllabus (procedural or implicit knowledge) category (see Table 4.73) in that the syllabus results on average clearly outrank the non-syllabus gains. The conspicuous fact is that the high-gain range does not comprise any of the non-syllabus items. Still, some of the non-syllabus items

contain respectable gains and will be discussed below. The difference in gains per item is visualised in Figure 4.14.

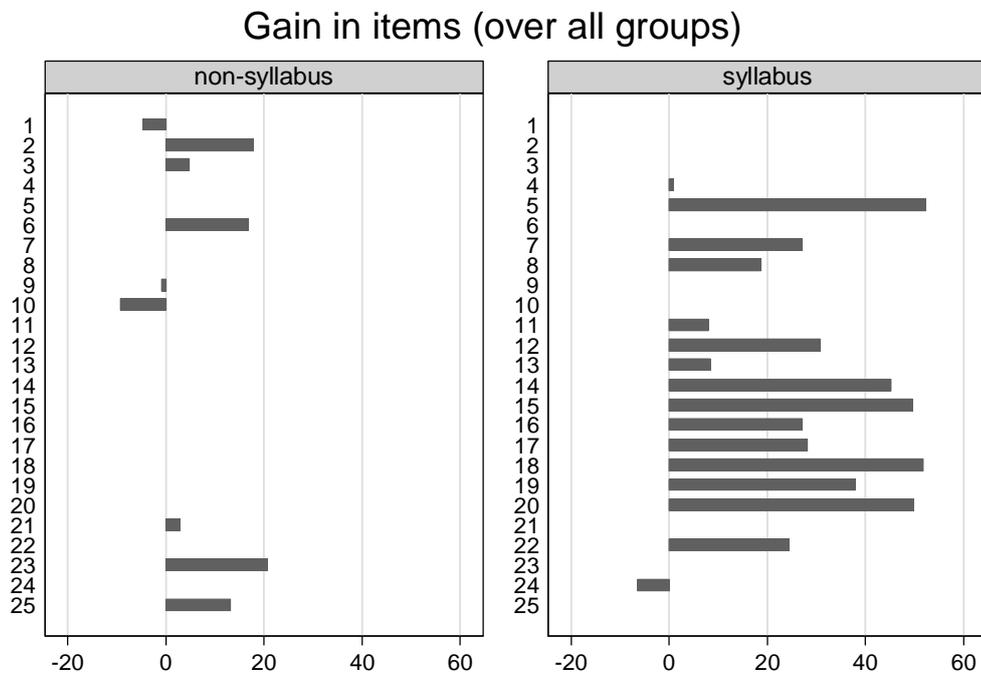


Figure 4.14 Gains per item, non-syllabus versus syllabus combined for all groups.

4.6.6.3.1 Syllabus item gain results

No doubt the spectacular gains in items 5, 14, 15, 18 and 20, which are all above 40 per cent points, may be due to a very low amount of knowledge at the entrance level. The items in question (see Table 4.69 and 4.70) comprise the sentence constituents of complement to the subject, co-ordinating conjunction, indirect object and the word class of adverbs. The pre-test sums of these items were in a range from 6.33 to 35.6 out of 107 possible (see Table 4.5 or Appendix VII).

The same situation seems to apply to items 7, 15, 16 and 22. This is not unexpected, indeed anticipations were that the pre-treatment knowledge in syllabus items would be limited, and the instruction was designed to remedy and abate this and bring the

students forward. This analysis of the tests confirms that the chosen instrument is a good vehicle for this knowledge.

Table 4.70
Syllabus gain results.

(n=107)

	item	non_syl	points	percent
1.	5	syllabus	56	52.34
2.	18	syllabus	55.4	51.78
3.	20	syllabus	53.3	49.81
4.	15	syllabus	53	49.53
5.	14	syllabus	48.36	45.20
6.	19	syllabus	40.45	37.80
7.	12	syllabus	33	30.84
8.	17	syllabus	30.15	28.18
9.	16	syllabus	29	27.10
10.	7	syllabus	29	27.10
11.	22	syllabus	25.95	24.25
12.	8	syllabus	20	18.69
13.	13	syllabus	9	8.41
14.	11	syllabus	8.58	8.02
15.	4	syllabus	1	0.93
16.	24	syllabus	-7	-6.54

Table 4.71
Word class (syllabus) gain (%).

Item	Category word class syllabus	Gain(%)
4	noun/verb	0.94
8	adjectives	18.69
11	nouns	8.02
12	prepositions	30.84
14	conjunctions	45.19
17	adjectives	28.18
18	adverbs	51.78
20	pronouns	49.81
24	noun/adjective	-6.54
Mean		25.21

Table 4.72
Syntax items (syllabus) gain (%).

Item	Category syntax syllabus	Gain(%)
5	complement to S	52.33
7	direct object	27.20
13	direct object	8.41
15	indirect object	49.53
16	S of main clause	27.10
19	subject	37.81
22	adverbial	24.25
Mean		32.45

4.6.6.3.2 Non-syllabus item gain results

The non-syllabus gains (see Table 4.73; also Appendix VIII) are concentrated in the lower half of the spectrum with two-digit gains for only three items (2, 6, 23) pertaining to tense/aspect (items 6 and 23) and that-ellipsis (item 2). Of the remaining six non-syllabus items, two have one-digit gains (items 3, 21), pertaining to adjectives and adverb placement, and the bottom three items (1, 9, 10) suffer losses ranging from –0.93% to –9.35%. The latter pertain to adverb placement and S-P concord. The effect for tense/aspect (items 6 and 23) and the lack of effect for S-P concord (items 1 and 3) seem established.

Table 4.73
Non-syllabus gain results.

(n=107)

	item	non_syl	points	percent
1.	23	non-syllabus	22	20.56
2.	2	non-syllabus	19	17.76
3.	6	non-syllabus	18	16.82
4.	25	non-syllabus	14	13.08
5.	3	non-syllabus	5	4.67
6.	21	non-syllabus	3	2.80
7.	9	non-syllabus	-1	-0.93
8.	1	non-syllabus	-5	-4.67
9.	10	non-syllabus	-10	-9.35

The effect on adverbial placement is less unified seeing that item 21 has a gain of 2.80 %, which is hardly comparable to item 10's negative gain of -9.35%, which in turn is

hardly comparable to the 17.76 % gain of item 2, notwithstanding the fact that all of these items pertain to adverbs/adverbials.

The gain ranking comparison across experimental groups (see Appendix IX) shows that for the Cand. Negot. BASIS group the distribution of non-syllabus items is spread out almost evenly over the scale of ranking when including all items with four items in the top half and five items in the bottom half, the first one as high as third place. The English BASIS group has no non-syllabus item before tenth place, item 2, which is also the first item in the Cand. Negot. BASIS group, and only three non-syllabus items in the top half of the rank. The distribution for the NON-VISL cohort is very similar in the two groups which both have one syllabus item in the top half rank; the Cand. Negot. NON-VISL has item 6 in ninth place and English NON-VISL has item 23 in eighth place. The second item for both groups is item 2 which is in fourteenth and fifteenth place respectively.

The VISL cohort (see Appendix IX) as such has no syllabus items in the top half rank, but the situation is more diverse when the groups are examined. The Cand. Negot. VISL group, which is the highest scoring group of all the experimental groups, has item 25 in eleventh place, and the English VISL group has item 23 in seventh place, and item 6 in twelfth place. Generally, the non-syllabus items are concentrated at the bottom of the gain rank. However, if non-syllabus items do appear among the top ranking items they are likely to be items 2, 6, 23 or 25.

The situation at entrance level (see Appendix VII) is that all but items 23 and 25 are in the top half of the scale (see Table 4.69 and Table 4.74). Thus, the post-treatment change in gain ranking may indicate that there is an accidental effect on some of these items, and this will be discussed below.

Item 10 suffers loss in five experimental groups and no gain in the Cand. Negot. VISL group. One might speculate about the reason why there is no loss for item 10 in the Cand. Negot. VISL group. One tentative explanation is that this group is generally so high-performing that all results are raised. The Cand. Negot. VISL group is also the one

with the lowest number of items which have no gain or negative gain (loss). Item 25 is highly affected in the VISL and BASIS groups of both cohorts, but not in the NON-VISL groups.

Table 4.74
Syntax (non-syllabus) gain (%).

Item	Category syntax non-syllabus	Gain(%)
2	that-ellipsis	18.78
10	adverbial	-9.34
21	adverbial	2.8
25	adverbial	13.09
Mean		6.08

Table 4.75
Morphology (non-syllabus) gain (%).

Item	Category morphology non-syllabus	Gain(%)
1	S-P concord	-4.68
3	S-P concord	-0.93
9	comp. adjectives	4.68
Mean		-0.31

The Cand. Negot. VISL group has only one loss item, which is item 9 with a loss of 11% affecting only three subjects out of 107 (see Appendix IX), and one item with zero gain - item 10 as described above. Item 9 belongs to the morphology group of items (see Table 4.75), a group in which the gain in percentage points collectively (-0.31) indicates a status quo pre- and post-treatment.

English VISL and English NON-VISL are the two groups with the highest number of non-syllabus items with no gain or negative gain (loss), namely five items and six items, respectively. In contrast, the Cand. Negot. VISL and the Cand. Negot. NON-VISL

groups are the highest achieving groups if measured by the lowest number of loss items and no gain items.

When the non-syllabus items are grouped and compared to the syllabus item groups the ranking in gains becomes clearer. It also becomes clear that the gain ranking is different from the post-test ranking (compare Tables 4.62 and 4.70; see also Appendix VIII).

The non-syllabus gain ranking, top and bottom, is basically the same across all treatment groups when the 107 item results are taken across all treatment groups (see Appendix IX). This indicates that the relative effect of the different experimental treatments is homogenous. It also indicates that the relative difficulty of these non-syllabus items remains the same. There are interesting details, though, when comparing post-test ranking and the gain ranking. This comparison reveals that items 6 and 23 are different from the other items; a difference which is compounded by the fact that they belong to the same grammar group, i.e. both items deal with tense/aspect. Items 6 and 23 are highly affected in all cohorts with the exceptions of item 6 in the English NON-VISL group (zero gain), and item 23 in the Cand. Negot. BASIS group (zero gain).

Table 4.76
Tense/aspect morphology (non-syllabus) gain (%).

Item	Category Syntax non-syllabus	Gain(%)
6	tense/aspect	16.83
23	tense/aspect	20.56
Mean		18.70

The student-based analysis in Chapter 4.6.2.2 established that, with the exception of the Cand. Negot. VISL group, there was no statistical difference between pre-test results and post-test results in the non-syllabus section. Therefore a closer examination of items 6 and 23 may throw some light on this issue which pertains to research question 5: *Can the syllabus instruction affect the non-syllabus results?* The ANOVAs on gains for the

syllabus and non-syllabus categories respectively (see Tables 4.35, 4.36 for Syllabus and 4.48, 4.49 for non-syllabus) can only inform us on the sections per se and not with regard to the discrete items.

It was therefore decided to conduct a post hoc statistical analysis of variances (ANOVA) of these two items against the other non-syllabus items. The result of the ANOVA was significant ($F(8, 106) = 3.59$) at the p-value of 0.0010. Supplementary tests were then carried out to pinpoint the difference and the three item-based groups of morphology, syntax and tense/aspect were tested against each other. The result was significant ($F(2, 106) = 7.73$) at a p-value of 0.0007. A comparison between the groups showed a significant difference (see Table 4.77 standard variation and p-value) between the tense/aspect group and the morphology group ($p=0.001$), and between the tense/aspect group and the syntax group ($p=0.025$), but not between the morphology and the syntax groups.

Table 4.77
Differences between non-syllabus item groups (n=107).

Row Mean- Col Mean	Morpholo.	Syntax
Syntax	.063863 0.295	
Tense/ aspect	.190031 0.001	.126168 0.025

Subsequently, ANOVAs were carried out to establish whether any differences between the experimental treatment groups were at a significant level. The tests were carried out separately for item 6 (see Table 4.78, standard variation and p-value) and for item 23 (see Table 4.79, standard variation and p-value). This turned out not to be the case for item 6 ($F(2, 106) = 0.48$; $p=0.6210$) as well as for item 23 ($F(2, 106) = 0.063$; $p=0.5346$).

Table 4.78
Differences (Scheffe) for item 6 between treatment groups (n=107).

Row Mean- Col Mean	BASIS	NON-VISL
NON-VISL	.116667 0.734	
VISL	.138889 0.642	.022222 0.989

Table 4.79
Differences (Scheffe) for item 23 between treatment groups (n=107).

Row Mean- Col Mean	BASIS	NON-VISL
NON-VISL	.174603 0.505	
VISL	.111111 0.755	-.063492 0.913

These results indicate that the possibility exists that the instruction in the syllabus content may influence the non-syllabus results as far as the tense/aspect category is concerned. The number of test items is so small that the result can only be an indication and more comprehensive investigations need to be carried out in order for the information to be firmly established. The subjects were exposed to many influences in addition to the experimental treatment and therefore some of the changes may be due to other causes than the treatment. Another effect of the small numbers may be the fact that it is impossible to see any statistical difference of the treatment according to treatment groups. This means that the observed differences between VISL, NON-VISL and BASIS cannot be said to be caused by factors other than natural variation in performance.

4.6.6.4 VISL, NON-VISL and BASIS discrete item result differences

The syllabus section contains two categories of items, viz. word class items and syntactical items. Both groups of items are classified as demanding declarative knowledge. In the pre-test the mean value of the scores in these sixteen items was

38.22% (see Table 4.50). In contrast the mean value of the non-syllabus section was 65.11% (see Table 4.51). The interesting question and measure of success is whether the instruction was focused enough to bring the syllabus section mean up to – and perhaps even beyond - the level of that of the non-syllabus section. The mean result of the syllabus section in the post-test was measured at 66.56% (see Table 4.50), which according to above mentioned criteria must be said to be satisfactory in that it came to approximately the same level as the non-syllabus section's post-test mean of 71.86% (see Table 4.51).

The experimental treatments, which were designed to affect the knowledge and understanding of grammatical concepts and terms, seem to be particularly well-suited to the task of accomplishing this with regard to the word classes and word-class distinctions as long as the class boundaries are not fuzzy as was seen in item 24. Item 24 was a borderline case of a substantival use of an adjective. These boundary problems between nouns and adjectives can be partly related to the fact that this issue receives little or no attention in the instruction at this stage, and apparently no transfer or inference takes place from knowledge about the two word classes. The combination of knowledge of either word class in a blend into one understanding of the conceptual characteristics, including similarities and differences, appears to be difficult. It should be noted that there is no problem with the noun-verb distinction (item19).

The items on syntax appear to be affected in the syllabus category (declarative or explicit), but very little in the non-syllabus category (procedural or implicit). However, the post-test results are almost equal in the two categories. With regard to the experimental groups there are only small differences between the VISL, NON-VISL and BASIS results. The results of the English and the Cand. Negot.cohorts are similar.

The tense/aspect group of items was only tested for in the non-syllabus (procedural or implicit) category, proved to be a special case among the non-syllabus items. Statistical tests showed significant gain results for the group compared to the other non-syllabus groups of items.

In an overall perspective of the efficacy of the instruction in general it seems that - with the exception of the Cand. Negot. BASIS group, the achievement of the students in acquiring the targeted metalinguistic knowledge is successful. The non-syllabus items, which in the pre-test ranking were in the uppermost end of the scale, have small or no gains, some even suffer negative gains (loss), and are therefore for the most part placed in the lower half of the rank (see Table 4.61).

There is a hierarchy in the successfulness of the different instructional methods in that the VISL cohort is more successful from the perspective of gains and from the perspective of bringing syllabus items forward. The NON-VISL cohort is slightly less successful at the top end of the scale, but less so at the bottom end of the scale. The two BASIS groups are so diverse that they cannot be described under one heading as the English BASIS is too different from the outset to be comparable. Consequently, it may be reasonable to let the Cand. Negot. BASIS serve as the normative group whose results without a doubt are less beneficial in reaching the set goal than the other groups and cohorts.

On a general scale the treatment or the instrument appears effective in the area it was designed to modify. It is clear that the syllabus items have large and significant gains. The non-syllabus section, with the possible exception of the tense/aspect⁸⁹ group, appears little affected by the treatment. Statistical analysis of the overall correlation of syllabus results to non-syllabus result gives a heterogeneous result. The Cand. Negot. cohort appears to have a small and non-significant but positive correlation in contrast to the English cohort where there is a small and non-significant but negative correlation. The influence, or instrument effect, of the syllabus instruction on the non-syllabus results appear to be precarious and invites further research.

⁸⁹ Salaberry & Shirai (eds.) (2002) refer to other studies which have yielded result that could indicate that verbal morphology acquisition may not be a morphological category acquisition issue. The learners tend to associate verbal morphology with lexical items rather than morphology (p.4). Bardovi-Harlig (2000:351) states in reference to her own and other studies: "We may conclude from these studies that the tense-aspect system is learnable in a classroom setting, but whether this result is due to the increased input or the specific noticing activities cannot be determined."

4.7 Attendance

The attendance of the subjects in all the experimental groups was recorded. The BASIS group attendance was not recorded and therefore no data are available for these students as will be apparent from Tables 4.80 to 4.84.

Table 4.80
Distribution of attendance all subjects.

Attendance(Freq)	Subjects	Per cent	Cum.
1-4	7	6.54	6.54
5-7	19	17.76	24.30
8-10	45	42.06	66.36
No data/BASIS	36	33.64	100.00
Total	107	100.00	

Table 4.81
NON-VISL treatment. Distribution of attendance.

Student	N	mean	sd	min	p5	p25	p50
Cand.Negot	20.00	7.25	2.20	3.00	4.00	5.00	7.50
English	15.00	8.20	1.52	5.00	5.00	8.00	8.00
Total	35.00	7.66	1.97	3.00	5.00	6.00	8.00

Student	p75	p95	max
Cand.Negot	9.00	10.00	10.00
English	9.00	10.00	10.00
Total	9.00	10.00	10.00

Cand. Negot.VISL as well as NON-VISL subjects have high attendance records with the mean value for NON-VISL at 7.25 and for VISL at 8.94, but distribution clearly shows that the Cand. Negot.VISL attendance is much higher (see Tables 4.81 and 4.82).

Table 4.82
VISL treatment. Distribution of attendance.

Student	N	mean	sd	min	p5	p25	p50
Cand.Negot	18.00	8.94	1.00	7.00	7.00	8.00	9.00
English	18.00	6.17	3.40	1.00	1.00	3.00	7.00
Total	36.00	7.56	2.84	1.00	1.00	7.00	9.00

Student	p75	p95	max
Cand.Negot	10.00	10.00	10.00
English	10.00	10.00	10.00
Total	10.00	10.00	10.00

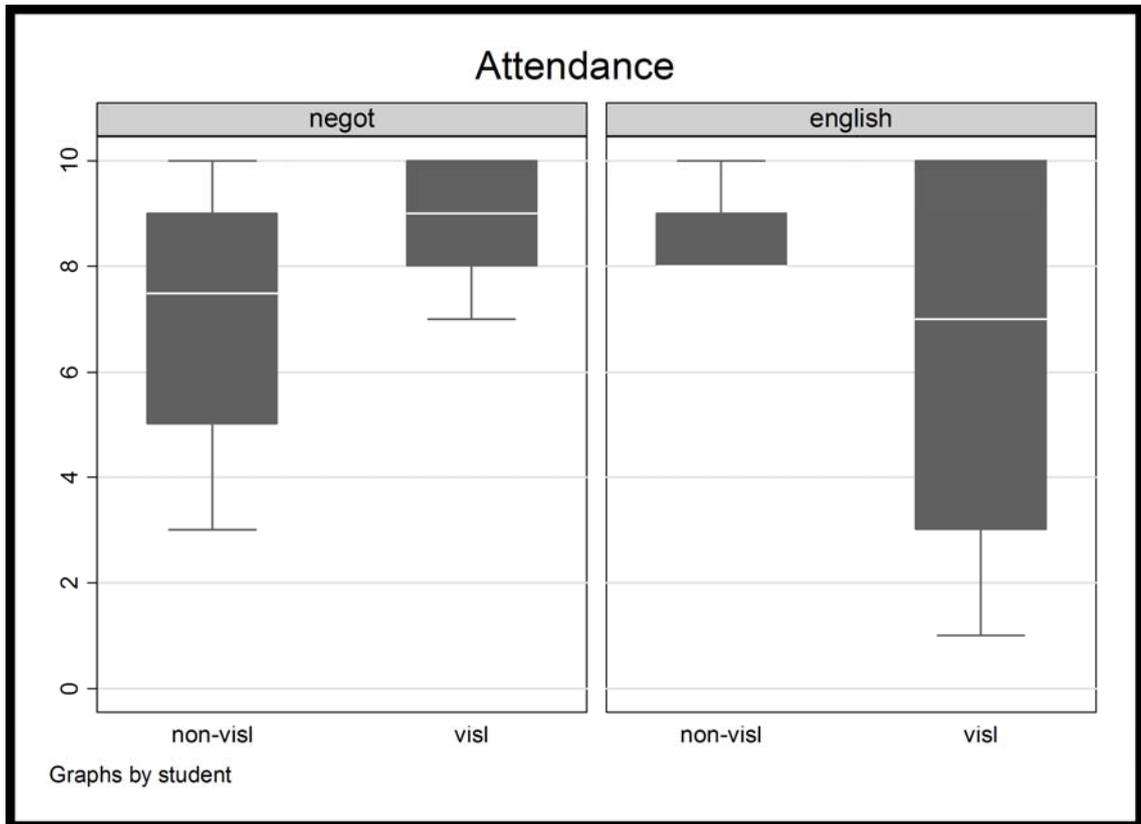


Figure 4.15 Boxplot of attendance by treatment group.

Table 4.83
Cand. Negot. High attendance distribution.

Attendance	Treatment		Total
	NON-VISL	VISL	
8	2	3	5
9	4	7	11
10	4	6	10
Total	10	16	26

The minimum for Cand. Negot. VISL is 7 and the 50% quartile is 9 in comparison to 3 and 7.50, respectively, for the NON-VISL subjects. In other words, the Cand. Negot.

VISL attendance is characterised as being high and concentrated in the high end of the scale, which is also expressed in the low SD of 1 in contrast to the SD for Cand. Negot. NON-VISL which is 2.20 indicating a higher spread for the Cand. Negot. NON-VISL attendance which shares the high attendance in general with the VISL group. The Cand. Negot. NON-VISL attendance is not as concentrated as the Cand. Negot. VISL attendance, however. The Cand. Negot. NON-VISL attendance comprises the low end of the scale as well as the high end of the scale. Tables 4.81 and 4.82 indicate the number of students with high attendance according to the treatment group, and it is clearly illustrated that the Cand. Negot. cohort has higher attendance than the English cohort at the high end of the scale, while especially the Cand. Negot. VISL group raises the attendance frequency.

Table 4.84
English. High attendance distribution.

Attendance	Treatment		Total
	NON-VISL	VISL	
8	5	1	6
9	4	1	5
10	3	5	8
Total	12	7	19

English VISL and NON-VISL subjects have unequal attendance rates, with English NON-VISL attendance rates being markedly higher than those found in the English VISL group (see Tables 4.81 and 4.82). The former has a mean of 8.20 and the latter a mean of 6.17. Furthermore, the spread among the English VISL group observations is more than double that found in the English NON-VISL group in that the SD for English NON-VISL is 1.52 and for VISL 3.40 (see also boxplot in Figure 4.13). The background for this is evident in the 5% percentile, where the attendance for English NON-VISL is at 5.00 and for English VISL only at 1.00, and the 25% quartile where the attendance for English NON-VISL is at 8.00 and for VISL at 3.00. The upper half is more equal with the difference at the 50% and 75% quartiles showing a reduction of the gap to 1. The fact that the English VISL median (7.00) deviates from the mean (6.17) to a much higher degree than those of the English NON-VISL as well as both of the Cand. Negot. groups hints that the distribution may not be a normal distribution.

The Cand. Negot. and English cohorts both have very different patterns of attendance for the VISL and the NON-VISL groups. The highest attendance is found in the English NON-VISL and the Cand. Negot. VISL groups. The English VISL and the Cand. Negot. NON-VISL groups show much larger variation and spread than the former. The distribution of attendance frequency is encouraging in the sense that the largest group of students is found in the high attendance range, i.e. 8 to 10 times with 10 being the maximum (see Tables 4.83 and 4.84). As many as 45 out of the 71 experimental subjects, or 63.36% attended. Attendance for experimental subjects who attended half the time or more is 64 out of 71, or 90% (see Table 4.80).

4.8 Summary of results

The experimental treatment groups of VISL, NON-VISL, and BASIS of the two cohorts of Cand. Negot. and English were examined in a pre-test/post-test design and results have been analysed from a student-based and an item-based perspective.

The student-based section first looks at the pre-test results of all items in each of the six groups in order to determine the initial level of knowledge and proficiency, and whether

the level in the experimental groups is of such a nature that the post-test results and the gain results can be compared in a meaningful fashion. All results are investigated with regard to total levels, and in separate levels for syllabus and non-syllabus sections. The pre-tests revealed that the English BASIS group has an initial level of knowledge which is much higher than that of the other groups, and that the English BASIS group as a consequence of this cannot serve as a proper control group for the VISL and NON-VISL results of the English cohort. The pre-tests further establish that the English cohort in general has a higher entrance level than that of the Cand. Negot. cohort.

The post-test and gain results reveal that the VISL treatment is highly effective and more so than the NON-VISL treatment, though not significantly so with the exception of the Cand. Negot. VISL group. The Cand. Negot. VISL group starts out at a lower level than the other groups and yet attains higher post-test and gain results than the others. The NON-VISL treatment also proves to be effective in comparison to the BASIS treatment. For the Cand. Negot. cohort there is a significant difference between the BASIS group and the NON-VISL and VISL group. Due to the high level in the English BASIS group this difference was not found for the English cohort. The most interesting result difference between the experimental treatments is found in scrutinising the results for the percentiles. It appears that the VISL treatment is better at raising the bottom level. This indicates the lower-achieving student benefit more from the VISL treatment than from any of the other treatments.

The difference between the syllabus and the non-syllabus results is quite marked. The instruction has a good effect on the syllabus items, but hardly any on the non-syllabus items. This is not unexpected as the instruction was designed to address the type of knowledge measured in the syllabus section, i.e. metalinguistic knowledge. The instruction was not aimed at improving or influencing general proficiency, which is what the non-syllabus section measures. That the results are so clear tend to support the view that the two types of knowledge are indeed different and are influenced and stored in different representations in the brain. In parallel to the syllabus section the Cand. Negot. VISL group result did show significant gains in the non-syllabus section.

The effect of the instruction on the items complements the student-based results. The purpose is not to investigate the learnability of particular linguistic features from an acquisition point of view; rather, the purpose is to evaluate the effect of the experimental treatments as such. The same treatment effects that are demonstrated in the student-based investigation appear in a similar fashion as regards the syllabus items. With regard to the non-syllabus items it becomes clear that the small effect that is measured in the syllabus section as a whole can be ascribed to significant improvements in the tense/aspect items. Syntax is measured in both the syllabus section and in the non-syllabus section. The syllabus group of syntax items shows improvement, but the non-syllabus group of items does not.

The conclusion is that the treatments are effective; the VISL treatment more so than the others. The treatments are effective in the areas they were designed to effective and only the VISL treatment appears to have an effect also on the non-syllabus section. The VISL treatment appears to have a higher impact on the lower achieving subjects. The results may have been influenced by the different pattern of attendance that can be observed for the groups as well as the cohorts. The English cohort has a lower attendance rate than the Cand. Negot. cohort does. This is particularly true for the English VISL group. This might be one reason why the English VISL group in general has slightly lower results than the Cand. Negot. VISL group.

PART THREE

5 PROCESSES OF LEARNING: REVIEW

Discussions in education, and especially in language learning, have in them an embedded element which concerns the nature of the learning processes. The processes by which learning in general takes place have been the object of investigation involving researchers in a variety of fields ranging from psychology, psycholinguistics, and neurology to educational research. The same is the case with language learning processes. At the centre of discussion is the issue of explicit and implicit learning and the connection, if any, between them. In the following I shall outline the role of consciousness, various definitions of implicit and explicit learning and knowledge, and subsequently some of the implications for language learning.

5.1 Consciousness

The discussion on implicit and explicit learning has been characterised by a variety of definitions of these two terms and by implication so have the associated areas of implicit knowledge and explicit knowledge. The difficulty of definition is partly due to differences in perception and construal of related concepts such as consciousness and awareness, their perceived content as well as the interaction between these concepts and their ascribed role in learning.

Any definition of implicit and explicit learning must include an indication of the role of consciousness pertaining to its ascribed influence on learning. It is this issue which seems to cause the great diversity of opinion with regard to learning. Cleeremans and Jiménez (2002) distinguish between cognitive and non-cognitive adaptations to changes in the environment maintaining that the non-cognitive adaptation, which occurs constantly in organic as well as in inorganic objects, cannot be recognised as learning. Even though an object reacts to and adapts to changes or influences in the environment with a subsequent permanent modification as a result, this cannot be classified as learning. A computer is a case in point. Computers can be programmed in such a way

that inductions can be causally efficacious in influencing its own set-up or that of the environment. A computer can even be programmed to re-programme itself in a pattern which may resemble learning. Cleeremans and Jiménez insist, however, that such a pattern of action and reaction does not constitute learning; nor does the environmental adaptation of which natural selection is the prime example.

There is ample evidence that learning can be the result of more than one mode of apprehension. It is less evident what these different modes consist of and what the quality is of the product that they result in. This opaqueness has led to a plethora of terms, such as incidental learning, implicit learning, procedural knowledge on the one hand and intentional learning, explicit learning, declarative knowledge on the other hand. These terms are sometimes used with a very specific frame of reference and sometimes as a more loosely described approximation of a given situation. This may be due to the ambiguity of such a central concept as consciousness. In their target article Perruchet and Vinter (2002a) have argued for their theory of consciousness, the self-organising consciousness (SOC), in which the cognitive unconscious has no place. They reach the conclusion that what may seem to be indications of unconsciousness are incidences of “self-organising properties of conscious representations” (ibid, p.327)⁹⁰. The disagreement in the field on whether learning without consciousness is possible or not is a continuous discussion though part of the disagreement seems to be a matter of definition of concepts and a variety in conceptions of observed phenomena, as well as disagreement about the terms which should be applied to the various concepts.

The challenge is to find a workable definition which will make us able to define learning and which will allow for the inclusion of the types of learning and knowledge of which we are not fully aware that we are engaged in. This opens for a scenario in which learning without full conscious awareness (in the sense that it cannot be verbalised) constitutes a possibility. Learning presupposes a cognitive system in that the condition of subjective experience involves thoughts and feelings not present in sheer adaptations in non-cognitive systems. Cleeremans and Jiménez (2002) give a lucid

⁹⁰ One common way of defining consciousness is to attach verbalisability to the contents. Perruchet and Vinter disagree with this: “The traditional collapsing of consciousness and language in many areas of research makes it necessary to emphasize that the contents of phenomenal experiences cannot be identified as verbalizable knowledge” (2002a: 328, note 4).

overview of the many facets of consciousness, and their proposition is that “..consciousness is not an all-or-nothing process or property but it affords many degrees and components” (Cleeremans and Jiménez, 2002:9). They make the case that it can be assumed that consciousness is graded and dynamic (see also Rosenthal 2000)⁹¹.

5.2 Implicit vs. explicit learning

Since the 1980s several empirical studies have investigated the notions of implicit and explicit learning and have tried to settle on definitions which could illuminate their differences concerning quality and potential. The above discussion of consciousness illustrates the problems involved in establishing commonly accepted definitions.

Psychologists are still in disagreement as to whether it is possible to make a clearly defined distinction, and there are varying perceptions of the issues. Stadler (1997) voices the opinion that “...if no such definition can be found, then the distinction between implicit and explicit learning will not be worth making.” (p.57). Perruchet and Vinter (2002a) claim that their model of self-organising consciousness can account for results stipulated as implicit learning (p.297), and Perruchet, Vinter, and Gallego (1997) state that “the very question of the implicit or explicit nature of the knowledge that emerges from a learning episode may not be meaningful” (p.43). This view should be interpreted to mean that they do not accept that learning can take place without consciousness.

Less stringently and using different terminology, Whittlesea and Dorker (1997) defend more or less the same view although with the concession that learning may take place under different circumstances. Any learning will result in new potential for new performance, but the individual may not hold a conscious idea of this new potential ⁹². Whittlesea and Dorker (1997) give the example of a person learning the word “road” and the word “street”. The two words are learned separately, and the person holds separate representations (or entries) of them. The ability to perform a comparison is not

⁹¹ The constructs of awareness and noticing and their role in learning are discussed in the review chapter on second language acquisition (Chapter 2).

⁹² To some degree this view parallels Perruchet and Vinter’s (2002a) SOC

computed, and is not represented, but the potential is there. This is not held by the person as new knowledge, nor as knowledge of new potential. Awareness comes only and if a situation should arise in which a given demand stimulates the person to call on this potential. Whittlesea and Dorker (1997) call this indirect learning, but they do not want to classify it as learning without consciousness. A similar example is given by Dienes and Perner (1999:736) to illustrate implicitness⁹³. This brings us back to the question of concepts, conceptions and interpretations of phenomena, and the terms applied to them. A problem raised by Eriksen (1960), who also discusses the role of verbal report as an answer to the definitional problem of awareness, “I have raised these objections and problems concerning the definition of awareness, not because I have a solution to offer, but solely to insure that the issues raised [...] will be evaluated in terms of the limitations of our current definitions” (p.281). This caveat appears to be necessary today as well.

In an overview Hulstijn (2005) offers concise definitions of terms related to the ones which are being discussed here. Hulstijn prefers to apply the terms ‘*incidental*’ and ‘*intentional*’ when referring to the process of learning. However, his definitions of the two appear quite similar to those of implicit and explicit learning. Interestingly, Hulstijn operates with two definitions of incidental learning, one for experimental situations where “participants are not forewarned of an upcoming retention test for a particular type of information” (p.132), and one for non-experimental situations which is more general in that it involves “the unintentional picking up of information” (p.132). The latter is very similar to his definition of implicit learning in that he makes “intention” the differentiating feature of explicit and implicit learning (p.131). Thus implicit learning, whose product will be implicit knowledge (Hulstijn, 2005:131-32), is the process which takes place without the learner having his or her attention directed towards particular features in the input. Either because they have been told the purpose was a different one or simply as a by-product in a given task or situation. In contrast, explicit learning, the product of which will be explicit knowledge, is the process which takes place when the learner has the “conscious intention to find out whether the input

⁹³ Dienes and Perner (1999:736) give “bachelor” as the example.

information contains regularities and, if so, to work out the concepts and rules with which these regularities can be captured” (Hulstijn, 2005:131).

The difficulties involved in separating implicit and explicit learning may be due to a possibility pointed out by Mathews (1997), namely that there may exist several different types of implicit as well as explicit learning whose properties may possess different qualities.

5.2.1 Implicit and explicit knowledge

Implicit learning produces implicit knowledge, and explicit learning produces explicit knowledge (Hulstijn, 2005). Explicit knowledge is often described as analysed knowledge (R. Ellis, 2004), which can be controlled and recalled at will. This is in contrast to implicit knowledge, which one may not be aware that one is in possession of, but which nevertheless may influence behaviour (Cleeremans and Jiménez, 2002). In light of this the difference is often given to be a contrast between procedural and declarative knowledge. Dienes and Perner (1999) apply a definition in line with the commonly accepted meaning of the words implicit and explicit in ordinary language. They illustrate this with the following example: “They didn’t say so explicitly; it was left implicit” (1999:736), and they argue, “...it seems to us, we should adhere to that existing meaning as far as possible and not impose some arbitrary “operational definition”, or else we make it difficult for the scientific community to share the same meaning, because the natural meaning is likely to keep intruding” (ibid.). Dienes and Perner’s (1999) theory of implicit and explicit knowledge comprises a model according to which it is possible to explicate various stages (or states) of explicitness and its components⁹⁴. The model presents the components and conditions for full explicitness in a hierarchy (see Figure 5.1) which presumes that an explicit representation at one level will mean that lower level components are also explicitly represented. Perner has

⁹⁴ Perruchet and Vinter (2002b) point out that it is quite common to associate consciousness with explicit knowledge, but this is a misconception. Consciousness is much more complex than that, and they link it to attention, “...the term “consciousness” designates the on-line content of the attentional focus, and not the explicit knowledge that subjects might have developed about the material, such as it might be revealed in post-experimental tests.” (2002:46, note 2). Berry and Dienes (1993) conclude that “One difficulty with free recall is that it may not elicit low confidence knowledge” (p.33). Reversely, however, it is not possible to verbalise knowledge of which one is not consciously aware.

suggested that in contrast to implicit knowledge, explicit knowledge is seen as compositional (discussed in Dienes and Berry, 1993:160 on the basis of a personal communication to them from Perner). The components of full explicitness are *Content*, *Attitude* and *Self* which themselves contain different structures and elements. A fully explicit piece of knowledge has content, i.e. a proposition, which the self holds to be true, or takes to be true at a given time.

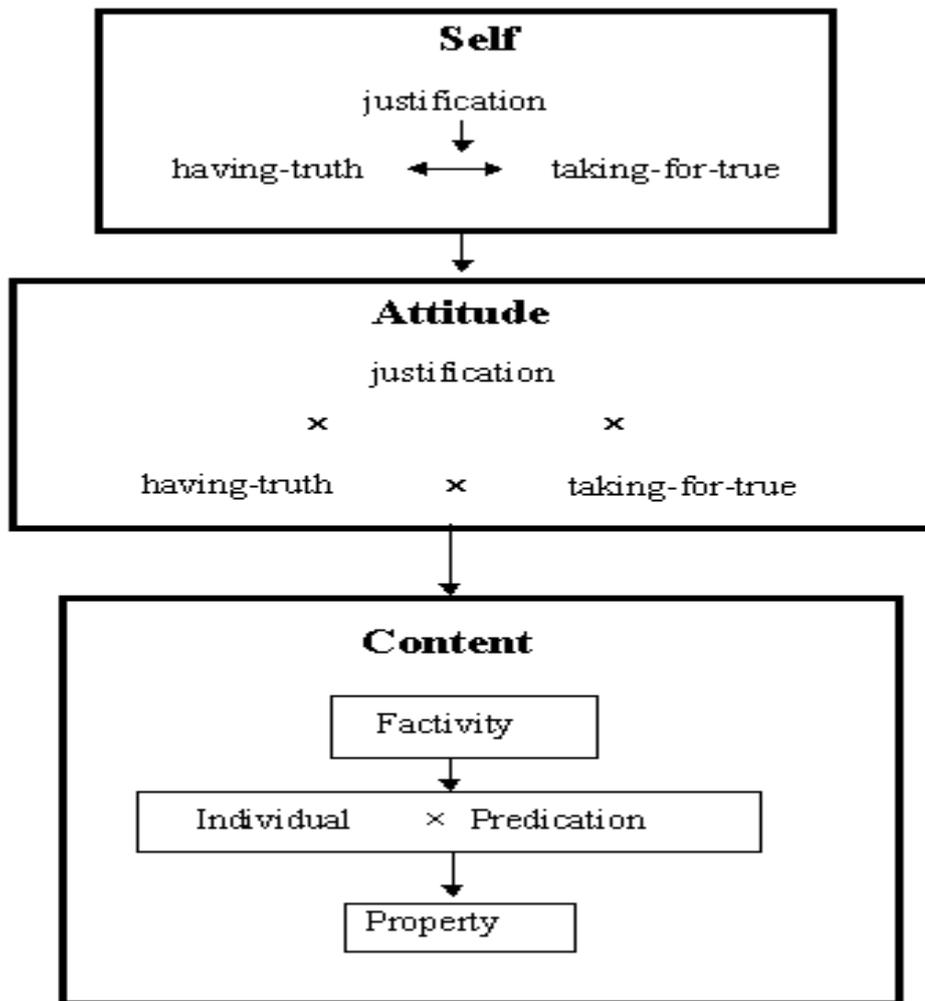


Figure 5.1 Graphic presentation of the hierarchy of explicitness
 Source: Dienes, Z., and Perner, J. (1999:740 Figure 1)⁹⁵

An explicit representation of a piece of knowledge may be accompanied by an implicit

⁹⁵ Dienes and Perner (1999:740. "Figure 1. Constraints on explicitness. An arrow denotes that explicitness of the item from which the arrow emanates entails explicitness of the item to which the arrow points. An "x" denotes that the explicitness of the two terms can be varied freely".

representation of the state of affairs. The example given by Dienes and Perner (1999, 2002) is the word “bachelor”. The proposition that someone is a bachelor implicitly presents that person as being male.

The model developed by Dienes and Perner emerged on the background of extensive and comparative studies of a host of psychological experiments as well as theoretical papers. They draw on insights gained in fields as diverse as phenomenology, memory studies and empirical studies of artificial language learning, among others. With their theoretical model they aim to answer some of the pertinent questions related to the implicit-explicit issues; first of all the model provides a clearly argued case for how explicit knowledge consists of several components, what these components are, and how it is possible that they in a given order become explicit before a representation may be said to have reached a level or stage where the combination of all the structures will constitute full explicitness of knowledge. The credibility of the model is underpinned by its comprehensiveness and the carefully detailed argumentation that is delineated in a structured approach which makes this model a logical choice as the basis for investigations of the think-aloud section of the present thesis (see Chapter 7.4). The model structure and the separate components (see Figure 5.1) at each stage of the process make it possible to operationalise the theory in a setting of actual learning situations.

5.2.2 The implicit-explicit interface

The interface between explicit and implicit knowledge is the object of unresolved discussion and a very contentious issue. The reason for this is that it is difficult to devise valid and reliable measures for these constructs, and consequently discussion tends to be based on assumptions and general beliefs rather than empirical evidence. The new scanning techniques give rise to optimism in some quarters for neurological evidence that might help confirm or discard various theories and perhaps give some pointers towards which areas might be fruitful for empirical experiments. The scanning techniques, such as PET scans and fMRI scans⁹⁶, make it possible to map the areas in the brain which are activated when subjects are carrying out experimental tasks.

⁹⁶ Positron emission tomography and functional magnetic resonance imaging, respectively.

There are three stances on the interface issue known as the no-interface position, strong-interface position, and the weak-interface position. The no-interface position entails two learning systems, one for explicit learning and one for implicit learning. Implicit knowledge has qualities which are different from those of explicit knowledge, and the usefulness and application of the two types are not the same. Explicit knowledge cannot be converted into implicit knowledge and vice versa. In contrast, it is assumed in the full interface position that knowledge of one type is freely convertible into the other type, given the right circumstances. One example of such circumstances could be repeated practice. The perspective is that the implicit/explicit distinction is not seen as a barrier to unhindered transfer between one type and the other. However, few contemporary scholars hold this view. A variety of the full interface position is the weak-interface position which holds that under certain circumstances implicit and explicit knowledge may be mutually convertible. The proponents of the weak-interface position do not think that practice may be the way to turn explicit knowledge into implicit knowledge. Rather, they argue that explicit knowledge may act as a facilitator for implicit learning. This is the position favoured by many second language acquisition researchers.

In a study of motor skills by Willingham and Goedert-Eschmann (1999) it was demonstrated that it is possible to learn skills implicitly at the same time as learning the same skills explicitly. The condition for this was that by pushing buttons the subjects acted out physically what they learned explicitly. The experiment showed that implicit knowledge can be developed in parallel with explicit knowledge. It was demonstrated that explicit learning may result in implicit knowledge as well as explicit knowledge if the right conditions are created which will allow it to happen. Thus, they found that implicit and explicit learning are not mutually exclusive (p.534). Since implicit knowledge is especially important for motor skills, this may be a result which could be true only for this particular area of learning, and the question is whether these findings can be said to have truth for other areas of learning, such as language learning. However, the study is interesting because it supports the notion that explicit learning may result in implicit as well as explicit knowledge, and that explicit knowledge is the facilitator of implicit knowledge. Willingham and Goedert-Eschmann themselves point to the importance

of their findings for the development of automaticity. They conclude that, "...the conscious, explicit process supports behaviour until the simultaneously acquired implicit representation is sufficiently well developed to support behaviour, at which time the explicit process is simply not used any longer; it does not transform into another representation" (p.534). In this they disagree with Anderson's (1993) ACT theory which holds that practice is causally efficacious in the transformation from one form of knowledge to another. Willingham and Goedert-Eschemann's view on automaticity is that it is characterised by demanding few attentional recourses which is also a characteristic of implicit learning; in this they agree with Anderson (1993) and Stadler (1995, 1997). Stadler (1997) comes to the conclusion that attention is important in the sense that the learner's intentional direction of attention may be the decisive factor in the difference between implicit and explicit learning, and he supports this with references to Dienes and Berry (1997) and Neal and Hesketh (1997a). Stadler (1997) proposes that explicit learning is the result of a hierarchical process (p.60) and that explicit knowledge is hierarchical. In this he is close to the theoretical foundation of Dienes and Perner (1999).

5.3 Artificial language learning

Disagreements about definitions, qualities, and interfaces have been quite as pronounced in language learning as it has in other fields of learning. Attempts towards theory building have sought foundations in natural language learning, including second language learning, as well as artificial language learning. Experimental studies in artificial language learning have been carried out by several scholars who have looked for ways to operationalise their theories of learning. The attractiveness of artificial language, or grammar, learning may lie in the fact that some variables are easier to control for and that results therefore may be more reliable. The most extensive research in this field has been carried out by Reber, and Reber and others, (1967, 1976, 1980, 1985, 1989, 1992, 1994, 1997), and he continuously examines the theoretical underpinnings of implicit versus explicit language learning. Reber's research also includes extensive reviews and examinations of studies by other researchers (1992, 1993, 1997). In Reber's experimental studies subjects were instructed to learn strings of

letters in a finite state grammar. The language consisted of a vocabulary which was constructed by combining the letters P, S, T, V, X (1967,1976) according to a set of rules which constituted the grammar. Strings, for instance TSSSXXVV, TXXVPXVV, PTVPXVV, PTTVPS (1976, 1980), consisted of between three and eight letters. Reber's grammar comprised 43 grammatical strings and 25 non-grammatical strings. In the experiment(s)⁹⁷ subjects were divided into groups which received implicit instruction, i.e. just told to learn and memorise as much as possible of the training stimuli, and groups which received explicit instruction, i.e. were told that there was a set of complex rules underlying the combination of letters, and told to look for those rules. One result (1980) was that the length of the strings had no effect on the ability of the subjects to classify the strings as either grammatical or non-grammatical. This indicates that subjects were not processing the 'sentences' in an explicit manner, but responded intuitively. New neuro-imaging research (fMRI) (Lieberman et al., 2004) appears to confirm Reber's (1969) finding that chunk learning as well as searching for rules play a role in grammar learning. Reber himself (1967) concluded that "the better performance of the three experimental groups in comparison with the control groups indicates that a blending of the two modes of learning [...] is still preferable to the use of one or the other" (p.501).

Reber's research has been most influential and has served as inspiration for replicationary investigations as well as further research and discussion. Especially the testing methods have been the focus of discussion. The tests used by Reber required subjects to respond positively or negatively to the question of whether the presented items were in accordance with underlying rules. Subjects were not asked what the rules were, only whether items could be said to follow a regularity in the grammar. If the performance of the subjects was at above-chance level, this was taken to indicate that implicit learning had taken place. The criticism raised against this method of testing has been the claim that once subjects were told that there was a rule-governed grammar they would tap into conscious processes (Perruchet and Pacteau, 1990) since the question itself referred to a previous learning episode, and this leaves open the possibility that the processes classified as implicit were not entirely without conscious awareness.

⁹⁷ The same finite state grammar was used in several experiments.

Perruchet and Pacteu set up three experiments based on the finite state grammar used by Reber and tested the knowledge of the subjects through grammaticality judgements. They found that the results might be explained not only by an unconscious representation of the grammar but by fragmentary conscious knowledge. Mathews and Cochran (1998) have further criticised artificial grammar learning per se "...the paradigm seems too artificial, too removed from anything people care about. Memorizing or judging strings of letters is a boring, tiresome task" (1998:223). And they continue "Even a pigeon can learn to make complex discriminations among stimuli" (1998:228).

The criticism of artificial grammar is mainly due to the fact that artificial grammar learning does not focus on generativity, which is what characterises natural language processing. Hulstijn (1997), in a review of laboratory studies, warns that "Because such research deliberate abstracts away from real-life learning situations, it simultaneously limits the possibilities to extrapolate their findings legitimately to real-life learning" (p.132). Hulstijn does stress, however, the usefulness of laboratory research in what he calls the "twin approach" (p.136), a combination of experiments with artificial or partly artificial language and a parallel experiment involving natural language. This method was applied by Hulstijn and de Graaff (1994) in their study of explicit and implicit knowledge and which conditions would best facilitate either type. The motivation for testing classroom research results in a more controlled laboratory setting is that the validity of the former can be supplemented with the higher reliability of the latter. The experiment reported by Hulstijn (1994) was set up to test the hypothesis that explicit knowledge facilitates the acquisition of implicit knowledge (see also de Graaff, 1997). Subjects followed either a self-study programme of an artificial language derived from Esperanto or a programme of four target structures in Spanish. The target structures were selected from Spanish, and the same structures were created in the artificial language. The artificial language subjects as well as the Spanish subjects were divided into two groups - one receiving explicit instruction and one receiving implicit instruction. Results from this experiment were reported in de Graaff (1997a, 1997b), and de Graaff concluded that explicit knowledge about language does have a facilitative

effect on implicit knowledge (1997b:165). The results were considered strong due to the replication of the experiments and the greater controllability of the artificial language learning element. Hulstijn (2005) has pointed out, though, one major difference between natural language learning and artificial language learning which one should bear in mind, namely that natural languages do not have the one-to-one relationship between form and meaning which characterises most machine languages.

5.4 Automatisatation

Explicit knowledge is often equated with verbalisability in contrast to the tacit nature of implicit knowledge. Automatisatation is interesting not only as a process but also as an expression of the interface between implicit and explicit knowledge. Accuracy and time are two parameters which are measurable variables related to automaticity. Implicit, internalised knowledge can be applied faster than controlled, explicit knowledge. One of the issues under discussion is whether practice can turn explicit knowledge into implicit knowledge (Sharwood Smith, 1981, 2004, 2005), or whether practice simply speeds up the explicit processes (Hulstijn, 2002a,b). Underlying this discussion is also the question of the qualitative differences between implicit and explicit knowledge and how a such difference could be measured. Explicit knowledge can be verbalised, but to reverse the definition to mean that knowledge which can be verbalised in explicit rules of language constitutes all of the explicitly represented may be a too simplified interpretation. The relationship is expressed by R. Ellis (2005), “Explicit knowledge is potentially verbalizable, although it exists in the minds of the learners independently of whether they can verbalize it” (p.150).

Dienes and Perner’s model (1999) embodies the idea that explicitness is hierarchical and that not all knowledge is either fully explicit or not explicit at all. It is generally recognised that one reason why some knowledge is unavailable for report is that it has been automatised. Schneider and Schrifin (1977) sum up the definitional difference between automatised and controlled processes remarkably concisely:

Automatic processing is activation of a learned sequence of elements in long-term memory that is initiated by appropriate inputs and then proceeds automatically – without subject control, without stressing the capacity limitations of the system, and without necessarily demanding attention.

Controlled processing is a temporary activation of a sequence of elements that can be set up quickly and easily but requires attention, is capacity-limited (usually serial in nature), and is controlled by the subject" (p.1).

DeKeyser (1996) sees automaticity as a continuum rather than an automatic-controlled dichotomy, and his research on grammar rule-learning (1997) confirmed that practice led to gradual automatisisation (p.207). However, with a reference to LaBerge & Samuels (1974), and Anderson (1992) he emphasises that besides consistent practice and spread of practice and feedback are also key issues in automatisisation of L2 skills (DeKeyser, 2001:145). The more automatic a performance is, the fewer cognitive recourses it will demand, thus freeing these recourses to be applied in controlling higher-order processes. The resulting effects should be measurable in shorter response times and a decrease in error rate.

Robinson and Ha (1993) sum up memory-based theories of automaticity thusly: "Performance is automatic when it is based on single-step, direct access retrieval of past solutions from memory rather than some algorithmic computation. This view provides several mechanisms by which automaticity can be acquired" (pp. 414-415). Robinson and Ha list three different memory-based mechanisms that can lead to automaticity: strengthening of connections, pattern recognition, and instance learning. Testing Logan's (1988, 1990, 1992) instance theory of automaticity, which was based on learning alphabet arithmetic problems, Robinson (1997b) found that the experimental groups, which comprised four different conditions, responded similarly in all groups to old stimuli in the transfer grammaticality test with regard to speed and automaticity (p.241). The experiments demonstrated no significant differences among the two conditions rule-based (focus on form) versus memory-based (no focus on form) conditions with regard to reaction times on trained items, but did show a difference in learning as tested in the grammaticality judgement tests.

Segalowitz (2003) suggests that automaticity may not be a unitary construct in the sense that it may be achieved through different routes. He points out the contrast between Anderson's (1983) theory and Logan's (1988) theory. Anderson sees automaticity arising out of rule-governed behaviour which becomes proceduralised. Logan sees automaticity arising from rule-governed behaviour which results in instances of

solutions being stored in memory. Segalowitz concludes that at the present the issue is unresolved. However, the quality of performance involving automatic behavioural behaviour is fast, stable, and demands fewer attentional recourses than non-automatic skill performance.

5.5 Summary

Discussion on implicitness and explicitness and the role of the two constructs in learning and language learning show disagreements on essential points such as the role of consciousness definitions and testing. On the one hand, there is the contention that learning entirely without consciousness is not possible, and that even what is categorised as implicit knowledge derives from conscious awareness at some point in the process. On the other hand, it is clear that some processes involved in learning and language learning are not accessible to conscious awareness in the sense that they cannot be verbalised. This is not disputed, the contention concerns the nature of these processes: are they self-organising processes operating on consciously perceived input or do they constitute implicit learning processes? Empirical studies have not rendered unambiguous results. The differing views on the nature, function and accessibility of knowledge also find expression in the research on automaticity. Implicit knowledge is by some seen as explicit knowledge which has become internalised, i.e. automatised through practice. Others support the view that implicit knowledge consists of unconsciously extracting regularities from the environment, i.e., in the case of second language learning, from the input. With the review of second language learning in Chapter 2 in mind, it is important not to lose sight of the totality of the many-faceted process in which other factors have a role to play and in which the implicit-explicit dichotomy is one issue among several others.

6 INTROSPECTION

6.1 Verbal report as a research tool

Introspection as a method to study cognitive processes has been used since the beginning of the 20th century⁹⁸ and was further developed after the seminal work by Ericsson and Simon (1984, 1993), which established the method as a genuine way to get a glimpse into the workings of the human mind. The method fell out of favour during the period of behaviourist theory dominance⁹⁹, but is now recognised as a method of value in the pursuit of new knowledge relating to, for instance, learner strategies, cognitive strategies, human-computer interaction, functionality of software and hardware, and not least as a test or assessment of the quality of the knowledge held (for the latter see further discussion below). The development of the method, which gained momentum in the eighties and nineties, was given expression in second language learning by Færch and Kasper (1987a,b) and has since then been adopted in a number of studies in the field (i.a. Haastrup, 1989; R. Ellis, 1991; Smagorinsky, 1994; Cohen, 1996). This does not mean that the criticism has disappeared entirely, and it is a method that must be applied and evaluated after careful consideration of its appropriateness in relation to the study objectives, context and saliency.

Nisbett and Wilson (1977:232) use the term "anti-introspectivist" to refer to cognitive psychologists criticising the use of introspective methods, and they quote Neisser, 1967:301, Miller, 1962:56, and Mandler, 1975:230, 241, 245, for the position that "we may have no direct access to higher order mental processes such as those involved in evaluation, judgement, problem solving, and the initiation of behavior." (Nisbett and Wilson, 1977:232). It is pointed out, though, that the three quoted authors did not support their position with data. Furthermore, the importance of the context is pointed out. Nisbett and Wilson argue that what might appear to be a refutation of the method arises from criticism of the application and the lack of saliency in the causal mechanism

⁹⁸ Frensch et al. refer to Wundt (1896) as the beginning with regard to reports on environmental regularities experienced by individuals.

⁹⁹ Ericsson and Simon (1980:215)

of stimuli. "Accurate reports will occur when influential stimuli are salient and plausible causes of the responses they produce, and will not occur when stimuli are not salient or are not plausible causes." (Nisbett and Wilson, 1977:231).

The theme under discussion above is what could be classified as stimulated introspection, and the Nisbett and Wilson article makes it clear how important a role classification and categorisation play for appropriate application and evaluation of methods and results. The stimulated introspection method has its problems and limitations as do all methods of introspection, but they should not be criticised for not producing reliable results in situations for which they were never intended. In addition Nisbett and Wilson outline the importance of not confusing content with process (ibid. pp. 255-256) and observe, "we are often capable of describing intermediate results of a series of mental operations in such a way as to promote the feeling that we are describing the operations themselves" (ibid. p.255), and they go on to give the following example:

... the confusion of intermediate output with process was provided by an acquaintance of the authors' who was asked to introspect about the process by which he had retrieved from memory his mother's maiden name. "I know just what the process was," he said. "I first thought of my uncle's last name, and since that happens to be my mother's maiden name, I had the solution." This only pushes the process question back a step further, of course, and our acquaintance's answer would appear to reflect a confusion of intermediate results with the process by which the final result was obtained (Nisbett and Wilson, 1977:255-56).

This type of content analysis can provide valuable information for researchers, if that is what they are looking for, and in the access to this type of content individual subjects have privileged access of a quality which any external observer would undoubtedly find hard to match. According to Nisbett and Wilson there are three kinds of content access at which an individual subject's accuracy would be superior (ibid:256), the third of which is of interest to the present study and concerns "attentional and intentional knowledge". However, any type of self-report should be used with great care, and for this particular reason, among other considerations, stimulated recall was not deemed appropriate for the present study.

A state-of-the-art work on verbal report as a source of data on cognitive processes which gave the field new insights and confidence in the method, and which today is still considered the most reliable reference was authored by K. Anders Ericsson and Herbert A. Simon in 1984 and revised in 1993.

6.1.1 Levels of verbalisation

Ericsson and Simon define three levels of verbalisation (Ericsson and Simon, 1993:79):

1. Vocalisation of covert articulatory or oral encodings.

This level requires no special effort, and the time frame of thought processes will be the same with or without verbalisation.

2. Description or explication of thought content.

Information (e.g. an odour) that is not originally orally encoded (a recoding process added).

The added processing load means that a subject who is verbalising will need additional time for the task whereas the structure of the processes remains the same.

3. An interpretative process which is not simply recoding and may thus change the structure of the thought process, and the actual motives and causes may not be represented correctly.

This definitional categorisation makes it possible to evaluate the Nisbett and Wilson¹⁰⁰ discussion above as it becomes plausible that they were discussing what Ericsson and Simon designate as level three and of which they say that it may involve a subject in making inferences which cannot be concurrent reports of the direct processes of the actual original task performance.

¹⁰⁰ Nisbett and Wilson were criticised by Peter White (1980:105) for "unwarranted assumptions about the relationship between conscious awareness and the process and the verbal report". White is not nearly as critical about the method as Nisbett and Wilson are, but he is critical of their methodology and thus their conclusions. Ericsson and Simon (1980), after close scrutiny of the studies in question, agree with much of the criticism raised by Nisbett and Wilson (1977) and conclude: "In fact, Nisbett and Wilson's own detailed summary of the conditions under which verbal reports can be assumed to be valid are consistent with the conclusions we have reached in this article" (Ericsson and Simon, 1980:247).

6.1.2 Character of reports

To be available for reporting, information must be held in Short Term Memory (STM) or be retrievable from Long Term Memory (LTM) for processing in STM. In their 1980 article, Ericsson and Simon (Ericsson and Simon, 1980:236) draws attention to a number of situations in which verbal report of processed information is either not possible or will result in incomplete reports¹⁰¹. Unavailability is generally the case for processes which do not use STM for storage of the intermediate processes but only for the product such as perceptual-encoding processes, motor processes, and processes which use direct retrieval from LTM. And they add, "In recall and retrieval of familiar information, unless it requires problem-solving with the aid of successive associations, we frequently find processes that leave only the final product as trace in STM" (Ericsson and Simon, 1980:236).

Incompleteness occurs if the information is not available in STM, and this is the case if previously processed information was not stored in LTM for later retrieval, or if the information was not attended to. Furthermore, it must be realised that not all information available is necessarily reported. If a subject is strained by the cognitive demands made by the task required, the verbal reporting may stop or be incomplete. The same mechanisms can be observed in cases where a process is being repeated leading to automatization.

Fully automatized skills or knowledge is not readily available for report. However, automatization is a gradual process during which mechanization of procedures increasingly will reduce the conscious awareness of the information being processed and stored in LTM and thus increasingly unavailable for report. Particularly the intermediate stages of the processing will remain inaccessible as these are no longer activated by cues which trigger retrieval from LTM. Experts who possess knowledge in a given field, and who appear to have ready answers to queries use direct access to information previously stored in LTM, and which was attended to in STM for a sufficient period of time to lead to "fixation" (Ericsson and Simon, 1993:115). It is this

¹⁰¹ See also Ericsson and Simon, 1993, Chapter 3, Completeness of Reports.

fixation which leads to learning, " ... learning does not occur without awareness and ... the information available in LTM is a subset of the information that, at some earlier time or times, was held in STM" (Ericsson and Simon, 1993:116). Learning is the result, or product, of the process which in the advanced stages leads to automatization and information being unavailable, or only partially available, for verbal report as intermediate stages are skipped. The ensuing manifestations of this process result in knowledge appearing to spring 'out of the blue' and leading to subjects being unable to answer probing as to 'how' and 'why'. Therefore the veridicality of reports, especially where verbalisations at level three are concerned, should be a consideration when researchers decide on designs which include either concurrent self-report or retrospective reports.

6.2 Retrospective reports

Level three verbalisations often take the form of retrospective reports. Subjects will be asked to vocalise their motivations for their acts or be asked to recall their thoughts at the time of their acts. This entails explaining, and consequently also retrieval from LTM of past events and intermediate steps in the process of which the result was the action or event which is now the object of study. Gass and Mackey (2000:51) operate with three types of stimulated recall: Consecutive Recall, Delayed Recall, and Nonrecent Recall. Some prompting is usually embedded in the method. The prompting may take various forms, e.g. interviews which seek answers to items of the students' thoughts during an event, or which elicit information from the students about their reactions during an event, or their strategies when carrying out acts.

A variation of the retrospective report is the procedure where the students are prompted by video recordings or audio recordings of themselves carrying out tasks and encouraged to recall what their thoughts were during the event. Characteristic of this probing are queries about the students' memories of how, why, when related issues, which, as discussed above, may lead the student to speculate in case the desired information is not available or incompletely represented in the student's mind. This does not mean that this method or this kind of information is worthless, only that there are limitations to the usefulness of verbal reports as data. One issue has to do with the

accuracy of the reporting. This is particularly the case in self-report and self-observational data. A second issue has to do with the type of memory structure used in recalls. With self-report and self-observational data, when the time between the event reported and the reporting itself is short, there is a greater likelihood that the reporting will be accurate" (Gass and Mackey, 2000:17).

The problems involved notwithstanding, the stimulated retrospective reports can in fact bring forth valuable information which would otherwise be inaccessible and hard to achieve in any other way. Several important studies have implemented this method (see Gass and Mackey, 2000:29-35 for an overview) to elicit many different types of information in a varied field of study, such as Færch and Kasper's study on translation (1986a), Kirsten Haastrup's study on vocabulary (1987, 1989), and Susan M. Gass's 1994 study on acceptability judgements, to mention a few.

6.3 Concurrent reports

Verbalisations which are vocal encodings of thoughts on choices, strategies and even attitudes need not be retrospective although that is often the case. Some researchers prefer concurrent reports because the temporal distance between event and report is a factor for accuracy and validity of the report. Concurrent reports are either self-reports of content as outlined above involving some sort of explanation on the part of the subject¹⁰², or think-aloud vocalisation of thoughts as they come into a student's mind during an event. The former type of report would be classed as a level three verbalisation according to Ericsson and Simon's typology, and the latter type would be level one verbalisation.

Andrew D. Cohen (1996) classifies this type of verbalisation as "self-revelation" and defines it as "stream of consciousness disclosure of thought processes while the information is being attended to" (p.7). In their preface to the 1993 edition of their book,

¹⁰² Ericsson and Simon classify these as retrospective. Presumably because they involve two sets of processes; one which is the process connected to the performance of the task, and one which is a retrieval of information heeded during the task and recoded into a comment, explanation, or other aspects wanted by the researcher.

Ericsson and Simon characterise the qualitative advantages of verbalisation concurrent with the thought processes as follows:

When subjects verbalize directly only the thoughts entering their attention as part of performing the task, the sequence of thoughts is not changed by the added instruction to think aloud. However, if subjects are also instructed to describe or explain their thoughts, additional thoughts and information have to be accessed to produce these auxiliary descriptions and explanations. As a result, the sequence of thoughts is changed, because the subjects must attend to information not normally needed to perform the task (1993, preface:xiii).

Some critics object to the application of level 1 and 2 verbalisations because the protocols often are unorganised with little coherence, few complete and well-formed sentences and sudden shifts in focus. Although verbalisation is a serial process, the protocols of level 1 and level 2 types do not give a linear and logical string of complete sentences in contrast to level 3 verbalisations which will allow the subject to encode into what Ericsson and Simon call "social verbalizations" (1993:preface, xiv), i.e. socially acceptable communication forms. From an information processing point of view, however, concurrent think-aloud protocols offer the most direct and unedited access to the cognitive processes because the "processing and verbal report would coincide in time" (ibid, xiii). Concurrent think-aloud protocols are of value for the study of cognitive activities involved in problem-solving processes as well as for intermediate steps or stages in these processes, and they provide information which is not obtainable from observational behaviour.

There are limitations to the capacity of the short term memory which¹⁰³ holds the information needed for only a brief period of time (i.e. seconds), and it can contain only a limited number of chunks of information at any one time. Information held in STM, and only that, is available for further processing and verbal reporting. Some of this information is new input from the sensory system and some may be previously held information which was stored in LTM and which is being retrieved for new processing

¹⁰³ In cognitive science short term memory is now often referred to as working memory, a term usually ascribed to Baddeley (1995; first published 1986).

in coordination with new information merging into new revised and updated information that may then result in fixation in LTM (see Figure 6.1)¹⁰⁴.

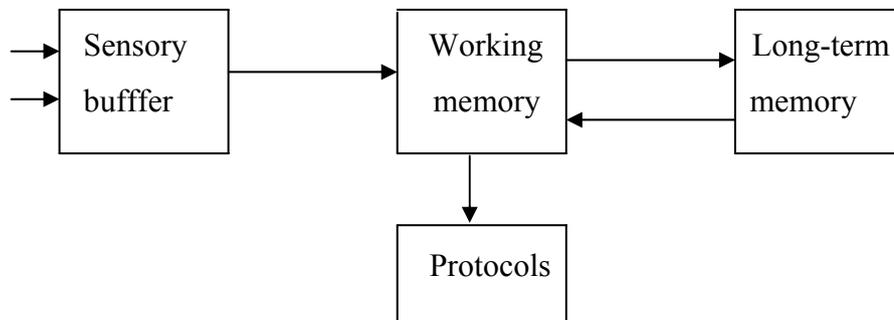


Figure 6.1 Memory model.

Source: van Someren et al., 1994:19

Critics of the method were worried that the task of verbalisation would influence the mental processes which the investigators were interested in studying. This is a worry that is not supported by Ericsson and Simon who posit¹⁰⁵ the following:

Two forms of verbal reports can claim to being the closest reflection of the cognitive processes. Foremost are concurrent verbal reports – "talk aloud" and "think aloud" reports – where the cognitive processes, described as successive states of heeded information, are verbalized directly.... We claim that cognitive processes are not modified by these verbal reports, and that task-directed cognitive processes determine what information is heeded and verbalized (1993:16).

The instructions given to subjects before the procedure must be carefully designed for the nature and purpose of the protocols. Inappropriate instructions may influence the protocols in an unintentional way with regard to the content. Interference during the process may change the course of the cognitive processes. Such interference may, for instance, consist of probes by the experimenter in the form of queries of a nature which

¹⁰⁴ See also Ericsson and Simon, 1993:17.

¹⁰⁵ A position which remains the same after looking closely at empirical evidence.

directs or alters the attention of the subject and consequently the content of the report. "Probes, if used, must be carefully worded to reduce the likelihood of intermediate and inferential processing, both of which could change the natural sequence of heeded information" (Green, 1998:9-10). When talking aloud or thinking aloud, subjects will invariably, especially in tasks involving problem-solving, pause at times, and occasionally they will need to be prompted by the researcher. Such prompts need to be as brief as possible and as little open to interpretation by the subject as possible. "When the subject is working on the task, the role of the experimenter is a restrained one. Interference should only occur when the subject stops talking. The experimenter should prompt the subject by just, and only just saying: "Keep talking". (van Someren et al, 1994:44). Ericsson and Simon mention several other feasible options, e.g. 'Please think aloud', "Please tell me what you are thinking", but they recommend "Keep talking" because it cannot be interpreted by the subjects as needing a direct response to the experimenter (1993:256).

7 THE QUALITATIVE STUDY

7.1 The context

7.1.1 General

The quantitative study (see Chapter 4) comprising 107 subjects was designed to bring forth data on the efficacy of the various experimental methods in the experiment and the differential effect of the treatments on student groups and discrete items¹⁰⁶. To complement the quantitative study it was decided to conduct a qualitative investigation whose purpose it was to focus on the cognitive processes¹⁰⁷ involved in the tasks connected with sentence analysis, and thus illuminate and support the quantitative results. This qualitative investigation took place the year following the quantitative study.

It was an important consideration for the validity of the qualitative study that the set-up should include a replication of the preceding year's quantitative elements, thus enabling the creation of a similar environment for the qualitative study from which could be drawn reliable parallels that could lend authority to the findings that might emerge.

7.1.2 The design background

In parallel to the quantitative study, subjects were included from the English study programme as well as the Cand. Negot. study programme and each programme was represented by one class. In each class half the students were assigned to a VISL group and the other half to a NON-VISL group. The total student population was considerably smaller than the preceding year. Consequently, the classes also comprised fewer students than the classes the preceding year, and in addition the dropout rate turned out to be somewhat higher was the case the preceding year.

From the outset the experimental groups comprised 19 subjects in the VISL group and 17 subjects in the NONVISL group. By the time the experimental period expired the

¹⁰⁶ After having commenced this section of the research it became clear that the thesis would benefit from other types of research which would enlighten the topic of research in different ways.

¹⁰⁷ In particular, I am grateful to Professor Nick Ellis who was kind enough to discuss my project with me.

number of students had been reduced to 13 subjects in the VISL groups and 8 subjects in the NON-VISL group. Of these, 8 subjects were chosen from each group to participate in the think-aloud (TA) protocol section of the study in which subjects were recorded while concurrently analysing sentences and thinking aloud. The high dropout rate among the NON-VISL students especially meant a limited option for selection of subjects to the experimental TA groups.

7.1.3 The preceding treatment

A pre-test was administered on the first day of university followed by a ten-week treatment period which was concluded by a post-test. The NON-VISL treatment consisted of traditional chalk-and-talk classes in sentence analysis, and the VISL treatment shared the same content but utilised the Visual Interactive Syntax Learning computer programme in the experimental classes. In comparison to the preceding year's large quantitative study there were no differences in hours, content or methodology; nor were there any differences in the testing material or methodology (see Chapter 4).

7.2 Methodology

7.2.1 Collecting the protocols

7.2.1.1 Rationale for the choice of type of protocol

The concurrent think-aloud protocol has the advantage over other types of retrospective methods, such as stimulated recall introspection, in that the collecting of information does not interfere with the thought processes if carried out properly. There may be limitations as to the content and scope of the protocols as discussed above in Chapter 6, but most importantly, the concurrent think-aloud method allows subjects to voice their thought as these thoughts appear in their heads with no interference from any medium or time lapse which might distort the outcome.

7.2.1.2 Instruction

The subjects were instructed prior to the recording sessions. During this instruction they were given examples of the different forms of protocols, i.e. between thinking out loud and retrospective reports. They were told that they should think out loud, i.e. vocalise any thought that came into their head while carrying out the tasks, and they were

specifically told not to speculate about or explain their performance. Subsequently, they were given the opportunity to carry out the procedure. In addition, at the recording sessions each individual was given written instructions which they were asked to read before embarking on the task. These instructions were written on the same sheet of paper as the task sentences. The task sentences were unknown to the subjects before the sessions and were presented to them at the recording sessions. The instruction to the subjects made it clear that the subjects were the decision-makers during the recording sessions. It was an integral part of the process that the subjects had to decide whether or when a task was finished, and/or whether to continue, how to continue, or whether to give up. Equally, the subjects were told that the language in which they were thinking, and therefore verbalising, was entirely a choice left to the participants themselves; it could be English, Danish, or a mixture. The quotes and TA excerpts are given in the language in which they were spoken. Remarks made in Danish have been translated, and the translated utterances added in brackets.

7.2.1.3 The recordings

Both the VISL and the NON-VISL subject were recorded by camcorder and the recordings depict their actions as well as containing their verbal records. For the VISL subjects this means that the computer screen was recorded to document their actions. For the NON-VISL subjects the camera was focused on the paper on which subjects wrote down their analyses as they built the tree diagrams. The camera was placed to the right/back beyond the visual range of the subjects. A cameraman from the IT service department conducted the technical tasks involved in the recordings and was present at all recordings but positioned behind the subjects and outside their field of vision. The recordings took place after two weeks of instruction and again after ten weeks of instruction. At each stage, recordings were placed on two separate days, one for English subjects, and one for Cand. Negot. subjects.

7.2.1.4 The subjects

The subjects for the think-aloud study were selected on the basis of their pre-test results (see Table 7.1). The aim was to have a broad selection of different levels of ability and attainment from each study programme as well as for the two treatment groups. Four

students representing pre-test scores ranging from low to high were selected from each of the following groups: English VISL, English NON-VISL, Cand. Negot. VISL, Cand. Negot. NON-VISL. These sixteen TA subjects were not given special instruction but followed the experimental classes for the groups they were assigned to originally.

Table 7.1
Entrance levels of TA subjects.

Subjects		Pre total (max. 25)	Pre Syll. (max. 16)	Pre Non-syll. (max. 9)
VISL				
English				
	TAS1	6.55	1.55	5
	TAS2	10.62	5.62	5
	TAS3	19.20	11.20	8
	TAS4	16.82	9.82	7
Cand. Negot.				
	TAS5	10.32	4.32	6
	TAS6	11.33	5.33	6
	TAS7	8.29	4.30	4
	TAS8	13.15	9.15	4
NON-VISL				
English				
	TAS9	15.13	8.13	7
	TAS10	12.40	7.40	5
	TAS11	3.00	0.00	3
	TAS12	11.91	7.91	4
Cand. Negot.				
	TAS13	7.79	3.79	4
	TAS14	9.05	3.05	6
	TAS15	13.32	7.32	6
	TAS16	5.55	1.55	4

7.2.1.5 The subject-matter content

The content of the recordings are the thoughts verbalised by the subjects while analysing simple sentences containing the most basic sentence functions and forms. Each subject analysed two sentences in the early (hereafter PRE) recording sessions, and each subject two sentences in the later (hereafter POST) recording sessions. The PRE sentences are: *George painted the house yesterday*, and *She is a very beautiful woman*. The POST sentences are: *Jill sold the car today*, and *It is a very good book*. Each analysed sentence constitutes one protocol, which means the corpus contains 22 NON-VISL protocols and 24 VISL protocols. In addition the VISL subjects were asked to analyse one sentence each on paper after having analysed the two POST sentences.

This sentence is: *She visited the hospital yesterday*. This increases the corpus with six protocols.

The sentences were deliberately kept at a simple level while simultaneously containing all basic syntactic elements. This policy was adopted for two reasons. The overall consideration of the purpose of this experiment was to gain insight into the cognitive processes underlying the product measured by the tests which were the foundation of the qualitative study. The insight into these processes was more important than the actual products in the form of analysed sentences. The attainment of the knowledge of the cognitive processes as elicited from TA protocols was more likely to be furthered by simple sentences which would give rise to unhindered verbalisation in contrast to more complex analysis tasks which might create overly heavy cognitive loads (see also Guan, Lee, Cuddihy & Ramey, 2006), and consequently lead to less verbalisation, given the premise that the tasks included the fundamental structure of English syntax¹⁰⁸. Ericsson and Simon (1980:237) state: “Subjects tend to stop verbalizing or verbalize incompletely in conditions in which they are giving indications of being under heavy cognitive load”. A pragmatic consideration was the need for tasks which could be carried out by the novices as well as "experts" since the first recording sessions were held at a time when barely any training in the subject matter had taken place.

7.2.1.6 Task analysis

Sentence 1 PRE and sentence 1 POST contain the same number of functions (7) and forms (7), and these functions and forms are identical from pre to post, the only difference being their semantic expression. Sentence 2 PRE and sentence 2 POST adhere to the same pattern as mentioned above with the same number of functions (9) and forms (9). Syntactically type 1 sentences and type 2 sentences differ with regard to the type of verbs contained in predicator and consequently with regard to sentence

¹⁰⁸ Waters & Caplan (2004:132), describing experiments on syntactic processing, though basically for comprehension and not verbalisation, state that “...low-span college students and elderly individuals with limitations in working-memory capacity do not have increased difficulty in processing syntactically complex sentences, such as garden path sentences and sentences with object relativization. [...] Subjects with reduced working-memory capacities are also not differentially impaired on syntactically more complex sentences under concurrent memoryload conditions”. Although this is not related to think-aloud experiments they express underlying conditions of syntactic processing which pre-conditions thinking aloud (see also Caplan & Waters, 1999) .

constituents, in that type 1 sentences have a transitive predicator and thus an object. Type 2 sentences have a copula predicator and thus a complement. Both type 1 and type 2 sentences contain the functions H and DEP at group level (see Table 7.2). Type 1 sentences contain two levels: clause level and group level. Type 2 sentences contain three levels: clause level and two group levels. Type 1 sentences contain a DOER and a DONE-TO plus an EXTRA (adverbial) whereas type 2 sentences contain a SPECIFIED and a SPECIFIER. All sentences are active declarative sentences (Bache and Davidsen-Nielsen 1997:196-197).

Given the way the VISL interface supplies support for structure (tree) building, it could be expected that VISL subjects would take advantage of this facility and complete the tree diagram structure before embarking upon the task of labelling constituents and parts of speech. In contrast this could not be expected to the same degree of the NON-VISL subjects, in fact expectations were that they would complete one level at a time.

Table 7.2
Task analysis for the think-aloud tasks.

```

+-----+
PRE-Sentence 1: George painted the house yesterday
-----
FUNCTIONS:      7                                FORMS:      7
      S      1                                noun      2
      P      1                                verb      1
      Od     1                                adverb    1
      A      1                                article   1
      DEP    1                                group     1
      H      1                                clause    1
      STA    1
-----
+-----+

PRE-Sentence 2: She is a very beautiful woman
-----
FUNCTIONS:      9                                FORMS:      9
      S      1                                noun      1
      P      1                                verb      1
      Cs     1                                pronoun   1
      DEP    3                                adjective  1
      H      2                                adverb    1
      STA    1                                article   1
                                           group     1
                                           clause    1
-----
+-----+

```

POST-Sentence 1: *Jill sold the car today*

```
-----  
FUNCTIONS:  7                FORMS:  7  
    S      1                noun   2  
    P      1                verb   1  
    Od     1                adverb  1  
    A      1                article  1  
    DEP    1                group   1  
    H      1                clause  1  
    STA    1  
-----+
```

POST-Sentence 2: *It is a very good book*

```
-----  
FUNCTIONS:  9                FORMS:  9  
    S      1                noun   1  
    P      1                verb   1  
    Cs     1                pronoun  1  
    DEP    3                adjective 1  
    H      2                adverb  1  
    STA    1                article  1  
                group   2  
                clause  1  
-----+
```

POST-Sentence VISL on paper: *She visited the hospital yesterday*

```
-----  
FUNCTIONS:  7                FORMS:  7  
    S      1                noun   1  
    P      1                pronoun  1  
    Od     1                verb   1  
    A      1                adjective 1  
    DEP    1                adverb  1  
    H      1                article  1  
    STA    1                group   1  
                clause  1  
-----+
```

7.2.1.7 The role of the experimenter

The experimenter was present at all recording sessions with one function in view, namely to secure that participants continually verbalised thoughts as they came into their heads, i.e. what Neil J. Anderson described as “thoughts running through their minds” (N. Anderson, 1996:7). In practise this meant that whenever subjects stopped talking, they were prompted by the experimenter and asked to ‘keep talking’ or similar words to that effect.

In accordance with Ericsson and Simon’s guidelines subjects were never asked questions only encouraged to continue the verbalised flow of thoughts. Furthermore, the experimenter was positioned behind the participants outside their range of view. Research has shown that in any student-teacher relationship, students consciously or unconsciously ‘read’ or react to what they perceive to be in accordance with the desires

of the teacher. "Through subtle cues – such as tone of voice, posture, smiling approvingly, and so on – a researcher can cue desired behaviors" (Smagorinsky, 1994:6, see also Smagorinsky, 1987:337-339). Having a visual impression of the experimenter might therefore influence not only the actions of the participants but their very thought processes. For the same reasons, the experimenter took care to be silent except when it was necessary to prompt subjects.

7.2.1.8 Protocols

A protocol comprises one analysed sentence. Each protocol has been transcribed from the recorded videotape and subsequently segmented in such a way that each segment contained one idea or one event. An event is either a verbalised cognitive process or an action designation. Pauses may also qualify as events. Thus a segment may be verbal as well as non-verbal. The verbal segments may vary from one word, e.g. '*Ja*', to a whole clause, e.g. '*It is hard to speak and think at the same time*'. In line with the variation in segment length, the protocols vary greatly in length as well as in number of segments. The shortest protocol contains 17 segments and the longest as many as 320 segments.

The long protocols are found in the VISL as well as the NON-VISL sections but PRE protocols tend to be longer than POST protocols (see below for interpretation of protocols). Short is a term that may refer to number of segments as well as duration in time of a protocol. For each protocol the time spent on the analysis task was recorded and time is a parameter used in parallel with other significant signs in the interpretation of the protocols.

7.2.2 Content of the protocols

7.2.2.1 Elicitation of information

Think-aloud protocols contain a fountain of information which is many-faceted, heterogeneous, dispersed, unorganised and less than lucid. Sentences are incomplete, trains of thought are broken, some never to be continued whereas other threads appear later in a new context. This 'raw' material needs to be organised, categorised, and interpreted. In the process inferences need to be made based on the context and content of the protocols. This is the responsibility of the researcher rather than the subjects. In

other types of introspection subjects play a role in that process but this is not the case in concurrent think-alouds:

The term that continues to surface in discussions with researchers who study verbal reports is "messy". The work is very messy indeed, and no doubt the difficulty of replication, the idiosyncratic nature of each investigation, the need to conceive hypotheses and data analysis procedures in mid-study, and other seeming indicators of imprecision have caused great consternation among those who find the methodologies "unscientific" (Smagorinsky, 1994:x).

Ericsson and Simon (1987, 1993) examined arguments of this nature and thoroughly examined empirical studies and their methodologies and reached a conclusion also voiced by Smagorinsky: " I argue that protocol analysis nonetheless offers a unique glimpse into the workings of the human mind..." (Smagorinsky, 1994:xiii).

Empirical research by Leow and Morgan-Short (2004) investigated the effect of concurrent think-aloud activities on the performance of participating subjects. They were especially interested in the effect on the processes involving attention and awareness and whether the TA procedure would result in reactivity and thus influence or invalidate results concerning the constructs under investigation. The conclusion from a reading task was that the think-aloud procedure neither facilitated nor detracted from performance when compared to the control group. Carefulness and timidity on the part of the researcher is called for, however, in order not to over-interpret results.

7.2.2.2 Categories of information

The contents of the present think-aloud study fall into three major categories:

1. Thoughts on the software including those which can lead to conclusions pertaining to human-computer interactions.
2. The immediate information available on the task of sentence analysis itself as a discipline and the mastery of it by the students. This information pertains to the product of the task rather than the process although the two sides are closely intertwined and may serve to inform each other. In particular, this information may help with the interpretation of some of the results of the quantitative study which

again may give rise to deliberations on the pedagogical implications of the results of the study. This type of information can be described as subject-matter content information which is extractable with little or no inferences on the part of the researcher. An evaluation of the nature or the quality of the knowledge whose presence is evidenced in the tangible content of the protocol, on the other hand, requires a theoretical underpinning which consequently entails a sophisticated coding system in harmony with the theoretical hypotheses. Needless to say, the development of such a coding system will demand a measure of inference and interpretation on the part of the researcher. The information extracted in this category might possibly enhance our understanding of the learning process and the acquisition of a particular kind of knowledge.

3. Process information, in casu information about the learning strategies employed by the subjects. This information requires some measure of inferences on the part of the experimenter and a classification, i.e. coding, system needs to be employed. Subsequently, an analysis is needed to describe first the strategies employed, their relative usefulness and finally the pedagogical implications for the future. Learning styles and learner types are related to the choice of strategy and will be considered when appropriate but will not be a subject of analysis in its own right.

These points will be elaborated on, explored and discussed below.

7.3 Interpreting the protocols

The interpretation of the content of the protocols is a long and composite process which is a combination of pre-conceived ideas of the purpose of the think-aloud, and the expected resulting content of the protocols. The development of a systematic instrument builds on the repeated reading of the protocols, which in itself entails looking for and discovering expected and unexpected manifestations of your target elements, revision of theories, and developing new theories in response to the actual material at hand.

The think-aloud method has been shown to give increased insight in studies involving reading, writing/composition, vocabulary, social interaction and other areas of study in language acquisition and processing, so model studies and coding schemes already exist

(see section above on Introspection; see also Bracewell & Breuleux, 1994; Gass & Mackey, 2000) which to some extent can meet the needs and difficulties in protocol analysis in general. However, the aim of the present study is to elicit from the TA protocols information as far as it is discernible with regard to the degree of explicitness of the knowledge held by the subjects. In order to achieve this goal it is attempted to deploy a coding system, or rather a system of interpretation of verbalised utterances, which finds its basis in a theoretical framework that can be operationalised in a way which will lend systematicity to the application. This coding system, or system of interpretation, has its theoretical underpinning in Dienes and Perner's (1999) theory of implicit and explicit knowledge (see 7.3.1).

The following protocol excerpt and its interpretation in terms of a brief account of the segments and their classification illustrate the process from protocol to classification and evaluation:

TA Excerpt 7.1

TAS4 (VISL). POST-Sentence 2.

+-----+	
POST-Sentence 2: <i>It is a very good book</i>	
+-----+	
Segment	Transcription
+-----+	
01	ahr..this bothers me that this machine is so..... slow,
02	hm.. (the blue/green boxes did not appear but half-formed words which he has to click on to make the boxes appear) <i>hm...</i>
03	then I have the predicator
04	which is this (points at <i>is</i> but does not click)
05	and subject which is <i>It</i> (does not click)
06	and then I have what it is, and it is <i>a very good book</i>
07	so I'm going to try to put these four words into <i>a</i> group (clicks for the first time)
08	and I'm correct
+-----+	

This short excerpt from the beginning of the protocol is, despite its briefness, very rich and composite. There is information on the software handling and software attitude. There is an affective element (student attitude). There is metalinguistic content (knowledge), information on the nature/quality of that knowledge, and the application of that knowledge (cognitive strategy), and finally an instance of self-monitoring. The

handling of the software is flawless, which the software itself is not. Segment 01 contains two negative remarks, one which can be classified as a negative remark on the software: "*it is so...slow*", and one which reveals that the subject is annoyed by it. This annoyance is expressed in two ways, first by the subject's words: "*this bothers me*" and by the pause between the intensifier '*so*' and the descriptive adjective '*slow*'. The subject virtually acts out his annoyance. The latter interpretation constitutes an inference on part of the researcher whereas the former is documented in verbal form. This inference is legitimised by the language which also justifies an interpretation of segment 02 as an expression or continuation of this grumbling attitude towards something about which the subject is clearly sceptical. Segment 02, "*hm*", pause, "*hm*"¹⁰⁹, points towards a reservation or scepticism of the workings of the machine as if the subject were thinking "what's it up to now - and when will it get on with it?"¹¹⁰. Thus the coding of segments 01 and 02 contain two instances of "negative software remark" and one instance of 'negative student attitude'.¹¹¹

The subject-matter content is present in segments 03 '*predicator*', segment 05 '*subject*', segment 06 '*what it is*', and segment 07 '*a group*'. The complete structure of the sentence at the sentence level (the group level structure of '*very good*' as a dependent group comes later in the protocol, is there, including the terms of the sentence functions, with the exception of 'subject complement', which comes later in the protocol (see Appendix X). The subject demonstrates a knowledge of the function although the term is saved for later, but the student has taken the first step towards terms for the forms of the sentence constituents in that '*a group*' is mentioned. This is a necessary prerequisite for the completion of the whole sentence structure that follows later in the protocol since the analysis of the next level depends on the knowledge that the sentence level structure contains a group. The subject-matter content analysis leads us towards an

¹⁰⁹ Bracewell & Breuleux, 1994, transcribed verbal pause utterance as "um". I chose to transcribe it "hm" because that represented my perception of the sound made by the actual subjects. Having read Bracewell and Breuleux's transcriptions, I considered changing my transcription. However, after having given it some thought I decided to retain my original transcription for the reason given above. My suggestion is that there may actually be a qualitative difference in the phonic manifestation of native speakers' pause or hesitation marker compared to that of non-native speakers.

¹¹⁰ Bracewell and Breuleux's study (1994) of the writing process contains an excellent and very exhaustive description of the process from transcription to coding and interpretation. They describe "um" as "predominant *typing resumes* sentiment", (Bracewell & Breuleux, 1994:61). This could be interpreted and extrapolated to mean 'activity resumes', and this may indeed be the case; however, I have taken a more differentiated view on this which is described in the section on pauses.

¹¹¹ Learner strategies were not part of the present study.

answer as to why the subject only points to predicator and subject but does not click until he comes to '*a very good book*'. The subject needs to combine the words in order to complete the sentence structure.

The protocol gives us insight into the cognitive and metacognitive strategies employed by the students and may allow us to draw conclusions with regard to learning strategies and their implications. The strategy employed by this particular student would be coded as a metacognitive strategy in the subset of planning. In this particular protocol there are linguistic 'markers' (i.e. verbal encodings) whose presence would indicate the metacognitive strategy of planning, such as 'then' in segment 03, 'and' in segment 05, 'and then' , 'and' in segment 06, and 'so I'm going to try to put' in segment 07. The excerpt contains one further instance of metacognitive strategy in segment 08 'I'm correct' which belongs to the subset of evaluation, but it might be difficult to say whether it belongs to self-evaluation or self-reinforcement as further subsets, or perhaps both. So far the extraction, coding and interpretation of the information contained in this short excerpt has been possible with very little inference being necessary since the protocol documents virtually all the elements explicitly (i.e. in verbally encoded thoughts).

The category of information that pertains to the quality of the knowledge held by the subject is also contained in TA Excerpt 7.1, and this demands a more elaborate coding scheme and a higher degree of inference which, to be reliable and of value, must be underpinned by cognitive theory. Such a theory emerges from the work by Dienes and Perner (1999, 2002). The theory and its inherent potential for operationalisation of the constructs will be outlined below followed by its application to the TA protocols. The degree of explicitness of the knowledge of the linguistic content taught during the treatment sessions will be of special interest.

7.3.1 Dienes and Perner's theory of implicit and explicit knowledge

The underlying premises of this theory are the representational theory of mind and the understanding that explicit knowledge is compositional and hierarchical (see Chapter 5 and Table 5.1). The foundation of explicit knowledge is the content that comprises a

proposition which is predicated to an individual. Property is central. A proposition is composed of a property predicated to an individual. To make the property and its predication to an individual into the content of a representation, its factuality must also be explicitly represented. Dienes and Perner give the example of *This is a cat* in order to illustrate the representational content of a proposition. Content consists of property (*cat* or *catness*), predication (*is*), individual (*this*), and factuality, that is, the proposition of *This is a cat* must be seen to be a fact and not just an assumption. If the proposition is not seen to be a fact, the content is not fully explicit. How can factuality be discerned to be held or represented by the system¹¹²? It can be verbally encoded in a verbal report, i.e. *It is a fact that this is a cat*. If the factuality, which is the higher ranging structure in the content box (see Figure 5.1), is present, the theory posits that the lower structures of property, predication and individual are also explicitly represented.

In the analysis of the TA protocols the first step is to find verbal expressions which can be said to hold the factuality of a given proposition (cf. Chapters 7.3.1 and 7.4.1.2). The next step is to establish whether the content of a proposition is represented as knowledge rather than guesswork or wishful thinking, etc., in order to ascertain the explicit nature of the knowledge. If the system represents content as knowledge, the ascribed value is seen as holding truth or as taken for truth at a given time. In the Dienes and Perner model this is labelled ‘*attitude*’, thus the propositional attitude towards the content is classified as knowledge.

Once the content is established as being knowledge, the question is whether this is knowledge held by the *self*. This can be verbally encoded as in *I know it is a fact that this is a cat*. If the self holds the knowledge, it means that the self is cognitively aware of holding the knowledge at the present time.

The fully explicit propositional attitude expressed verbally leaves nothing to be interpreted or inferred. Often, however, not everything about a proposition is verbally encoded in that some aspect is left unsaid, i.e. to be implicitly understood. In Dienes and Perner’s model, functional as well as structural aspects can be left as implicit (see

¹¹² The system is the cognitive system of the individual.

Chapter 5). The implicit aspects are either supporting facts or inherently embedded and necessary for the understanding and truth of the explicitly expressed propositional attitude. The hierarchical nature of Dienes and Perner's model predicts the relationship between implicitness and explicitness. In the example mentioned above, for instance, the temporality of the statement remains implicit. It is understood implicitly that the self holds the truth of the proposition at the current time of expressing it.

For a representation to be possessed by the system as fact it must have been properly caused. In the case of the sentence about the cat this means that a fact has been established through visual perception (p.739). Dienes and Perner point out that the constraints of the system apply only to single representations, and it is certainly possible to combine one representation containing explicitness of fact with another representation that is implicit at a lower level independently of the former. This is illustrated with the example "I know that there is some fact involving F" (p.740). Here *factuality*, *attitude*, and *self* are explicit, but the predication of F to an individual is not implicitly represented as this is not a supporting fact of the explicit content, nor is it inherent in the structure of the explicit meaning of the representation. When the TA subject says "I have the predicator, which is this, and subject, which is *It*" (segments 03-05, TA Excerpt 7.1 in section 7.3), this is interpreted to indicate knowledge, not guesswork or supposition. Especially, the words "I have" and "which is" indicate a degree of explicitness of knowledge which is emphasised by the unhesitating style of presentation and the terseness of expression. In the terminology of Dienes and Perner we can say that there is in the system an explicit representation of predicator and subject. The two properties are predicated to individuals 'this' and '*It*', respectively. The former by pointing to *is* and verbally encoding the physical indication 'this', and the latter by the verbal encoding '*It*'. The propositions are *content* explicit. The *content* is ascribed as being held by the self, '*I*'. The *attitude* expressed about the *content* is 'have'. The interpretation of this word is that it is the visual input which is the cause of the subject's statement. The verbal statement is an indication that the subject 'has' or possesses this knowledge. It is not a supposition or a guess, but explicitly represented knowledge verbalised in the TA protocol. There are no pauses, no hesitation; the verbal report is an expression of confidence in the knowledge. The statement is an example of

fully explicit knowledge of which the *self* represents explicitly *content* and *attitude* (cf. Figure 5.1 and Table 7.3).

7.3.2 Verbal manifestations and content categories

The various verbal encodings of the representations held by the cognitive system of the participating TA subjects need to be examined and classified in the light of Dienes and Perner's theory and their hierarchy of explicit-implicit constraints. The present thesis is especially interested in the degree of explicit knowledge acquired by the TA subjects during the treatment. Research question 4 was: *Do students achieve full explicitness of the subject matter?* Consequently, establishing the presence and subsequent classification and interpretation of verbalisations of expression which can be interpreted to encompass '*self*', '*attitude*', and '*content*', and their respective structures will be the first step towards clarifying aspects of the learning processes of grammatical constructs.

7.3.2.1 Fully explicit knowledge

Fully explicit knowledge is the goal of the treatment involved in this experiment, and the aim of the think-aloud is to find indications of the degree to which the treatment has been facilitative in promoting the success of the participating students. The elicitation of information from the protocols rests on the divisions which are natural to the subject matter. This means that the division between forms and functions will be maintained when discussing the results of the TA manifestations.

Full explicitness entails that *self*, *attitude*, and *content*, and full explicitness are classified only as such if the protocols contain verbal elements of a calibre that justifies such a coding category. If, however, the TA protocols reveal that the subjects verbalise fragmented knowledge, this is registered at the lower levels of the hierarchy; this means at the level of attitude or content. Content is especially crucial for the whole set of dependencies as it contains the factuality of the proposition. If the cognitive system of a subject represents something as a fact, e.g. '*It is fact that this is a subject*', it is implied that it is the self of the subject who is the holder of this knowledge. This is linked to consciousness in that the higher-order thought of being aware of one's own thought becomes possible once the first order thought of establishing something as a fact has

taken place “...once factuality is represented explicitly, explicit representation of self and attitude is also possible” (Dienes and Perner, 1999:741). With a reference to Gordon (1995) they point out that “within one’s own perspective there is a one-to-one correspondence between what is a fact for me [...] and what I know” (p.741). Thus, once ‘*It is a fact that this is a subject*’ is represented, this is a sufficient condition for the self to be consciously aware of holding and giving verbal expression to this as knowledge.

The internal structure of *content* contains *property*, *predication*, *individual*, and *factuality* (see Figure 5.1). If factuality, being the highest in the hierarchy, is explicit, the lower structures are also assumed to be explicitly represented. Furthermore, Dienes and Perner make the case that factuality is a central hub, about which, when expressed, it can be assumed that the knowledge is possessed by the individual: “Because explicit factuality implies explicitness of predication, individuals, and properties, we can conclude that explicit representation of self or attitude implies explicit representation of the content” (Dienes & Perner, 1999:739).

The procedure of the analysis of the TA protocols will be to look for vocabulary indication of factuality first, followed by *predication*, *individual* and *property*. If a protocol indicates full content explicitness, the lower structures are not catalogued. The lower structures become of interest in the search for the degree of explicitness in case full explicitness is not evidenced in the protocol. If there is no full content explicitness, the content and the attitude will be the levels at which the categorisation will take place. An illustration of this is TA Excerpt 7.1, Segment 07, “*so I’m going to try to put these together*”. The knowledge of the elements belonging to the group (i.e. *it is a very good book*) is not fully explicit; there is no factuality, there is no *self*, there is no *attitude*. What can be discerned from the verbal expression is that the *property* of the group is present, and it can be predicated to an individual instance (i.e. the words pointed to), but there is no factuality. The subject “*is going to try*”, which indicates hypothesis testing and not factuality. The illustration exemplifies a case where lower structures are of interest and informative. If the subject performing the task in the excerpt had said ‘I know for a fact that these are a group’, then factuality is taken to be expressed, just as

the *attitude* is explicit, and the *self* is explicit. Since the *self* is the highest element in the hierarchy, the analysis would register that there is full explicitness and therefore there would be no need to catalogue the lower levels. The analysis of TA Excerpt 7.1, Segment 07, shows that it allows for a promising development in the construction of new knowledge, though. The subject is close to being able to form a new assurance about ‘groups’, which is verbalised in the final segment (08) “*I’m correct*”.

7.4 Results

The case outlined above with fully explicitly represented and verbalised knowledge is very rare. Very few segments, if any, contain the self, the attitude, the factuality. ‘ I know that it is a fact that this is a subject because...’ with the dots illustrating some justification for this attitude (see R. Ellis, 2004:233 for a similar example pertaining to the relative pronoun *that*). This fully explicit example is typically what might be the outcome in a retrospective protocol, in which subjects are asked to comment on their own actions and thoughts, which would be the case also in a concurrent report for which subjects had been instructed to do the same. The protocols in the present study represent a ‘stream-of-consciousness’ type of protocol, and therefore the verbal expressions are not perfectly formed communication. Rather, they are fragmentary and broken expressions; indeed they are authentic raw material of knowledge consciously available at the time of verbalisation, not rational narratives in retrospection, or intelligent comments in well-organised sentences (see also Chapter 6). Consequently, inferences need to be made when eliciting information from the protocols. The following chapter sections illustrate the typical forms that the protocols take, and the inferences that can be made from them. The inferences can then be made into coherent interpretations of the content of the explicitness of the knowledge represented in the cognitive systems of the subjects in the study.

7.4.1 Inferences in the interpretation process

7.4.1.1 Self

Fully explicit knowledge would, in this type of protocol, rarely have indications of the *self* expressed explicitly. If you are thinking aloud and revealed your own thoughts, i.e. the internal cognitive processes of your own mind, they differ from own types of

concurrent protocols in that they are not communication from the self to another entity, rather, it is the self in its own cognitive universe. Thoughts are not directed towards anyone and in consequence the expression of ‘self’ is rare because the self is not in contrast to anything or anyone else, it is a given. This is the case when knowledge has been internalised and the process is fast. Speaking slows down the process and if the internal process runs fast and smoothly less is said. The vocabulary of ‘self’, however, does occasionally appear in the TA protocols in that ‘I’ and ‘Jeg’ [I] appear in several functional situations. There are examples of the fully explicit statements discussed above:

TA Excerpt 7.2

TAS7 (VISL). PRE-Sentence 1a.

PRE-Sentence 1: <i>George painted the house yesterday</i>	
Segment	Transcription
01	I know that <i>George</i> is
02
03	the
04 (scrolls)...
PROMPT	
05	...
06	t h e subject.
07	he’s doing the paint,
08	and he is of course the noun

This protocol excerpt illustrates self explicitness (“I”), attitude explicitness (“know”), and content explicitness (“*George is the subject*”). In addition, we can see from segment 07 that the attitude has been properly caused (“*he is doing the paint*”). In other protocols, other instances of the expressions of *self* appear to be a kind of self-instruction “I need to make a tree of this”, or the subject is directing comments to someone other than the self, in which case the protocol needs to be interpreted as an expression of a metaprocess, e.g. TAS7 continues after a prompt:

TA Excerpt 7.3

TAS7 (VISL). PRE-Sentence 1b.

-----+

PRE-Sentence 1: *George painted the house yesterday*

Segment	Transcription
09	I'm just trying to put <i>the house</i> as object
10	but I'm not sure if it's an indirect object
11	or what I have to do.
12	I can't
13	...
14	I know

-----+

The wording of segment 09, in particular the word “just”, tells us that this cannot be an inner monologue, and it is clearly a response to the prompt by the experimenter.

Segments 10 to 14 become directed more towards the subject herself, and there is a return to the inner processes. Some protocol segments will be interpreted as having full explicit representation despite the protocol containing no verbalised ‘I’ or ‘Jeg’ directly connected to a particular constituent, for the reasons given above, when other indicators can be interpreted to contain and support the interpretation that the *self* is embedded in the attitude. One example is the following protocol segment from another VISL subject:

TA Excerpt 7.4

TAS2 (VISL). PRE-Sentence 1a.

-----+

PRE-Sentence 1: *George painted the house yesterday*

Segment	Transcription
01	I'm looking for something
02	...
03	it didn't the first time,
04	but it's the predicator
05	and <i>painted</i> a verb
06	and <i>George</i> is the subject
07	and a noun
08	and

-----+

The subject has some difficulty finding the wanted buttons with the correct linguistic tags, and it seems the courseware is a bit slow in responding to the commands given by

the subject as can be seen expressed in segments 01 to 03. This does not in any way make the subject doubt his own knowledge, or even speculate about alternatives, despite the lack of response from the courseware. Consequently, the protocol is interpreted as having demonstrated a fully explicit representation of subject and predicator.

7.4.1.2 Attitude

It must be stressed that when using TA protocols as a source of information on cognitive processes, it is important to seek supporting evidence of one's interpretation, if possible.

The protocol of TAS7 continues from the quoted segments above (TA Excerpts 7.2 and 7.3) as the analysis of the sentence progresses. TAS7 is a VISL subject only two weeks into the training. At this early stage in the experimental treatment, the handling of the courseware is precarious, and TAS7 makes several errors which result in the unresponsiveness of the VISL courseware, which repeatedly cause an inability to progress in the sentence. In fact, the subject would have been justified in thinking that she was not analysing the sentence correctly, and consequently, that her knowledge were deficient. This is a real test of how stable her knowledge is. The protocol excerpt above indicates that the knowledge of the two sentence constituents of subject and predicator is fully explicit, and the subject's behaviour under stress supports this. Despite several setbacks and the unwillingness of the courseware to execute the given commands for the reasons outlined above, the subject does not budge - she knows *George* is the subject, and she knows *painted* is the predicator. Segments 09 and 10 cast doubt on the *attitude* concerning the object of the sentence. Is *the house* the object or an indirect object, the subject asks; doubt has arisen after failed attempts to get the VISL courseware to respond to the command. However, as the analysis continues and the protocol unfolds, it is revealed that the doubt concerns the distinction between indirect object and (direct) object. There is no doubt that it is an object. The distinction between a direct and an indirect object is not discussed. The protocols can only be interpreted to mean that the subject has no knowledge of indirect objects beyond the *property* (see Figure 5.1 and Table 7.3).

The likelihood of other ‘attitudes’ being represented is a definite possibility especially, of course, in the pre-treatment protocols. The training leading up to the pre-treatment protocol has been so brief that the possibility of new knowledge being cemented and represented as knowledge is slim. In view of the subject-matter, the expectations for fully explicit knowledge should pertain to the functions of subject, predicator, object and should be valid for all subjects. The same broad appreciation should pertain to the forms of noun, verb, article, adjective. The expectations for the functions of subject complement, adverbial and especially head and dependent would be that ‘attitudes’ other other than knowledge would be prevalent, e.g. supposition and guesswork.

The difficulty in interpreting the verbal reports would involve being able to discern knowledge from supposition, and guesses. As discussed above, ‘I know’ is occasionally verbalised. The same is the case with ‘I guess’, ‘I think’, which can definitely rule out the content as being represented as knowledge. In less transparent cases inferences need to be made on the basis of other signs, for instances long pauses, trial and error attempts, adverbials such as ‘perhaps’ or modals such as ‘could’. The following protocol excerpts illustrate the procedure applied.

TA Excerpt 7.5

TAS2. PRE-Sentence 1b.

+-----+	
PRE-Sentence 1: <i>George painted the house yesterday</i>	
+-----+	
Segment	Transcription
+-----+	
09	and <i>yesterday</i>
10	...hmmm...
PROMPT	
11	ahmm..
12	I’m just thinking what it is
13	don’t know
14	...
15	I’m not sure
16	...adverbial
17	...ah...Ehmm...puh...
PROMPT	
18	I’m wondering what it is yesterday as a form
19	...eh...

the constituents of the sentence is a matter of a few seconds, and the inference is that the knowledge is fully explicit despite the absence of verbal indications of *self*, or *attitude*.

TA Excerpt 7.6

TAS9 (NON-VISL). POST-Sentence 1.

POST-Sentence 1: <i>Jill sold the car today</i>	
Segment	Transcription
01	And <i>Jill</i> is subject
02	and <i>sold</i> is predicator
03	and <i>the car</i> is object
04	and <i>today</i> is adverbial
05	and then I need to make a tree of this

7.4.1.3 Content

The knowledge of the subject matter is embedded in the *content* of a given proposition. This *content* has a structure which encompasses *individual*, *property*, *predication* and *factuality* (see Figure 5.1 and Table 7.3). A given representation for which the system explicitly holds all four components will have full content explicitness.

Dienes and Perner (1999) make use of a notation system similar to mathematical symbols, e.g. F, G (i.e. capital letters) symbolises *property*, the *individual* is symbolised in small case letters, e.g. a, b, etc. and the *predication* of the *property* to the *individual* is denoted by the combination of the capital letter and the small case letter, e.g. Fb. Left implicit is the temporal element of ‘*now*’, the time at which the property is true about the individual. Factuality also is left implicit and no symbol is used to indicate factuality. In verbal report the factuality may be expressed, and factuality is inherent in the system, i.e. held implicitly by the cognitive system in such a statement as ‘*I have R*’ (see *ibid*, note 8) which is explicit to the level of ‘*self*’. R denotes ‘representation’ and the fully explicit representation of a piece of knowledge has been reached when the ‘*self*’ is explicit. The factuality implicit in the statement cited thus must also have as explicitly represented the following truths which makes up the

attitude of knowledge (ibid:739) :

- (0) R is possessed by the system
- (1) R is accurate (true)
- (2) R is used by the system as an accurate reflection of reality
- (3) R has been properly caused (must not have come about by accident but have a respectable causal origin, which, when made explicit, serves to justify the claim to knowledge).

Table 7.3 Implicitness, explicitness and possible combinations

Table 1. *Possible combinations of implicit and explicit knowledge of aspects of facts (Factuality stands for factuality and/or time)*

		Represented	
		Explicitly	Implicitly
1.	Property		Individual + predication + factuality
2.	(a) Property + individual		Predication + factuality
	(b) Property + predication		Individual + factuality
3.	Property + individual + predic.		Factuality
4.	Property + + factuality		None

Source: Dienes and Perner (1999:739)

In establishing something as knowledge, factuality of the proposition is the pivotal element. A verbal expression related to the subject matter of the present TA protocols could be *This is a noun*. There is an individual, *This (=b)*, there is a property, *a noun (=F)*, and the property is predicated to the individual, *This is a noun (=Fb)*. The factuality and the time is left implicit. However, *This is a noun*, may form a given context and the inferences which can be cautiously made from the verbal expressions in the context can then be interpreted to indicate: It is a fact that this is a noun. Complete certainty that this has been properly caused (see point 3 above), which would be a precondition for it to be classified as knowledge, is difficult to ascertain in some cases such as a concurrent think-aloud protocols. The researcher therefore needs to exercise

constraints in interpreting protocols. On the other hand, there are cases of taciturnity which combined with gestures and/or actions could lead to an inference of factuality.

Dienes and Perner themselves illustrate the case with a reference to Strawson's (1959:206, referred to by Dienes and Perner 1999: 737) 'naming game' in which an object is presented to the subjects who then call out the name. In this case the *property* is the only structure which has explicit representation. That this property is predicated to an individual – i.e. the object or figure shown to the subject, - and that it is a fact that this property is possessed by the individual on the screen or card are left implicit. An example of such an instance could be a subject working on the VISL tree diagram pointing or clicking but verbalising only 'noun'. A parallel in a NON-VISL protocol would be a subject pointing with the pen or writing without verbalising. Such examples could be envisioned in instances when a degree of automaticity has been reached. Consider the following TA excerpt:

TA Excerpt 7.7

TAS8 (VISL). POST-Sentence VISL on paper. a.

Segment	Transcription
01	to start with I just find the predicator and the subject
02	and I ask what did <i>She visit</i>
03	and <i>She visit the hospital</i>
04	and an adverb (has written S P O A above the words)
05	so
06	...

TA Excerpt 7.7 shows that the student gives limited verbalisation of her actions. The verbal protocol, and the action protocol are both needed to give the observer a fair representation of the knowledge of the student. The student has analysed the sentence and physically predicated the properties to the constituents without giving verbal report of the individuals to which the properties are predicated. This happens quickly and efficiently and from the protocol can be inferred that the student is confident in her analysis. From the verbal protocol we know there are subject, predicator, object and

adverb. We know the individual instances of subject, predicator, object: *She visited the hospital*, but we do not know the individual possessing the property of adverb. The word is not mentioned at all. The action is quicker than the verbal articulation. Nor are we told specifically which properties are possessed by which individuals as the ascription of the respective properties happens through the act of writing the designations on the paper above the appropriate words. There is explicit and verbalised knowledge of the properties of subject, predicator, object and adverb[ial], there is predication of these properties to individual constituents, but the factuality is left implicit. The factuality finds expression and becomes explicit as the protocol continues. There appears to be a summing up or evaluation of the process which in the protocol is indicated by segment 05 'so', which is followed by a pause (segment 06).

In Excerpt 7.8, which is a continuation of Excerpt 7.7, it becomes clear that segments 05 and 06 are indicative of a shift in cognitive processes. Segment 06 appears to be a cognitive pause (cf. Scollon and Scollon, 1995) which is followed by a less automatic mode of behaviour. Knowledge becomes explicit, "We have subject", and it is now coordinated with action, 'writes S under *She*'. The expression "We have subject" explicitly states there is knowledge of subject in that 'have' not only indicates predication of property to an individual but also that the content is taken to hold truth and attitude explicitness (see Figure 5.1). The subject *has* the representation of the property of *subject*; the property is predicated to an individual *She*, and the verbalised attitude embedded in *have* makes the proposition attitude explicit, that is explicit knowledge.

Attitude explicitness is structurally at a higher level than content structures, and Dienes and Perner's theory stipulates that explicitness at a higher structural level indicates explicitness of any lower structure. The knowledge of the subject in the given sentence is thus held as explicit knowledge. This knowledge is held by "we", from which the inference can be made that the self is included: "we" indicating the self 'I' and 'anyone who were to look at the sentence'. In other words, "we" has generic reference. The cognitive system of this particular student has representations of the concepts of subject, predicator and object which are content explicit, attitude explicit and self explicit, which

makes this knowledge fully explicit. These representations are taken to be true by the self.

TA Excerpt 7.7 differs from TA Excerpt 7.8 in the degree of explicitness of the knowledge. From TA Excerpt 7.7 alone, the explicitness of *attitude* and *self* is missing, and even the factuality of the proposition is missing. It is the continuation of the protocol in TA Excerpt 7.8 which opens the extent of the explicitness of the knowledge to the observer.

TA Excerpt 7.8

TAS8 (VISL). POST-Sentence VISL on paper. b.

+-----+	
POST-Sentence VISL on paper: <i>She visited the hospital yesterday</i>	
+-----+	
Segment	Transcription
+-----+	
07	and we have subject (writes S under <i>She</i>)
08	and we have predicator (writes P under <i>visited</i>)
09	...
10	<i>She visited,</i>
11	and we have the object
12	and it's a group (writes O, but not g under <i>the hospital</i>)
13	and we have the head (<i>hospital</i>)
14	it's a noun
15	and we have dependent (<i>the</i>)
16	it's an article
17	and we have <i>yesterday</i>
18	and it's adverb (writes A:adv)
19	ja [yes]
20	ok
+-----+	

That the representation of the knowledge has been properly caused becomes clear from segments 10 and 11 in TA Excerpt 7.8 ‘*She visited* (segment 10) and “we have the object” (segment 11). By means of identifying the activity, the participant roles of DOER and DONE-TO, the subject reveals the cognitive processes involved in reaching the goal of establishing ‘*the hospital*’ as the object of the sentence.

The above analysis illustrates how the structural constraints on which the theory builds can assist in establishing explicitness at a lower level despite the fact that this knowledge is not represented in the verbal protocol. However, since the constraints posit that explicitness at a higher structural level is only possible if the lower structures are explicitly represented by the system, the full analysis will reveal that the knowledge is explicitly represented in the verbalised structures, which means in the case of above example that to establish the explicitness of the factuality in the content structure the protocol needs to be analysed to a higher structure.

TA Excerpt 7.9

TAS11 (NON-VISL). PRE-Sentence 1a.

PRE-Sentence 1: <i>George painted the house yesterday</i>	
Segment	Transcription
01	hvad skete der [what happened there]
02	det må [it must]
03	...
PROMPT	
04	ja [yes]
05	han malede [he painted] (writes o under <i>painted</i>),
06	og hvem malede? [and who painted]
07	det gjorde <i>George</i> [George did] (writes x under <i>George</i>),
08	og hvad malede han? [and what did he paint]
09	han malede huset [he painted the house] (underlines <i>the house</i> , writes Δ)
10	og så ved jeg ikke, hvad den sidste er, [and then I don't know what the last one is]
11	men det er så...de to der [but this is then...those two] (indicates <i>the house</i>)

Dienes and Perner (1999) point out that the constraints apply only to one representation at a time. However, this does not preclude the possibility of two independent representations, one making something explicit at the higher level and the other representing something at the lower level implicitly. For example:

- (a) "I know that there is some fact involving F"
(i.e., explicitly representing attitude and factuality).
- (b) "F" (i.e., implicitly representing predication of F to b).

This is possible, but the point is that (a) does not implicitly represent the fact Fb. Rather, it explicitly represents the knowledge that there is something concerning the property F. In that case there is no implicit knowledge of Fb being a fact. That this is not implicit in (a) can be seen from the fact that Fb is not a supporting fact of (a), that is, one can know that there is something about F without the fact Fb. (Dienes and Perner, 1999:740).

In the TA protocols there is a case which illustrates Dienes and Perner's point (see TA Excerpt 7.9 and TA Excerpt 7.10). The verbalisation in TA Excerpt 7.9, segments 05 to 10, invites the inference that there is explicit knowledge relating to subject, predicator and direct object. However, a second reading of the protocol allows only a more restrictive interpretation. The concepts verbalised are not really the subject, the predicator and the object, but x, o, and Δ. These are the graphic symbols used in Danish elementary schools for subject, predicator and direct object, respectively (see more below).

From TA Excerpt 7.10 it appears that TAS11 is aware of some fact involving the symbol of a square, the graphic symbol used by Danish elementary schools to indicate indirect object.

TA Excerpt 7.10

TAS11. PRE-Sentence 1b.

PRE-Sentence 1: <i>George painted the house yesterday</i>	
Segment	Transcription
12	og så er der en [and then there is one]
13	...
14	her [here]
15	jeg ved ikke, hvad det er for en, [I don't know what it is]
16	jeg må hellere lave en firkant [I'd better make a square] (writes □ under <i>yesterday</i>).

This excerpt (7.10) involves explicitness at the very lowest level. There is a property represented, manifested in the symbol of the □, however, the property of the symbol is

not fully explicit. There is a memory trace of something involving the symbol of the square, but no specifics of the quality nor of any instance which might qualify for the property involved. The subject cannot predicate the property to any individual, but since *yesterday* is the only element left to be accounted for, the property is ‘parked’ there, one assumes until such a time when it can be put to proper use or assigned a proper place.

There is no representation of the fact Fb (i.e. the TA subject has no representation of the property embodied in *yesterday*, nor of the property of \square , nor of factuality of either) despite the property being ascribed to an individual, i.e. *yesterday*. It is clear from the protocol that there is no knowledge, “jeg ved ikke hvad det er for en” [I don’t know what this one is] (segment 15), and the marking of the individual there is no predication nor any factuality. There is no content explicitness and even the property is only fragmented and may not even be traceable in an indirect test. Such a test would, for instance, be to show a word for a few milliseconds with the consequence that a semantically related word would be processed or identified faster than a semantically unrelated word. For instance, if the word *doctor* was shown for just long enough for the visual system to perceive the word, an indirect test would subsequently show that the word *nurse* would be identified as being a proper word faster than a semantically unrelated word (Dienes and Perner, 1999:742; Marcel, 1983).

In TA Excerpt 7.10 there is no identification of the property, only the symbol of the property, and there is no predication of the property beyond an attachment to an entity which has been left unaccounted for by the process. The process of analysing the sentence has activated the field of analytic symbols from prior experience, but the field is too broad or too general for it to be of any use in the particular situation, because it is the field of symbols or the construct of analysis that has been activated, not the field of concepts. The narrower scope embedded in the concept behind the symbol of the square has not been activated. There is a representation of sets of symbols, i.e. x , \circ , Δ , \square , but only x , \circ , and Δ can be mapped on to appropriate concepts. As far as the word *yesterday* is concerned, there is no representation of that word, nor does there seem to be any representation of the structural concepts which could apply to a description of it. In the light of the student trying to rely on the recall of old knowledge from prior experiences

with grammatical analysis, it is not surprising that the concept of ‘adverbial’ is absent from the mental inventory, nor is it surprising that the symbol of the square is activated since the symbols used in the Danish tradition would contain only the four particular symbols listed above. The student has brought up the full catalogue of symbols and since there is one symbol left and one word unaccounted for, it seems only logical to pair the two. The process illustrates perfectly Dienes and Perner’s point about how the constraints apply only to one representation, and how one representation can be at one level when another representation is at another level. Thus, the theory helps illuminate the processes of learning as well as the shortcomings of learning. This point becomes even more evident as the protocol continues and it will be further discussed below.

There is explicit knowledge of the roles, symbols, and relationship of the constituents as far as activity, doer, done-to is concerned (x, o, Δ). These properties are predicated to the individual sentence constituents, and the factuality appears to be explicit for the activity, doer and done-to in contrast to the last item (*yesterday*), to which there is no factuality attached, in fact the subject explicitly states that there is no knowledge: “*og så ved jeg ikke, hvad den sidste er*” [and then I don’t know what the last one is]. There is explicit knowledge of content as well as attitude (knowing) and self (I) for x, o, Δ , but not for \square . The next question is whether this knowledge been properly caused and justified? The causal efficacy of the x, o, Δ are evident from TA Excerpt 7.9, whereas TA Excerpt 7.10 tells us that as far as \square is concerned there is only explicit knowledge of the symbol of a property, but it cannot be predicated to anything in particular: “*jeg må hellere lave en firkant*” [I’d better make a \square] (writes it under *yesterday*), nor is it taken to be a fact: “*Jeg ved ikke, hvad det er for en*” [I don’t know what this one is], and evidently it has not been properly caused. TAS11 has a representation of the property of the symbol, but is not able to predicate it to an individual instance nor establish its factuality.

The system (TAS11) holds the following representations:

R1 = George is x /Fb is a fact

R2= painted is o /Fb is a fact

R3= the house is Δ /Fb is a fact

R4= there is something involving □

For R1,2,3, the system holds:

- (0) R is possessed by the system
- (1) R is accurate (true, reflects the fact Fb)
- (2) R is used by the system as an accurate reflection of reality (judged to be true)
- (3) R has been properly caused (must not have come about by accident but have a respectable causal origin, which, when made explicit, serves to justify the claim to knowledge)¹¹³

For R4 the system holds only the initial stage:

- (0) R is possessed by the system

The system (of TAS11) does not hold R4 as true, nor as an accurate reflection of reality, nor as having been properly caused. TAS11 is only aware of the existence of a property; this property represents only fragmented knowledge because the only the symbol representing the property is held by TAS11.

7.4.2 Explicit knowledge and consciousness

If the above-mentioned conditions 0-3 hold, then the system represents its attitude of knowing explicitly. There is knowledge of the fact that George is the subject, i.e. x , even though this says nothing explicitly about *self*, which demands a metarepresentation such as *I know that it is a fact that George is the subject*. Only when *content*, *attitude*, and *self* can be represented explicitly does knowledge count as conscious.

Consciousness involves being aware of our mental states (cf. Higher-Order-Thought Theory of Consciousness¹¹⁴, Dienes and Perner, 1999:41): *I am conscious of George being x* (a higher-order thought)→ *I know that George is x* (i.e. it is not just imagined). It does not require the actual entertaining of a higher-order thought but only the

¹¹³ The hierarchical system outlined by Dienes & Perner, 1999:739); see also this thesis p.257.

¹¹⁴ In order to represent knowledge as explicit it needs to be at least attitude explicit, and this requires a second-order thought involving consciousness of representing something as a fact; see Dienes and Perner, 1999:796, Rosenthal, 1997; and cf. Carruthers, 1992.

potential for forming such a higher-order thought (ibid.:741). That potential is present for x , o , Δ , as is clear from TA Excerpt 7.9, but not for \square (see TA Excerpt 7.10).

In order to be able to form a metarepresentation of the content of one's knowledge, it is necessary first to entertain first-order thought, e.g. *George is x*. Content needs to be represented explicitly, and the necessity for first-order thought ties consciousness and explicitness together, and verbal access to knowledge is tied to consciousness. The ability to address the content of one's knowledge verbally characterizes conscious and explicit knowledge, that is, verbal reference requires explicit representation of content. Tied to explicit representation is declarative knowledge and voluntary control. Declarative knowledge is accessible; it consists of 'knowing that' in contrast to the procedural knowledge of 'knowing how' (Anderson, 1980; Dienes and Perner, 1999:743; De Keyser, 1997; see also Ryle, 1949).

Implicit knowledge is not verbalisable in that verbal reference requires explicit representation of content, particularly factuality. In the protocol TA Excerpt 7.10 "*Jeg ved ikke, hvad det er for en, jeg må hellere lave en firkant*" [I don't know what this is so I'd better make a square] (segments 15-16) there is no explicit representation of content beyond the suspicion that there is some fact involving \square (F). Therefore, there is no voluntary retrieval, there is no knowing, and there is no memorial state of awareness. Only the fragmentary property is represented, and the combination of the symbol of \square and the word *yesterday* is based on guessing, but there is an implicit understanding that the symbol of the square belongs to the same construct as the other three symbols.

Dienes and Perner's theory of implicit and explicit knowledge builds on the tenet that explicitness is a matter of degrees and that the structural constraints on explicit knowledge can be explored from *content* explicitness to *attitude* explicitness to *self* explicitness as the structure progresses up the hierarchy. The lowest element of explicitness they operate with is *property* explicitness, e.g. cat or catness. However, they do not discuss the internal structure of 'catness' which must be a composite concept in its own right. In parallel, the same must be true for noun or 'nounness'. A cat has many similarities with a dog, what exactly are the knowledge structures which

excludes doubt about the property of catness? Again, in parallel, what exactly are the structures which make us certain about the classification of nouns? What does it take to make us certain, or in the formula of Dienes and Perner, what makes us able to say *I know that it is a fact that this is a noun*? Knowledge of the conceptual structure of a noun or any other linguistic concept is structurally complex and includes at the minimum morphological, syntactical and semantic features. The student who is asked whether this or that word is a noun needs to compile and process all these pieces of knowledge in order to be able to establish ‘nounness’ or ‘non-nounness’. The truth of this is, perhaps, self-evident, but the relevance of Dienes and Perner’s theory is that fragmentary knowledge will be revealed without the need to establish exactly which piece of the puzzle it is that is missing before ‘nounness’ or ‘non-nounness’ can be established; in other words, no test of the composite elements of ‘nounness’ is needed because the test of the knowledge of a given property, in this case nounness, is built into the hierarchy of Dienes and Perner’s model. If the property of ‘nounness’ is incomplete knowledge, then it cannot be predicated to an individual, nor can it be established as a fact, and therefore it cannot be attitude explicit, i.e. count as knowledge. It can, of course, be activated as guesses or hypotheses but not as bona fide knowledge. There is no doubt that linguists and educators would find it of interest to be able to establish which of the composite parts of, for instance, ‘nounness’ students lack or find difficult. For the operationality of Dienes and Perner’s theory it is not essential, though. The quality of the knowledge, indeed if there is knowledge at all, will be revealed when one looks for factuality or attitude explicitness. The following protocol excerpt comprises a case which demonstrates the mechanisms involved:

TA Excerpt 7.11

TAS7 (VISL). PRE-Sentence 1c.

-----+
 PRE-Sentence 1: *George painted the house yesterday*

Segment	Transcription
15	ahm...there...and we have hm..the..
16	I mean <i>the house yesterday</i> ...that’s..
17	we have to put it together because it’s a group <i>the house</i> is one group
18	and <i>yesterday</i> is a group,

19 *the* it's an article
20 .. mmm.. an article we have it here
21 and *the house* is the object the direct object (she has not
grouped the two words, and it doesn't work)
22
23 no
24 ...
25 it
26 ...
27 was
28 ...
29 wrong
30 ...
31 let's say it's an object
32 ..hmm
33 it doesn't work I guess
34 ok
35 at least I know that *house* is a noun
36 ...
37 and
38 ...
39 let me see,
40 build trees
41 .. hrmm...(goes to Tools and Build tree)
42 ...(clicks on the boxes and art disappears from *the* and n from
house)
43 ...
44 what happened
45
46 god
47 ...
48 they'll
49 ...
50 I tried
51 ...
52 hmm
53 ...
PROMPT
54 ahm...
55 I'm just trying
56 *the house* is the object,
57 and it doesn't work
58 (laughs)
59 ... (clicks on noun)
60 ok (the software accepts this input)
61 ...
62 it's a noun,
63 and it's an object

64 ...
65 I'm not quite sure about if it's
66 ...
67 it must be an
68 ...
69 a direct object
70 ...
71 it doesn't work
72 (laughs)
73 let's see (scrolls)
74 ...
75 hmm...
76 it doesn't work anymore
77 ...
PROMPT
78 I'm just trying to put
79 ...
80 *house* is the object
81 I mean *the house* is what you group together
82 ...
83 this is the object
84 ...
85 hm...
86 and *the* the article (accidentally removes P/v),
87 and it's still...
88 ok (now has art and n in place).
89 and it's adverbium (should be dependent)..
90 no...(discovers that P/v has disappeared)
91 I don't know what's going on.
92 hm...
93 den bliver ved med at slette det der [it keeps erasing that]
94 det har jeg allerede gjort [I already did that]
95 *painted* is ...(scrolls)
96 and ...
PROMPT
97 (Laughs)
98 I'm lost. (Scrolls first Function then form)
PROMPT
99 It's a verb.
100 It's a noun word.
101 We have it
102 ...no...
103 it's the predictor
104 ja [yes]
105 ... ok
106 I'm not quite sure about *yesterday*
107 ...
108 let me have the trees

109 oh my god what's going on (goes to Tools/Build tree and
 starts over from the top)
 110 oh my god...
 111 ok fine
 112 ... god
 113 ...
 114 sorry about that
 115 hm...
 116 the subject
 117 and we know it's a noun.
 118 so we go to the next which is
 119 ...
 120 predi...¹¹⁵
 121 preticator
 122 ...
 123 and
 124 ...
 PROMPT
 125 I'm just putting it in the right place,
 126 and I know that *painted* is the pretigor
 127 ...
 128 there
 129 ...
 130 like this
 131 and the *house*
 132 ...
 133 it's a noun
 134 ...
 135 and *the* it's an article if I'm
 136 ...
 137 and *the house* is the
 138 ...
 139 object (still hasn't grouped them)
 140 ...
 141 ahm...
 142 maybe because it's wrong it doesn't put it
 143 or
 144 ...
 145 indirect object
 146 no
 147 ...
 148 pardon
 149 ...
 150 I have to
 151 ... (finally groups them)
 152 hm

¹¹⁵ TAS7 keeps having difficulties in pronouncing 'predicator' correctly.

153 like this
 154 ok
 155 I know that it's the object
 156 ...
 PROMPT
 157 hm...
 158 I'm just wondering
 159 ahm...
 160 no
 161 ...
 162 it doesn't work correctly
 163 it must be me
 164 ... (laughs)
 165 ...uhm...
 166 suddenly
 167 ...
 168 I don't know what's going on
 169 .. (laughs)
 PROMPT
 170 I'm just trying to put *the house* as object
 171 but I'm not sure if it's an indirect object
 172 or what I have to do
 173 I can't
 174 ...
 175 I know
 176 ...
 177 this (points the cursor at *the*) is adverbium (should be
 dependent)
 178 ...
 PROMPT
 179 it's an adverbium
 180 it doesn't work so it's not.

In TA Excerpt 7.11, segment 21, the subject states that *the house* is the object and then adds that it is the direct object¹¹⁶. This is a subject from the VISL group so the software has to be manipulated to enter the analysis performed and perceived by the subject. However, the computer does not respond in the desired way, partly due to inept handling by TAS7. The subject's reaction to this malfunction tells us something about the nature of the knowledge held by TAS7. If this knowledge was not fully *content* explicit, the subject would not be able to sustain her own effort to make the computer comply with her endeavours to enter the correct analysis. From segment 21 to 130 there

¹¹⁶ The beginning of the TA protocol of TAS7 is found in TA Excerpt 7.2.

is no doubt that *the house* is the object, the direct object of the sentence, and it is stated in segment 21 and restated in segments 31, 46, 53, 59, 70, 73 and 129. The inference that can be made from this persistence is that the knowledge is represented as a fact, i.e. the factuality of the content makes it content explicit. It can be further inferred that this is knowledge held by the *self* of the subject.

Some doubt is expressed in segments 132 (“*maybe it’s because it is wrong*”) and 135 (or...indirect object) it may seem that the failure of the software in responding to the input is beginning to rattle TAS7 to such an extent that she shows signs of doubting herself and her knowledge. If one reads the protocol carefully, however, it becomes clear that there is no doubt that *the house* is an object, there is only doubt as to which type of object. The subject is astoundingly self-assured and even when faced with repeated failure to get the computer to respond to the input TAS7 is not considering changing her judgement. She considers looking into variations on the theme (TA Excerpt 7.11, segments 135 and 160). TAS7 is persistent to a degree which can only support the inferences made above. After having considered the option of indirect object, TAS7 returns to her statement of direct object in segment 145 (“*I know that it’s the object*”).

The soundness of the inferences are revealed at the end of TA Excerpt 7.11 when the classification of *the* is at issue. She tries to call it an ‘adverbium’ (clicks on adverbial), and immediately when the computer refuses to respond, she accepts that it is not. The reactional pattern is so different from that revealed by the protocol concerning the object that the two response patterns warrant two different interpretations, and it demonstrates the stability of TAS7’s knowledge that *the house* is an object. TAS7’s knowledge of how to classify *the* is not explicit knowledge. It is more like a guess; the content is not explicit and the lack of knowledge is not a matter of knowing the right term (compare the object variations above). The subject discloses no knowledge at all. The subject knows there is something involving *the*, but the concept is not content explicit, and the correct terminology is not represented at all. Through one hundred and fifty-nine segments, TAS7 maintains that *the house* is the object, but it only takes three segments for her to admit that she is wrong when she says that *the* is an ‘adverbium’

(segment160), because the computer does not allow this input – this despite the fact that she had the same difficulty with the object.

It is clear from the verbal protocol of TAS11 (see TA Excerpts 7.9, 7.10 and 7.12) that the cognitive system of this subject has no representation of the connection between symbols of x, o, Δ, □ and the more appropriate denominations of S, P, O, A. There is knowledge of the terms subject, object, and adverbial (see TA Excerpt 7.12, segments 20, 47 and 55). The representations are not of the same quality, though. Subject and object are represented simply as terms, but they are not predicated to any structure nor are they in any way connected to the symbols of x, 0, Δ, □, which the subject applied in the initial analysis of the sentence (see TA Excerpt 7.9, 7.10). The protocol of TAS11 demonstrates that the knowledge of x, o, Δ is properly caused. TA Excerpt 7.9 demonstrates equally strongly that the knowledge of the concepts predicator, subject, and object was not explicit beyond the property nor did the cognitive system of TAS11 represent any relationship or connection between the old knowledge and the new knowledge. It can be concluded from the protocol of TAS11 that the instruction has so far failed to make the subject apprehend the underlying common concepts and constructs (see TA Excerpt 7.12). It needs to be said, though, that this protocol is a PRE-Sentence recording, which took place a mere two weeks into the experimental treatment. The tenuous impact of the instruction becomes more and more obvious as the protocol continues.

TA Excerpt 7.12

TAS11 (NON-VISL). PRE-Sentence 1c.

+-----+
 PRE-Sentence 1: *George painted the house yesterday*

Segment	Transcription
17	<i>painted</i>
18	det er så [that is]
19	... hmmm...
20	subject (writes S under <i>painted</i>),
21	...
22	og [and]
23	...
PROMPT	

24 ahm...
 25 jeg kan ikke huske [I can't remember]
 26 ...
 27 om [if]
 28 ...
 29 *George*, om det er [*George* if that is]
 30 eh...
 31 ehm...
 32 det er navneord [it's a noun] (writes n)
 33 og det er *the house* også [and *the house* also is] (writes n
 under *house*)
 34 og [and]
 35
 36 der er et eller andet deroppe (points at *She* which is in another
 sentence), [there is something up here]
 37 ...
 38 det er [it is]
 39 ...
 40 hvad er det det hedder? [what is it called]
 41 ...
 42 det hedder [it is called]
 43 ...
 PROMPT
 44 ja [yes]
 45 men jeg har ikke lige noget i tankerne lige nu. [but I have
 nothing in my mind right now]
 46 ...hrmppp.
 47 subject, object
 48 ...
 49 sådan noget der (writes S under *George*) [something like that]
 50 og så tror jeg at [and then I think that]
 51 ...
 52 *the house*
 53 det er [it is]
 54
 PROMPT
 55 *yesterday* det er i hvert fald adverbial [*yesterday* that is
 definitely adverbial] (writes A under *yesterday*)
 56 og så *house* det er [and then *house* that is]
 57 ...
 58 det er head [it is head] (writes H under *house*),
 59 og det andet det er det er en SUB [and the other one that is a
 SUB] (writes SUB under *the*).
 +-----+

The one item which stands out in TA Excerpt 7.12 is the verbalisation of the analysis of the adverbial *yesterday*. The representation held by the system with regard to this item

appears to be different from those of subject and object. TA Excerpt 7.10 makes it clear that the subject's awareness of subject and object is at the same level as that of the symbol of the square in TA Excerpt 7.9, i.e. not content explicit as there is no factuality, and nor predication to an individual; even the property is a fragmented structure. The knowledge made explicit goes no deeper than the awareness of the existence of the concepts. Segment 55 in TA Excerpt 7.12 demonstrates that the representation held of *yesterday* is different from the others. There is a property, (adverbial), which is predicated to the individual *yesterday*, and the remaining issue is whether it can be established as a fact. The linguistic expression used by TAS11 in segment 55 reveals a certainty on the part of the subject that lends itself to the inference that there is a representation of the fact that *yesterday* is an adverbial. This means that there is content explicitness of this representation. The verbal expression "i hvert fald" [definitely], and the context which frames the verbalisation both point towards the inference that there is no doubt in the student's mind that this is the case.

The context is significant in that the verbalisation follows a prompt by the researcher. The prompt does not lead to pauses or irrelevant comments nor is there any sign of hesitation. This is one thing of which the student has ready knowledge which can be produced under pressure. The difference in representations of the concept and instance of 'adverbial' and the verbalisation related to the other constituents in the sentence in question support the inference that the representation of adverbial is different. This becomes evident when comparing segments 43-54 with segment 55. The TA Excerpts 7.9, 7.10, and 7.12, which together comprise one protocol containing the analysis of *George painted the house yesterday*, illustrate the premise that the constraints in the explicitness hierarchy only hold true for one representation at a time, and that the knowledge of complex facts contain a combination of representations making up a composite object. A given piece of knowledge is in existence alongside and in combination with another piece of knowledge, and the two are constrained at different levels independently of each other. That this is the case with the two systems of analysis which the TAS11 protocol reveals becomes apparent from the very assured and unproblematic analysis of x , o , and Δ . By activating prior knowledge, the subject has analysed the sentence with regard to these three sentence functions, but leaves *yesterday*

unaccounted for. When activating new knowledge, *yesterday* is the only word whose sentence function the subject is sure of, underlining the tenet of the theory regarding co-existence of different levels representations and constraints in that the *yesterday*/adverbial representation is the only one which is attitude explicit.

TA Excerpts 7.9 and 7.12 reveal a difference in strategy. In TA Excerpt 7.9 the strategy is clearly problem-solving, and the subject justifies the knowledge of the first three sentence functions. That knowledge does not extend to the adverbial *yesterday* for which there is nothing to draw on in the prior knowledge. However, in TA Excerpt 7.12, there is no doubt that *yesterday* is an adverbial, and the strategy is no longer one of problem-solving but one of drawing on the memory of a particular instance. The protocol in segment 55 is an emphatic expression which leaves no room for doubt about the representation. In addition, the verbalisation comes so quickly after a prompt that this leaves no room for consideration or hypothesis-testing. The analysis of these elements will point towards a case of ‘instance’ recall. In other words, the subject remembers specifically from a previous instance that *yesterday* is an adverbial. This leaves us to conclude that there is explicit knowledge of the word *yesterday* being an adverbial, but there is no explicit knowledge of the concept of adverbial which could be transferred to other instances of the category and lead to a problem-solving strategy in new instances.

The difference in quality between old knowledge and new knowledge is laid open in TA Excerpt 7.12 segments 25 “*jeg kan ikke huske*” [I can’t remember], 40 “*hvad er det det hedder?*” [what is it called?], 45 “*men jeg har ikke lige noget i tankerne lige nu*” [but I have nothing in my mind right now], and 50 “*og så tror jeg at*” [and then I think that]. In segment 25 there is a specific reference to memory, and this is taken up again in segment 45 when it becomes clear that the subject is trying to retrieve knowledge from memory and there is nothing there. The subject cannot verbalise something of which there is no consciousness. In segment 50 there is a reference to supposition rather than knowledge, which indicates that the knowledge has not been properly caused, and thus the knowledge is not bona fide knowledge because it has been placed outside the “knowledge box” (Dienes and Perner, 1999: 737), therefore the student’s cognitive

system has no representation of this as a fact. The subject states that “*så tror jeg*” [then I think that], in which statement “think” is the decisive factor ruling out factuality, which is a necessary¹¹⁷ condition for conscious knowledge (ibid.:741).

7.4.3 Knowledge, accessibility and control

Different types of knowledge accord with different types and degrees of performance. Procedural knowledge allows task performance of a quality on a par with or even better than declarative knowledge. One is no better than the other as per definition; it all depends on the task demand; the latter type has traditionally been associated with formal education and the former with execution of automatic action or motor skills. In the paradigm of language learning both procedural and declarative knowledge are desirable, and both play a significant role in linguistic proficiency. Dienes and Perner (1999:743) make a case for connecting procedural knowledge with implicit knowledge and explicit knowledge with declarative knowledge, and they support this with references to Karmiloff-Smith (1986; 1992) and Squire (1992). Basically, implicit knowledge is procedural in nature and to be likened to “knowing how”; whereas explicit knowledge is declarative and to be likened to “knowing that” (see section 7.4.2). As we have seen, fully explicit knowledge is verbalisable, i.e. is declarative, which means that it is conscious, can be accessed, and therefore can be described as being under voluntary control, i.e. it can be controlled at will.

Implicit knowledge cannot be verbalised, i.e. it is tacit rather than declarative, and it is not conscious in the sense that it can be described by the holder of the knowledge. This raises the question of accessibility and control. Explicit knowledge can be recalled at will, this is not the case with implicit or procedural knowledge, but this does not mean that it is inaccessible, “All knowledge must be accessible in some way or it would not qualify as knowledge [...] in any case there would be no evidence that there was any knowledge at all” (Dienes and Perner, 1999:743). Procedural knowledge is accessible in that it will affect performance and can be tapped in order to carry out certain tasks, but it is not available for verbalisation. Since a distinguishing factor is the explicitness of factuality, it becomes clear that procedural knowledge may be an efficient type of

¹¹⁷ Explicit representation of factuality might also be a sufficient condition for conscious access (Dienes and Perner, 1999:745).

knowledge for some tasks; but the type of knowledge needed for hypothesis testing and problem solving must include explicitness of factuality, which will allow the transfer of application from one item to the next.

Procedural knowledge may be fast, but it is limited in application to one specific area. The control that can be exercised over procedural knowledge does not extend beyond its specificity, since factuality is not explicitly represented, and thus it cannot be used to evaluate new areas, since this would demand a second-order thought or a meta-representation. The think-aloud protocols illustrate this difference (see for instance TA Excerpts 7.9, 7.10, and 7.12). This may inform our understanding of the learning processes, and thus help us gain new insight into the mechanisms which need to be modified in order to enhance learning.

The limited capacity of procedural knowledge and the narrowness of applicability is illustrated in the protocol of TAS11 when TA Excerpts 7.9 and 7.10 are combined with TA Excerpt 7.12. The subject is incapable of combining his knowledge of *yesterday* as an adverbial (TA Excerpt 7.12) with his knowledge of sentence analysis (TA Excerpts 7.9 and 7.10). The reason for this might be the difference in sources (representations) which this performance draws on. TAS11 analyses the sentence with his old knowledge, which is declarative and explicit. It does not include adverbials. His new knowledge of adverbials is explicit and declarative as far as his memory of a specific instance of the word *yesterday* is concerned¹¹⁸, but it does not extend to the property of adverbials. The knowledge of the concept is not factuality explicit and cannot be of any use in hypothesis testing; nor can it be transferred to the area of sentence analysis in general. This could be the reason why the TA protocol excerpts of TAS11 read as if they were two different protocols rather than being a continuous one. Indeed, they seem to be posing two different analyses of the same sentence because from the perspective of knowledge type, accessibility, and control they are. In contrast, the protocol of TAS3 shows how conceptual knowledge can be used to hypothesise about the role of *yesterday*.

TA Excerpt 7.13, segments 16-17, “*og så er det er noget med tidsangivelse*” [and then

¹¹⁸ The very first chapter in Bache et al. (1999) has as its third sentence *He left his wife yesterday*.

there is something involving a definition of time] “*og det en adverbial*” [and that’s an adverbial], demonstrates that the explicit knowledge of the property of adverbials has been properly caused, and it is predicated to an individual in segment 22 (*yesterday*).

TA Excerpt 7.13

TAS3 (VISL). POST-Sentence VISL on paper.

+-----+

VISL POST-Sentence on paper: *She visited the hospital yesterday*

+-----+

Segment	Transcription
[01-15]	
16	og så er det er noget med tidsangivelse [and then there is something involving a definition of time]
17	og det en adverbial [and that’s an adverbial]
18	nej [no]
19	adverb og adverbial [adverb and adverbial]
20	eller omvendt var det [or the other way round it was]
21	eh...
22	og det er <i>yesterday</i> [and that is <i>yesterday</i>]
23	sådan [that’s it]

+-----+

The property of adverbial is not in question only the appropriate term to apply to it (segment 20). The knowledge is declarative and can therefore be used for hypothesis testing, i.e the student will be looking for an individual to which the quality related to definition of time can be predicated, and this is found in *yesterday*.

An examination of the two specific test results from the quantitative study makes it possible to relate the qualitative study to the quantitative achievements with regard to items on adverbials in the pre-test. The pre-test success rates for the items relating to adverbials seem to correspond to the elicited information from the TA protocols of TAS3 and TAS11. In the quantitative test items several types of adverbials were tested for: simple ones (item 10), which were also adverbs (here, now, hardly), and complex ones (item 18), which were primarily prepositional groups (e.g. in New Street, down a long hill, on warm evenings). The pre-test results of TAS11 (the TA Excerpts cited were from a PRE-Sentence) show that none of the adverbials were correctly identified and categorised. In contrast, the post-test result of TAS3 (the TA protocol cited is a POST-Sentence) reveals that the short adverbials whose forms were adverbs were all

identified, and so were the complex ones¹¹⁹. If TAS11 had control over the possessed knowledge of adverbials, that is, if the identification of *yesterday* as an adverbial was more than a memorial trace of a specific incidence, the expectation could be that the short adverbials/adverbs would be identified, which was not the case.

Instances which demonstrate the necessity of explicit knowledge for hypothesis-testing, and a necessary condition for the quality of being intentionally applicable, which explicit knowledge possesses, can be found repeatedly in the TA protocols. In the post-treatment protocols there are examples of students who carry out the analysis tasks with very little verbalisation, but if the VISL programme makes them aware of some erroneous input, or if they get to a point in their analysis (this is especially true for the NON-VISL subjects) where they are uncertain about how to proceed, the usefulness of explicit knowledge becomes clear in that this type of knowledge can be activated and applied intentionally (i.e. voluntarily) in a systematic manner, which is not an option with implicit knowledge.

TA Excerpt 7.14

TAS3 (VISL). PRE-Sentence 2.

+-----+ PRE-Sentence 2: <i>She is a very beautiful woman</i> -----	
Segment	Transcription

[01-76]	
76	<i>very</i> det må være [<i>very</i> that must be]
77	...
78	jeg prøver og se om det er adverbial [I'll try if it's an adverbial]
79	næh [no]
80	...
81	<i>very</i> , jeg skal lige prøve og se hvad der er [<i>very</i> , I'll just try what there is]
82	...
83	nåh [oh]
84	det er selvfølgelig dependent igen, fordi det er en gruppe [it's of course a dependent again because it is a group]
85	og så er det [and then it is]
86	...

¹¹⁹ TAS3 also identified the adverbials in the pre-test.

87 herovre adverb [over here adverb]
88 ja sådan [yes, that's it]

The issue verbalised in TA Excerpt 7.14 concerns the classification of *very*; first with regard to function and subsequently with regard to form. There is no attitude explicitness. The subject's cognitive system has not placed *very* as an adverbial in the 'knowledge box', but TAS3 has knowledge of the concept adverbial and *very* could fit that, so this is the first hypothesis (segment 78). The protocol demonstrates that there is no explicit representation in the cognitive system of *very* as an adverbial: "*jeg prøver og se om det er adverbial* "[I'll try if it's an adverbial] (segment 78). The words express a possibility (i.e. hypothesis-testing), but there is no factuality.

The VISL courseware does not accept the input of adverbial so TAS3 re-evaluates (segments 80-83), and the protocol tells us very little about the cognitive processes, of which TAS3 verbalises very little as they are ongoing, but from the resulting verbalisation the observer can infer that visual input from the courseware has activated knowledge about groups and their structure (segment 84). TAS3 must have surveyed the screen, and traced the larger context from which the item in question originated. The knowledge of the conceptual framework can be used for scrutiny of the context and its possibilities. TAS3 has a representation of groups and their structure of *Heads* and *Dependents* and is able to place *very* in that framework. The subject verbalises the justification for the analysis, and we can see that it has been properly caused. Subsequently, the form of adverb poses no difficulties, and the subject is again in control of the knowledge (segment 87). The protocol of TA Excerpt 7.14 exemplifies two different control processes. In the case of the function of *very*, the control pertained to the conceptual structure of the knowledge over which the subject could exercise voluntary control and construct the knowledge necessary to classify *very* as a dependent in a group.

A different case can be observed in TA Excerpt 7.15. TA Excerpt 7.15 is similar to, yet different from, TA Excerpt 7.14. The protocol reveals some knowledge, but it also reveals that this knowledge is fragmentary.

TA Excerpt 7.15

TAS5 (VISL). PRE-Sentence 1.

+-----+
PRE-Sentence 1: *George painted the house yesterday*

Segment

Transcription

[01-60]

61 Then I want to find out what *yesterday* is
62 I think
63
64 it's an adverbial
65 yes
66 and ah.... ah ..
67 no.. (laughs)(picks article)
68 ah..... hmmm...

PROMPT

69 ah...
70 I'm not sure
71 I thought it was an adverbial
72 maybe I think it's a pronoun
73 I just gonna try some (picks pronoun)
74 ...
75 no, it isn't (laughs)
76 ahmmmm
77 ...

PROMPT

78 ya
79 .. ahm... (sighs)
80 it tells me something about time
81 ... (sighs)
82 ah...
83 it's
84 ...
85 conjunction perhaps
86 ...
87 no
88 ah...uhm
89 I don't know (picks adverb)
90 oh
91 yes
92 I didn't know what it was so I just
93 ...
94 well, tried some things.

+-----+

Segment 62, “*I think*”, tells us that the idea that *yesterday* is an adverb is not *attitude* explicit, and TAS 5 repeatedly uses expressions of similar uncertainty in segments 71, 89 and 92, which demonstrates there is no factuality. The term attributed to *yesterday* should have been ‘adverb’ rather than ‘adverbial’, and the fragmentary nature of the knowledge, which gives rise to this error, provides further evidence that TAS5 is acting on a supposition. There is knowledge that there is some fact involving ‘adverbial’, but it cannot be predicated to *yesterday* as a fact. In the following segments TAS5 tries out several possibilities, but these are just guesses. The protocol in some ways resembles TA Excerpt 7.12, segment 55, and there is a memory trace of *yesterday* being an adverb. TAS5 is relying on memory to give the correct answer to her quest. Only when faced with failure does TAS5 engage in a strategy to make use of her conceptual knowledge “*it tells me something about time*” (segment 80). However, this is semantic information, which TAS5 cannot make useful in the present situation, and it demonstrates how fragmentary her knowledge is in that this representation cannot be linked to the analysis. The situation in TA Excerpt 7.14 allows the activation of conceptual knowledge, and therefore TAS3 is able to exert control over her knowledge where TAS5 is not.

TAS5 has a representation in which factuality is not explicit, and therefore it cannot be used with causal efficacy. Categorisation tasks may be influenced by implicit knowledge from episodic instances, but it will depend on the strength of the similarity factors between the item in question and the prior examples (Neal & Hesketh, 1997:26). Some researchers (Knowlton et al., 1992) maintain that categorisation depends on abstract knowledge rather than association with prior examples. Implicit knowledge is modular (Fodor, 1983), and it cannot easily be transferred to other areas because it is not knowledge that can be scrutinised, discussed or described. “Implicit learning tends to be associated with observation and memorization conditions rather than deliberate hypothesis testing” (Dienes and Berry, 1997:5). This is not the same as saying that it is not accessible. As discussed above, all knowledge is accessible in some way, or it would not be knowledge, but implicit or procedural knowledge is only accessible under certain circumstances and cannot be controlled and recalled at random. Like explicit knowledge it can be activated intentionally but only to carry out some task to which the

procedure is attached. The holder of the knowledge remains unaware of the processes since it is not possible to form a second-order thought about implicit knowledge. Consequently, implicit knowledge cannot be used to monitor performance, nor can it be easily modified.

Performance which rests on implicit or procedural knowledge which happens to be erroneous is difficult to correct or modify because there is no conscious awareness of the knowledge content. This is present only in content explicitness. There needs to be explicit knowledge in order to correct and control representations, and subsequently to enhance mastery, becomes evident especially when it is absent. The protocols can assist in the effort to ascertain to which degree the instruction is successful, and how far the acquisition process is able to proceed in the course of the instruction programme.

7.4.4 The differential effects of the treatment

Performance is influenced by explicit as well as implicit knowledge, which means that some students may be able to analyse a simple sentence such as the experimental sentences consisting of subject, predicator, object, adverbial (SVOA), or subject, predicator, subject complement (SVC) without being in possession of full, explicit knowledge of the involved concepts and linguistic terms. In most cases, as has been illustrated above, the protocols will contain revealing clues and cues which will justify inferences as to the fragmentary nature of the knowledge, despite an eventual successful analysis of the experimental sentence. Furthermore, the analysis of the TA protocols indicated that the treatment had a differential effect on the various experimental groups with regard to task completion rate and the processing mode. From the quantitative assessment methods emerged additional differences which were very informative with regard to the epistemological and pedagogical aspects of the experiment. These findings will be examined in the following sections. The basis of the description will be the TA pre-test syllabus scores (see Table 7.4) which form the hierarchy of quartiles against which the TA post-test achievements (see Table 7.5) will be put into perspective.

Table 7.4
Think-aloud subjects pre-test ranking for syllabus

	Subject	Treatment group	Point score (max 16)
1.	TAS3	V	11.20
2.	TAS4	V	9.82
3.	TAS8	V	9.15
4.	TAS9	NV	8.13
5.	TAS12	NV	7.91
6.	TAS10	NV	7.40
7.	TAS15	NV	7.32
8.	TAS2	V	5.62
9.	TAS6	V	5.33
10.	TAS5	V	4.32
11.	TAS7	V	4.30
12.	TAS13	NV	3.79
13.	TAS14	NV	3.05
14.	TAS16	NV	1.55
15.	TAS1	V	1.55
16.	TAS11	NV	0.00

Table 7.5
Think-aloud subjects post-test ranking syllabus and exam

	Subject	Treatment group	Point score (max 16)	Ranking change	Exam ¹²⁰ (max 13)
1.	TAS3	V	14.78	same	11
2.	TAS9	NV	13.75	up 2	10
3.	TAS5	V	13.69	up 7	9
4.	TAS6	V	12.41	up 5	9
5.	TAS10	NV	11.91	up 1	7
6.	TAS8	V	11.90	down 3	9
7.	TAS14	NV	9.27	up 6	8
8.	TAS1	V	8.08	up 7	9
9.	TAS15	NV	7.99	down 2	6
10.	TAS2	V	7.40	down 2	*
11.	TAS13	NV	5.77	up 1	5
12.	TAS7	V	2.96	down 1	*
13.	TAS11	NV	2.90	up 3	*
14.	*TAS4	V	*	*	*
15.	*TAS12	NV	*	*	*
16.	*TAS16	NV	*	*	*

*not available (dropout)

¹²⁰ The exam took place six month after the experiment ended. The exam scale consists of the following marks: 13, 11, 10, 9, 8, 7, 6, 5, 03, 00. Pass is from 6 and up.

The ranking table for the TA post-test results includes the subsequent exam marks achieved by the TA subjects in their regular grammar exam, in which sentence analysis is an integral part, and the requirements are that the sentence analysis section has to score at least a pass (i.e. 6) in order for the exam as such to be passed.

7.4.4.1 The high-achievers

The qualitative study included subjects representing the top, middle, and bottom of the scoring tables. The high-achievers from both the VISL and the NON-VISL comprise subjects who from the start had some knowledge of the subject-matter, that is, all four subjects in the top quartile were above chance level. These subjects, who are able to obtain a high level of explicit knowledge, show a correspondingly high level of control, which allows them to engage in productive problem-solving, which in turn is conducive to learning. The following post-treatment protocols belong to the two highest achieving students in the VISL (TAS3, TA Excerpt 7.16), and NON-VISL (TAS9, TA Excerpt 7.17) groups, respectively.

TA Excerpt 7.16

TAS3 (VISL). POST-Sentence 1.

-----+
 POST-Sentence 1: *Jill sold the car today*

Segment	Transcription
01	<i>Jill</i> , det er subject [<i>Jill</i> is subject]
02	og det er et noun
03	... (n appears a bit delayed, and she doesn't see it)
04	måske et pronoun (sees the n)
05	...
06	og <i>sold</i> det er predicator,
07	og det er verb
08	og hun solgte bilen [and she sold the car]
09	og det er object [and that's object]
10	direct object
11	og det er en group [and it's a group]
12	...
13	og <i>the</i> det er dependent til <i>car</i>
14	og det er article
15	og så er <i>car</i> head i den gruppe [and then <i>car</i> is head in that group]
16	og a noun [and a noun]
17	nej [no]

18 ...
 PROMPT
 19 [yes]
 20 jeg ved bare ikke hvad det skulle være, hvis det ikke er
 [I don't know what it is supposed to be if it is not]
 21 ...
 22 jeg synes altså at en bil det er et navneord [I do think that car
 is a noun]
 23 ...
 24 men jeg kan da godt prøve med noget andet [but I can try
 something else]
 25 ...(tries noun again)
 26 nå det vil den gerne have det [well, that it is willing to
 accept]
 27 ok
 28 man skal åbenbart ikke give sig [one should stick to one's
 guns]
 29 og så *today* [and then *today*]
 30 det har jeg [I have that]
 31 ...
 32 vil jeg sige er adverbial [I put down as adverbial]
 33 den skal lige have det et par gange [it just needs to be told
 more than once]
 34 og [and]
 35 eh...
 36 det er også noun [also it is a noun]
 37 nej [no]
 38 ...
 39 det er måske adverb [it could be and adverb]
 40 ja [yes]
 41 og så er jeg færdig med den [and I have finished]

-----+

TA Excerpt 7.17

TAS9 (NON-VISL). POST-Sentence 1.

-----+
 POST-Sentence 1: *Jill sold the car today*

Segment	Transcription
[01-11]	
12	and subject
13	that's a noun
14	<i>sold</i> predicator verb
15	and object (writes O- g)
16	and <i>today</i> (writes A- adv)

17 *the car the* is dependent
 18 and *car* is head and noun
 19 yeah
 20 that's right

The two protocols are prototypical of their two respective experimental groups in that VISL protocols tend to be longer than NON-VISL protocols. The VISL courseware necessitates a certain structure and sequence of events, and this takes time. The protocols reflect the fact that the computer is slower than the human mind and not as dexterous. Length thus becomes an inherent feature in the VISL protocols under the given premise, since it is that the subjects talk aloud all the time while carrying out the task at hand. The TA Excerpts 7.16 and 7.17 illustrate the qualitative difference that the two experimental conditions give rise to, which is that the VISL subjects are forced to reflect on their own errors, e.g. segments 36-40: “*det er også noun* [also it is a noun], nej [no], ..., *det er måske adverb* [it could be and adverb], ja [yes]”. VISL subjects are given the opportunity to reflect on their own knowledge, e.g. segments 15-25: “og så er *car head* i den gruppe [and then *car* is head in that group], og a noun [and a noun], nej [no], ..., ja [yes], jeg ved bare ikke hvad det skulle være, hvis det ikke er [I don't know what it is supposed to be if it is not], ..., jeg synes altså at en bil det er et navneord [I do think that car is a noun], ..., men jeg kan da godt prøve med noget andet [but I can try something else], ... (tries noun again)”, as well as the subject matter, e.g. segments 08-10: “og hun solgte bilen [and she sold *the car*], og det er object [and that's object], direct object”.

The interaction with the computer and the function of the courseware give occasion reflection in cases where the response of the computer is slow or unexpected, e.g. segments 26-33: “nå det vil den gerne have det [well, that it is willing to accept], ok, man skal åbenbart ikke give sig [one should stick to one's guns], og så *today* [and then *today*], det har jeg [I have that], ..., vil jeg sige er adverbial [I put that down as adverbial], den skal lige have det et par gange [it just needs to be told more than once]”.

The NON-VISL protocol is short in terms of its number of segments, and this indicates a level of cognitive processing which is very fast, if not automatic. The time span over

which the protocol expands is necessarily very short as the fastness of the execution of the task leaves little time for verbalisation. The action description accompanying the verbal protocol documents how the subject carries out two actions, but the subjects only finds time to verbalise one act, e.g. segments 15-16 “and object (writes O- g) and *today* (writes A- adv)”. The two subjects whose protocols have been cited here are two out of four subjects in the top quartile. By the time of the post-test, one (VISL) subject has dropped out, and one (VISL) drops to the middle group (see 7.4.4.2 for further comments). This drop, however, is only in ranking as the subject in real terms moves from 9.15 to 11.90 score points. The two subjects whose protocols are cited above are the two highest scoring subjects in the post-test.

All four subjects in the top TA pre-test quartile have scores above chance level, which for the syllabus test section is 8 points (max.16). The two highest scoring subjects in the TA pre-test syllabus section subsequently scored the two highest exam marks. And all the TA VISL subjects received above average exam marks (for the actual exam marks, see Table 7.5; for means see Table 7.6).

Table 7.6
Key figures for TA Subjects

Category	VISL	NON-VISL
Exam means	9.4	7.2
Dropouts by post	1	2
Dropouts by exam	3	3
Mean gain points	6.08	3.72
No. of subjects	3	1
ranking change more than +/- 3	(+7,+7,+5)	(+6)

One should be cautious, however, to place too much significance on the numeric achievements of the TA subjects since their number is relatively small, and the purpose of this chapter is not to say anything about the quantitative achievements per se of the VISL group subjects in relation to the NON-VISL subjects, but rather to find evidence in the protocols about the differences in cognitive processing between the two experimental groups. For that purpose the numeric indications can be used for

comparison and clarity as they put the TA results into perspective, but no more than that.

7.4.4.2 The middle group

The second quartile (pre-test) scores lie between 49% and 35%, and the quartile comprises 3 NON-VISL subjects and 1 VISL-subject, and none of them do particularly well. One NON-VISL subject drops out before the post-test, the others remain basically the same. Only one (NON-VISL) subject manages to attain post-test scores above chance level.

Three subjects dropped out between pre-test and post-test therefore any change in ranking which can be said to constitute significant change will have to be more than three places up or down the hierarchy. The three remaining subjects (2 NON-VISL, 1 VISL) in the second quartile experience little change from pre-test to post-test. TA Excerpts 7.18 and 7.19 may throw some light on the reasons for this.

TA Excerpt 7.18

TAS10 (NON-VISL). POST-Sentence 2.

+-----+	
POST-Sentence 2: <i>It is a very good book</i>	

Segment	Transcription

01	<i>is</i> det er predicator, verbum [<i>is</i> that is predicator, verb]
02	<i>It</i> er subjektet, som er pronoun [<i>It</i> is subject, which is pronoun]
03	fordi det står i stedet for noget andet [because it is there in stead of something else]
04	ahm...
05	<i>very good book</i> det er en objekt group, hvor <i>very good book</i> [<i>very good book</i> that's an object group where <i>very good book</i> that's]
06	...ahmm....
07	head det må være <i>book</i> , som noun [head that must be <i>book</i>]
08	og dependent er article <i>a</i> [and dependent that is article <i>a</i>]
09	ligemeget (skriver ingenting) [no matter (doesn't write anything)]
10	og dependent er [and dependent that is]
11	...
12	<i>very</i> som er adjektiv [<i>very</i> which is adjective]

13 og dependent igen som er adjektiv, *good* [and dependent
again which is adjective *good*]
14 ja [yes]

The difference in the quality of representations comes out in the contrast between TA Excerpt 7.18, segments 02-03 on the one hand and segment 12 on the other hand. We can see that the knowledge of *It* as a pronoun has been properly caused, unlike the classification of *very* as an adjective. The inference from segments 02-03 is that the proposition “*It* is a pronoun” constitutes *content* explicitness. Segment 12 could be interpreted as an instance of fast automatic-like processing. Two indicators make this doubtful; first, the hesitation in segment 11, which compounds the hesitation in segment 06, and second, the fact that it is wrong. The first indication that TAS10 has problems with the sentence is evident as early as segment 04 when the hesitation indicates uncertainty of the whole sentence.

The protocol ends because the subject is of the opinion that he has completed his task, but the analysis of the sentence is incomplete and incorrect in essential aspects. The complement group *a very good book* is taken to be an object, and *very good* is not seen as a subgroup but as two individual modifiers of *book* with *very* taken for an adjective. This subject manages to achieve a post-test result above the 50% level, but in the final exam he earns a mark just below average, namely 7.

TA Excerpt 7.19

TAS15. POST-Sentence 1.

POST-Sentence 1: *Jill sold the car today*

Segment	Transcription
01	<i>sold</i> det er predicator [<i>sold</i> that’s predicator]
02	<i>Jill</i> , det er subject [<i>Jill</i> that’s subject]
03	...
04	(sighs)
05	<i>the car</i>
06	ehh...
07	it’s a group
08	but..eh...
09	and object I think

10	<i>today</i> hmm...
11	det kan jeg ikke huske [I don't remember]
12	...eh...
13	nå [well]
14	jeg går lige til den næste [I'll just take the next one] [goes on to do sentence 2 and analyses sentence 2] [returns to sentence 1]
15	jeg kan altså ikke huske <i>today</i> [I really don't remember <i>today</i>]
16	så jeg tror, jeg bliver nødt til at stoppe [so I think I have to quit]

-----+

This subject (TAS15) has very little explicit knowledge, and her retrieval strategy is to rely on memory retrieval, or to attempt retrieval from memory of a specific instance of the word *today* (segments 17-21). When this fails there is no ability to use problem-solving strategies, a circumstance which indicates a lack of explicit knowledge. The subject's knowledge of the object is not fully explicit either (segments 03-09) which the subject makes clear by pausing (segments 0, 04, 07, 08), and by stating, 'I think' (segment 09). These words leave little doubt that this proposition is not attitude explicit, i.e. this is not knowledge but supposition. A comparison of the TA Excerpts 7.18 and 7.19 and the difference in ability between these subjects make it easier to understand why TAS10 is able to retain the position in the middle of the spectrum while TAS 15 is not able to move above chance level in the post-test score.

7.4.4.3 The low-achievers

The low-achieving group consists of those eight subjects (4 VISL and 4 NON-VISL) who managed between nil and one-third of the score items in the pre-test. Presumably, this is a weakly founded group of students, and therefore it would be valuable to know something about the possibilities of bettering the attainment level of this group of students.

Of the original sixteen think-aloud subjects, eight, i.e. half the group, belong to this group. By the time of the exam, which took place six months after this experiment had ended, three had dropped out, one failed to pass the exam, and one had passed with a minimum, i.e. six. The drop-out rate is no higher in this group than in the high-

achieving group, though. A remarkable fact about this group is that only one NON-VISL subject passed the exam from this group, in contrast to the VISL subjects of whom three passed (one dropped out), and these three all improved their ranking by seven (2 subjects) and five (1 subject) places, respectively. Two VISL subjects moved into the upper quartile, and TA Excerpt 7.20 below belongs to TAS5, who moved up seven places in ranking to the third highest position in the post-test, and whose score improved from 4.32 to 13.69 score points.

This protocol is very similar to the protocols of the two highest scoring subjects, who were in the upper quartile from the beginning. The features that these protocols have in common are the automatic-like speed and brevity, and the assurance of the subject in her own knowledge. These features allow for an inference that explicit knowledge of the proposition is held by the cognitive system.

TA Excerpt 7.20

TAS5 (VISL). POST-Sentence 1.

+-----+ POST-Sentence 1: <i>Jill sold the car today</i> +-----+	
Segment	Transcription
+-----+	
[01-03]	
04	and this (<i>sold</i>) is the predicator
05	it's a verb
06	and the subject (also picks noun)
07	and then I group these two (<i>the car</i>)
08	ja [yes]
09	and it should be the object
10	and it's a group
11	this (<i>car</i>) is the head
12	it's a noun
13	it's dependent (also picks article)
14	eh..
15	this (<i>today</i>) must be adverbial
16	and I think it's an adverb
+-----+	

The only hesitation appears in segment 15 and pertains to the form of *today* of which the statement in segment 16 '*I think*' makes it apparent that attitude is not explicit - there is no explicit *factuality* of the proposition, which cannot be established as a fact.

Typically, and in parallel with the protocols of other upper quartile excerpts, the subject is so fast that the verbalisation cannot keep up with the action of the mind and hand (segments 06 and 13) so that two things happen but only one of them is verbally encoded.

7.4.5 Summary of VISL and NON-VISL differences

Performance is influenced by explicit as well as implicit knowledge, which means that some students may be able to analyse a simple sentence such as the experimental sentences consisting of subject, predicator, object, adverbial (SVOA), or subject, predicator, subject complement (SVC) without being in possession of full, explicit knowledge of the involved concepts and linguistic terms. In most cases, as has been illustrated above, the protocols will contain revealing clues and cues which will justify inferences as to the fragmentary nature of the knowledge despite an eventual successful analysis of the experimental sentence. As to the result of the analysis of the experimental sentences in the pre-treatment protocols, there are four unfinished sentences in the VISL protocols, two SVOA and two SVC, concentrated on two subjects. The VISL subjects realised that the sentences were unfinished, of course, because the VISL courseware left no doubt about that, but the subjects decided that they could not finish successfully and gave up. In the NON-VISL protocols there are no examples of students giving up, but their results are poorer than those of the VISL subjects since none of the analysed sentences were completely correct in all details. The obvious difference is that the NON-VISL subjects, unlike the VISL subjects, had no feedback which could tell them that their performance was incomplete or incorrect.

The careful reading of the protocols seems to provide evidence that the lack of feedback in itself provides the background for assuming a differential effect of the VISL courseware. During the treatment period efforts were made to offer feedback to the NON-VISL groups. Regardless of the quality of the feedback in the NON-VISL classroom, there are some features of the VISL courseware feedback which – given the number of subjects present in a group – it is impossible to imitate, namely the individualised feedback of the VISL courseware, which is prompt, and which cannot be ignored or pushed aside. In many cases it will not be possible to continue without

rectifying incorrect input. The students are compelled to devote attention to their own errors, and as the TA protocols demonstrate, they are given the time and occasion to evaluate their own knowledge. This awareness of one's own performance and the focussing on problem areas appears an important factor which enhances learning.

The second round of TA task executions and recordings, which took place after ten weeks of instruction, shows that with regard to the SVOA sentence, the VISL and NON-VISL subjects are equally good except in the category of adverbials. One NON-VISL subject does manage to identify this constituent. The SVC sentence gives a completely different distribution, in that all but one of the NON-VISL subjects cannot identify the subject complement; in contrast all but one of the VISL subjects can. These results may be said to corroborate the interpretation that the nature and timing of the feedback play a role in the learning process, and thus constitute an elementary difference between the two methods at a functional level.

The learning processes and the resulting quality of knowledge as exposed through the application of the method based on Dienes and Perner's theory support the functional findings of the differential effects observable in VISL and NON-VISL subjects. In general, the VISL subjects achieved a higher or more stable level of explicit knowledge. This is especially pronounced in the lower middle/upper bottom segments of students. It appears that the VISL courseware has the ability to induce a cognitive approach to the subject-matter and its apprehension, a feature which allows the students just below the middle of the spectrum to construct their own knowledge in a way which may often result in low-achieving students becoming high-achievers. The same results cannot be seen for NON-VISL students at a comparable level; NON-VISL students remain basically at the same achievement level. The exam marks achieved by the students in the two groups also vary, with the VISL subjects (mean 9.4) clearly outperforming the NON-VISL subjects (mean 7.2). Of the 8 VISL subjects who participated in the TA experiment, 5 passed the exam; of the 8 NON-VISL subjects who participated in the TA experiment, four passed the exam.

Due to the small number of TA subjects who completed the whole run of pre-test, post-test, and exam, these results must not, of course, be given more weight than they deserve. However, the results of the analysis of the think-aloud protocols, and the application of the method based on the theory of implicit and explicit knowledge, have a qualitative validity which is independent of the number of participants. The sentence analysis results achieved by the participating TA subjects as well as the TA protocol content point in the same direction, which is that the VISL courseware has a functionality and effect which it is difficult to achieve in the traditional classroom without a heavy toll on teaching resources. From a learning perspective, the analysis has confirmed that the subject-matter is mastered better by students who either possess an analytic learning style, or by those who manage to acquire it during and through the instruction given. This is another reason the VISL courseware can facilitate the desired process.

PART FOUR

8 DISCUSSION AND CONCLUSIONS

The combination of the quantitative method and the qualitative method has turned out to provide a fuller answer to the research questions than either method would have been able to give separately although the individual research questions may be directed at and answered by one or the other type of methodological inquiry.

8.1 Discussion of the findings for the research questions

The first research question, which was: '*Is VISL as good as traditional classroom instruction?*', could be answered in the affirmative. The VISL subjects did as well as the NON-VISL subjects, generally speaking. Compared within each cohort, the English VISL outperformed the English NON-VISL by average post-test success rates (total) of 71.4 % versus 68.2%. For the Cand. Negot. cohort the average post-test success rate (total) for the VISL group was 73.1% versus 66.7% for the NON-VISL group. The VISL subjects performed not only on a par with the NON-VISL subjects, but slightly better. When the tests were screened for only the syllabus section items, the difference in post success rates remained manifest. The English VISL group achieved a post-test syllabus success rate of 70.6% against 65.8% for the NON-VISL post-test success rates. The Cand. Negot. cohort syllabus post-test success rate for the VISL group was 73.2 % against 66.7% for the NON-VISL group. The syllabus post-test success rates may seem to be the more relevant result than the total post success rates since the syllabus section of test items comprises the same type of knowledge contained in the teaching content, and thus, perhaps, a more certain gauge of the results of the treatment than the total figure which includes the non-syllabus items.

Gain rates revealed a slightly modified distribution in results. For the English cohort the difference was slightly in favour of the NON-VISL group. This group had very high pre-test levels, especially the BASIS group, and this meant that the gains were smaller

despite the higher success rate for the VISL group, but the results are quite close, i.e. less than one percentage point. For the Cand. Negot. cohort, the VISL group outperformed the NON-VISL group with respect to gains also.

The difference in results from the two experimental treatments indicates that the computer is a valid instructional; it performs on a par with traditional classroom instruction, and in some cases it produces better results. With regard to success rates – measured in terms of student results in their subsequent exams – the computer method gives results above the traditional instruction for both student cohorts, i.e. for Cand. Negot. as well as English students, but only significantly so for the Cand. Negot. students.

Embedded in the presentation above is the answer to the first half of the second research question, which was: ‘*Is VISL equally good for English students and Cand. Negot. students?*’, in that it outlines a difference in the Cand. Negot. and the English cohort. The Cand. Negot. cohort benefits more than the English cohort from the VISL treatment. The syllabus post-test success rate for English VISL is 70.7%, which is somewhat lower than the Cand. Negot. syllabus post-test success rate of 73.2%, and when it comes to gains rates the difference is even more marked. As for the syllabus gains, the English VISL gain is 28.9% versus the Cand. Negot. VISL syllabus gain of 39.0%. The Cand. Negot. cohort had a lower entrance level than the English cohort so there was more room for improvement, which is what the difference in the gains figures indicates.

The really telling result, however, is the fact that the Cand. Negot. VISL group started out at a lower level than the English VISL group, and yet the group managed to achieve a post-test success rate level which was higher than that of the English VISL group. The difference in syllabus pre-test levels was 7.6 percentage points in the English VISL group’s favour. At the syllabus post-test level the difference was 2.6 percentage points in the Cand. Negot. group’s favour.

The third research question was: *'Is VISL equally good for all achievement groups, i.e. high, middle and low?'*, and the answer to this was sought answered by means of the quantitative as well as the qualitative method, including the think-aloud data. Think-aloud protocols provide an insight into the cognitive processes as they play themselves out in a fashion which can be likened to a stream of consciousness. Theoretically, verbalisation is possible of all conscious thought. Concurrent reports enable us to observe the actions, thoughts, and approaches in which the students are engaged without being filtered significantly by deliberation, speculation, or memory capacity. The participants in the think-aloud study were selected in such a way that the whole spectrum of competence levels was represented, and from comparing the pre and post protocols from the various ability groups a pattern emerged. The upper group of students became almost automatic in their ability to analyse sentences. The qualities of the linguistic concepts were recognised and structures mastered. This was a development which took place in both the VISL and the NON-VISL groups. The upper-middle group behaved in basically the same way as the high-achieving group of participants. For the low-achievers there appeared to be no difference in behaviour or achievement between the VISL and the NON-VISL groups.

The interesting development appeared in the lower end of the middle-achievers. The think-aloud VISL subjects¹²¹ progressed beyond expectation, as they were able to break their ranking and move beyond the expected gain of their group. This did not happen in the think-aloud NON-VISL group. The low-middle think-aloud VISL subjects moved up a category and into the upper-middle category. But, the results need to be interpreted with caution as the number of individuals in the think-aloud study was relatively small, and therefore this finding needs to be related to the results of the quantitative study. In this light, it is therefore interesting to see the boxplots of the large quantitative study, which clearly support the findings from the qualitative study (see especially Figure 4.7). When the boxplots from the VISL and the NON-VISL groups are compared, it is evident that the floor is raised in the VISL groups, which is to say that the mean is raised. The VISL method appears to include more students in the groups of well-

¹²¹ The results referred to here are the ones relating to the think-aloud participant and not the results from the large quantitative study. The achievements of the TA subjects in test results are in line with the findings of the quantitative study, and hence the two studies inform each other.

performing students. The question that needs to be asked following this discovery is whether it is possible to discern any difference in the student behaviour which could explain such a difference between the VISL and the NON-VISL method. The think-aloud protocols reveal some differences which might indicate at least one explanation. The high- and upper-middle achievers all employed a very structured approach to the task. This was not the case with the lower-middle and the low-achievers, who were erratic and random in their task performance, and it appears that the VISL courseware was able to induce a better and more structured approach to the task of analysis in the originally lower-middle performers with.

Dienes and Perner's theory of implicit and explicit knowledge holds that the two types of knowledge coexist in a hierarchy, which again means that knowledge explicitness can be present in varying degree until full explicitness is reached. This theory was applied to the think-aloud data and operationalised in a set of interpretations of the verbal expression of the cognitive processes. By this method of analysis, and through the systematised interpretations of the expressions of explicit knowledge, it became clear that VISL was able to increase the level of explicitness and awareness in the students at the lower end of the spectrum to a degree which was not quite matched by the traditional classroom instructional method.

A combination of qualitative and quantitative methods was also applied in order to answer the fourth research question: '*Do students achieve full explicitness of the subject matter?*' This quantitative study measured the mean success rate of the syllabus content to lie from 50.9% (Cand. Negot. BASIS group) to 73.2% (Cand. Negot. VISL group). The application of the method based on Dienes and Perner's theory qualifies the quantitative results in that the think-aloud data revealed that only the few high-achieving students reach fully explicit levels of knowledge. On the basis of the analysis of the think-aloud data in accordance with the hierarchy of explicitness developed in the Dienes and Perner model (see Table 7.3 and Figure 5.1), it becomes possible to conclude that the explicitness of knowledge rarely reaches a level at which it becomes represented as knowledge held by the student. It never reaches the level in the hierarchy which is characteristic of full explicitness, i.e. a level which makes it possible for the

student to say '*I know that it is a fact that this a noun*'. The cognitive representation for the majority of students does not reach the 'knowledge' state according to this definition, but stays at the supposition or hypothesis-testing level. It is clear from the protocols that several students reach a level at which the hypothesis-testing stage is reaching a point in the learning process which immediately precedes the level at which it becomes possible to construct such knowledge.

The main research questions were related to the effectiveness of the instructional methods in the acquisition of the metalinguistic knowledge. This was tested in the syllabus section of the test items. The students were also subjected to test items on their general second language proficiency in the non-syllabus section of test items. Here students could react intuitively to test items. The instruction was aimed at having an effect on the syllabus test items, which it did, but could there be an effect also on the non-syllabus items? The fifth research question was therefore: '*Can the syllabus instruction affect the non-syllabus results*'? The expectation was that there would be no effect in the non-syllabus section. Indeed this question was a side issue as the experiment was set to demonstrate the effects on the acquisition of the syllabus by way of the two methods applied. The fifth research question should be seen in the light of the discussion of the relationship between explicit knowledge and implicit knowledge, and any possible interface between them. Most SLA researchers adhere to the weak-interface position, and the results in the non-syllabus section were therefore first of all measured and evaluated as a particular way of relating and contrasting the syllabus results. As expected there was a significant difference in results between the syllabus and the non-syllabus section, with marked improvements in the syllabus section, but only modest changes in the non-syllabus section. From the beginning the English cohort had higher test-levels in the non-syllabus section than the Cand. Negot. cohort. This difference was maintained at the post-test level, but the Cand. Negot. cohort had the highest gains. Though there is some change in the overall level in the non-syllabus section, this could be due to other factors than the treatment in question. A closer examination of the two non-syllabus items with the highest gain rate revealed that the two test items 6 and 23, which contained tense/aspect issues, were indeed different from the other non-syllabus items. The results from all treatment groups were therefore

pooled and tested against the other non-syllabus items in groups comprising word-class items, syntax items, morphology items, and tense/aspect items. The tense/aspect items appeared to be affected more by the treatment than was the case for the other groups.

The difference in effect of the instruction on the syllabus results and the non-syllabus results will support the claim that there is no direct linkage between implicit and the explicit knowledge in the sense that metalinguistic knowledge translates into improved linguistic proficiency. The non-syllabus results did indicate that the performance was not static, which supports the belief that there may be an interface between the two where the categories of knowledge will in fact interact. To some degree, the present study was not designed to delve into that particular issue but the present findings do not preclude such a possibility. Indeed the tense/aspect findings might be taken to support a view that there is a connection of sorts.

The noticing hypothesis, as developed by Schmidt (1990), is an important part of the answer. Noticing and awareness-raising are the beginnings of the construction of new knowledge. Those particular cognitive activities constitute the interface of metalinguistic knowledge which was the object of this study, and the intuitive linguistic performance which saw only modest or no improvement except for the Cand. Negot. VISL group. Metalinguistic instruction may facilitate better and more efficient internal processing, or it might provide a frame or context which will condition subsequent patterns of learning. In a reference to VanPatten (1996) among others, N. Ellis (2005:325) claims, "Metalinguistic information connects with implicit learning, and they meet and interact in processing. It is a dynamic interface." (see also VanPatten and Cadierno, 1993). This is an interpretation which may be said to be supported by the results of this study, although these say nothing about the exact character of the quality of the interface.

8.2 Conclusions and perspectives

The results outlined above warrant a contextual discussion of the role of the computer in language instruction at the tertiary level. First, there is the issue of creating suitable

courseware. The basic language learning has already taken place when students enter university. From a learning point of view we are dealing with adults who are proficient communicators but whose needs, as dictated by the objectives of the study programmes are to improve their own proficiency and knowledge (especially with regard to correctness and variety of register), as well as to acquire a metalanguage and metalinguistic knowledge. The object of investigation in this thesis was metalanguage and metalinguistic knowledge, and to what degree this can become explicit and declarative knowledge. The investigation into the general linguistic proficiency level was a subsidiary issue, and the tests would need to be designed differently to place the main focus on this aspect of language learning. Equally, the VISL tool might not be the best suited courseware for such a purpose.

The Cand. Negot. students seem to learn better from the VISL tool than the English students, and there is no ready explanation for this. The English students generally have higher levels in both proficiency and metalinguistic knowledge so there is more room for improvement with the Cand. Negot. students. This is not the full explanation, however, as the Cand. Negot. students reach an overall higher level than English students apart from the BASIS group. The high level in the English BASIS group makes comparison with this group difficult. The two study programmes may attract different types of students, but this is a very speculative issue. The description of the subjects in Chapter 4.2 outlines a difference in entrance exams held by the students. The majority of the English students have a general academic background (i.e. "Gymnasium") either from the language line or the mathematics line. The Cand. Negot. students represent a large segment whose background is the "HHX" (business related) background. The background research of this study also looked at student differences in terms of interest in IT, gender differences, and differences in time spent abroad, and none of these factors appeared to have any influence, which was the reason why they were not included in the thesis itself.

One aspect of learning which was not included in this study is the issue of learner strategies which is a prominent study area in its own right. A focussed study of learner

strategies might provide some of the answers which it was impossible to include in the scope of this study.

A further point to make is that the overall aim of this study has been to operate as close to realities of the current study programmes as possible in order to give the study ecological validity. The learning and the instruction take place under certain given conditions, and therefore it seems relevant to study learning outcomes under exactly those conditions with a view, of course, to discovering new aspects which may or may not lead to the conclusion that the given conditions should be changed to optimise learning conditions. One difference between the English and the Cand. Negot. students that could be offered for further investigation is that the English students are very diverse in their interests, and therefore the linguistic aspects of the study programme may not be in focus to the same extent that it is for students enrolled in the Cand. Negot. programme. Cand. Negot. students are typically focussed on improving their linguistic prowess, whereas the focus of English students has a wider span of options and offers in literature and history far exceeds that of the Cand. Negot. study programme.

Awareness is a critical ingredient in learning. The instruction in the experimental treatments was aimed at raising awareness. This study has not been able to demonstrate that increased metalinguistic awareness will result in increased proficiency. The tense/aspect items did appear to be affected, but that was only when all data were pooled, not when the results were divided according to experimental groups. This underscores the conclusion that the direct link between explicit and implicit knowledge is not strong, but the results do not rule out the weak-interface position as a possibility. The central question is how much time is needed; how many instances in how many different contexts are needed for the construction of fully explicit knowledge? It also needs to be stressed that this study was not set up to investigate this issue specifically, it only emerged as a natural subsidiary issue along the way. The small effect which is seen in the non-syllabus section could be the result of the processing of the 54 experimental sentences rather than from the instruction or the acquisition of the formal aspects of language. Added to this is the large quantity of English texts the subjects read during

the first semester. Consequently, this result can at best be described as an invitation to further investigation in that it leaves the issue open.

The VISL interface has an inbuilt ability to additionally raise the user's awareness of certain aspects in the input. In addition to the nature of the feedback, this is the main difference between the traditional approach and computer assisted learning. The colour scheme of the VISL interface enhances the difference between form and function, for instance. The colour of the connecting slant lines in the building of the sentence structure, and the red colouring of words when the end level has been reached are further features. Input enhancement may be one of the reasons why VISL is as good as or better than traditional instruction. The permanence and well-ordered structure of the screen could be an additional factor in the comparison to pen and paper. The users must build the structures themselves, but once the structure is there, it is neat and permanent, unlike the often messy drawings on paper. With VISL the overview of structural levels is easier to get and maintain.

The issue of feedback also needs to be mentioned. The VISL students received immediate and individual feedback, and despite the attempts to give quick and efficient feedback to the NON-VISL students, there is no doubt that there often was a time gap. During the time between student action and teacher feedback, the student had often pursued a line of thinking and action that had to be revised, and this factor probably added to the difference between VISL and NON-VISL results.

The fact that the VISL-based instruction gives an advantage to the lower-middle range of students is encouraging, but the study provides no ready explanation for this. The think-aloud data may hold a clue to the explanation in that it became clear that the top students all worked with a structurally based approach. After the treatment it became evident that the particular lower-middle students who did well were the VISL students. These students even moved up ranks in achievement category. The difference in their task approach was that they were now able to work in a structurally based fashion. This finding is in line with the difference between the VISL interface, and the pen and paper 'do-it-yourself' method. The upper-middle and the top students have no difficulties in

providing structure and order to their task approach, but the lower-middle students are not able to do this. It appears that VISL can facilitate this ability. The really low-achieving students cannot avail themselves of this facility and the high-achieving students have no need for it. It must be concluded that the drop-out rate could be reduced and student experience of success and achievement could be enhanced if VISL was used in a more systematic and targeted manner.

It is worth noticing that when the experiment was carried out, the VISL interface was not as developed as it is now. The students using the earlier version of VISL had to scroll up and down to find the term (form or function) or function button that they wanted. The present interface presents all the available possibilities in two menu bars, which are permanently on screen. One might therefore surmise that if the experiment were carried out today the results would be even more in VISL's favour.

The application of the method based on Dienes and Perner's theory of implicit and explicit knowledge to the think-aloud data revealed that the degree of full, explicit knowledge of the subject matter was limited. This raises the question of whether the objectives of the different study programmes have been met to a satisfactory degree. The overall success rate for the BASIS group was only slightly above fifty percent, or not much more than chance level. Even for the VISL and the NON-VISL experimental groups the up-take rate was at best slightly below a mean of seventy-five per cent. Especially the low BASIS mean value is worrying since this represents the actual instructional level of the study programmes.

In itself, this level of instruction does not constitute the end result of the formal instruction, and the bulk of instructional time and content will follow in both programmes, but the subject matter investigated in this experiment forms the basis for success in the subsequent instruction. As pointed out above, awareness is important, even critical, in the learning process. It might therefore be fruitful to raise student awareness of the intentions and implications of the formal instruction, and its place in their own learning in order to enhance motivation and results.

Students of today have been educated with the computer as a natural learning vehicle, and the perspective is that they expect the computer to have the same role at the university level that it has played in their pre-university education. The situation today is that very little suitable courseware is available of a quality that can meet the proficiency level of Danish students of English. The majority of disciplines and areas of study in the university study programmes are unsuited to transferral to computer-based instruction, therefore it is of importance that material is developed in the areas which lend themselves to the medium. Linguistic skills meet this central criterion, and certainly the metalinguistic syllabus would be an obvious choice for students to train on the computer. Blended learning in a combination of CALL and traditional classroom instruction - and discussion - could be a viable way forward. Furthermore, the research on human-computer interaction from a learning perspective is limited, and it would be interesting to pursue this aspect in order to develop the pedagogical aspect of CALL, a somewhat overlooked aspect. The VISL courseware constitutes one of the few tools of quality with a potential for facilitating language learning at the tertiary level.

English summary

This thesis is a synthesis of interest and investigation in three interrelated areas: Second language acquisition, computer assisted language learning, and the learning processes involved. At the heart of the thesis is a sincere interest in the cognitive processes that constitute learning. The processes are interesting in their own right, but the transformation from theory to application, is not less so. This thesis is devoted to examining aspects of learning which are found to constitute an interface between the computer as a learning tool, the metalinguistic aspects of learning a second language, and the processes involved, as far as they can be discovered and interpreted from the data collected in the quantitative as well as the qualitative studies.

The thesis comprises two studies: a quantitative study which examines the differences in learning between three experimental treatments, and a qualitative study which examines the learning processes. The thesis is structured in four parts. Part one contains the introduction and the outline and rationale for the research questions. Part two contains the literature reviews of the research areas connected to the quantitative study, i.e. second language acquisition (SLA) and computer assisted language learning (CALL), and the quantitative study itself including its results on the experimental treatments: Visual Interactive Syntax Learning (VISL), NON-VISL (the non-computer treatment) and BASIS (regular classes). Part three consists of the literature reviews of the research areas connected to the qualitative study, i.e. implicit and explicit learning and knowledge and introspection, and the study itself including its results. Part four consists of the discussion of the findings pertaining to the research questions, and the conclusions and perspectives.

Part one focuses on the research questions and their relevance to the project. The five research questions are:

1. Is VISL as good as traditional classroom instruction?
2. Is VISL equally good for English students and Cand. Negot. students?
3. Is VISL equally good for all achievement groups, i.e. high, middle, and low achievers?
4. Do students achieve full explicitness of the subject matter?

5. Can the syllabus instruction affect the non-syllabus results?

The VISL – Visual Interactive Syntax Learning - tool developed by researchers at Odense University (now University of Southern Denmark) is one of the few tools which are suitable for advanced students at the tertiary level and as such it is a natural choice for an investigation of the computer as a learning tool in language learning.

The three first research questions clearly relate to the efficiency of VISL in comparison to the traditional classroom instruction but also to the question of whether the interaction with the computer results in a different pattern of learning than the traditional methods.

The last two research questions concern the quality of the learning and the issue of explicitness and implicitness in learning and knowledge. The VISL tool is tailored to the metalinguistic syllabus at the English study programme and the Cand. Negot. study programme. The metalinguistic syllabus is an integral part of academic linguistic studies for the advanced students of these two programmes and has a place in its own right. The issues in research questions four and five concern the nature of cognitive processes in the learning of the subject matter, and the relation, if any, between the metalinguistic knowledge and students' linguistic proficiency. The studies into the experimental treatments and their effects embody the three major strands of the thesis: CALL, SLA and cognitive processes.

Part two outlines the aim and the methodology of the quantitative study in the framework of SLA and CALL. The research in CALL has been technology-driven to a large extent. However, this is now changing and the area is struggling to find its own research paradigm. The turn of CALL interests towards more pedagogical aspects makes the field look towards SLA. After all, CALL and SLA share some of the same pedagogical interests and methods.

In many ways, VISL places itself at the heart of current trends and discussions in CALL and SLA. VISL is a grammatical tool, and the issue of the role of grammar is central to the fundamental theoretical framework in SLA. The overview chapter on SLA outlines

three fundamental beliefs about the relation between metalinguistic (explicit) knowledge, implicit knowledge, and learning. It is the latter which is thought to be the basis of linguistic performance and fluency. The no-interface position states that there is no connection between the two types of knowledge. According to Krashen (1981, 1985) there are two different and unconnected types of knowledge, one is learned (explicit, metalinguistic) and one is acquired (implicit). The strong interface position claims that there is free interchange and conversion between one and the other. The weak- interface position, which is held by many present-day researchers in the field, states that there can be a connection under certain given circumstances. Such a circumstance is found in attention and awareness. If the students can focus attention and become aware of linguistic features it facilitate acquisition.

VISL has the potential to help raise awareness through its interface which supports input enhancement features through the colour scheme and structural tree-building facility. VISL also incorporates immediate feedback which raises student awareness of the quality of their input. The quantitative study investigates the difference in student-based learning results between the three groups, VISL, NON-VISL, and BASIS, distributed over the English cohort and the Cand. Negot. cohort. The study was conducted in a pre-test, post-test design. The treatment period was ten weeks and 107 subjects participated. The tests contained a section of syllabus items and a section of non-syllabus items. As expected, the syllabus items revealed that the entrance level knowledge of the subject-matter was low. The non-syllabus items, which subjects could respond to intuitively showed a much higher entrance level. In other words the explicit metalinguistic knowledge was low whereas the general proficiency as measured by the non-syllabus items was much higher. The detailed analysis of the student-based results also revealed that the VISL treatment was able to improve the achievements of the lower achieving students to a higher degree than the other treatments.

The treatment affected the syllabus and the non-syllabus items differentially. There was also a differential effect for the three treatments. The treatment had a good effect for the syllabus items, but very little for the non-syllabus items. Of the three treatment groups, the best results were achieved by the VISL and NON-VISL groups and the BASIS

groups were the least affected by the treatment. It should be noted, though, that the English BASIS group was so high-performing from the beginning that it is difficult to use it for comparison. The difference between the VISL and the NON-VISL groups was to the advantage of the VISL groups. Especially the Cand. Negot. VISL group was very high performing. The non-syllabus section showed little change from the initial pre-test results to the post-test result. Some improvement did take place, but the contrast to the syllabus improvement was marked. The treatment was at least four times as effectual for the syllabus section as it was for the non-syllabus section. An examination of the results based on an analysis of the discrete items showed that the non-syllabus improvement was highest for the tense/aspect items. The efficacy of the treatment judged by discrete items showed a good effect on the syllabus items, but poor effect on the non-syllabus items whose subject-matter content was the same. Syntax is a case in point. The results appear to confirm the weak interface hypothesis, with the possible exception of tense/aspect, of limited connection between the two types of knowledge.

Part three contains the qualitative study which was designed to supplement and support the quantitative study results. The qualitative study therefore represents the application of a method which allows a glimpse into the cognitive processes of learning. The raw material for the study was obtained through think-aloud protocols, and the theoretical foundation of the subsequent analysis of these protocol was found in Dienes and Perner's theory of implicit and explicit knowledge (1999). The theory is operationalisable as a research method because it presents a very clear hierarchy of explicit knowledge and the possible combinations with implicit knowledge. The research on explicitness and implicitness and their inherent qualities is discussed in this part of the thesis, especially the link to consciousness and how consciousness should be defined. There is major disagreement among the scholars of the field. The literature review pointed to the possibility that some of the disagreements could be due to the lack of common terms and definitions of the field.

According to Dienes and Perner's theory, fully explicit knowledge encompasses three structures, namely *content*, *attitude*, and *self*. Each of these levels has an internal structure with several layers of explicitness. A central point of explicitness is

'factuality', and whether a given piece of information is seen and perceived as fact. In the exposition of their theory Dienes and Perner give the theoretical and empirical research background on the basis of which they have developed their theory. One important piece of evidence for the qualitative study of the present thesis is the way they demonstrate that it is credible to link explicitness with verbalisation and declarative knowledge, on the one hand, while, on the other hand, it is possible to link implicitness with tacit and procedural knowledge.

The use of verbal protocols can take many forms, and each form has advantages and disadvantages, but a common feature is that it is one of the only available means for insight directly into the cognitive processes of learning. The method is not favoured by all, and results have to be evaluated carefully. For the present study the concurrent think-aloud type of protocol was chosen. Ericsson and Simon (1993) describe in detail several different types of protocols, and the one chosen for the present study is of the type where the subjects are asked to verbalise their thoughts as they arise. This type of protocol was chosen in order to avoid interference in the cognitive process due to imprecise recall or attempts to speculate.

Verbal protocols were selected from students selected on the basis of their pre-test results in order to have the whole spectrum of achievements represented. The subjects in the think-aloud study were not enrolled in the quantitative study but they were given the same treatment as the students participating in the quantitative study, but without the BASIS groups. The verbalisations were analysed according to the hierarchy developed by Dienes and Perner. The area of interest of the present thesis was the degree to which the metalinguistic knowledge becomes explicit and to which extent the syllabus instruction is able to affect the non-syllabus results. The quantitative results measured the effect of the instruction. The point of the analysis of the protocols was to gain insight into the mental processes involved and perhaps attain answers which could help us understand the overall results as measured by the statistics.

It turned out that few students achieved full explicitness of the subject-matter. However, the analysis of the protocols revealed an interesting variation among the students. The

high-achieving students from the pre-test remained high-achieving and a couple of them became almost automatic. However, the explicitness of their knowledge could be observed when they ran into difficulties. It is one of the characteristics of explicit knowledge that it is controllable and analysed. This could be observed when students had to recall the knowledge and apply it in an analytic fashion in order to complete the task. The lower-middle students from the pre-test achieved differential results depending on whether they were given the VISL or NON-VISL treatment. It turned out the VISL treatment was able to help the students become more analytic and structured in their approach and the VISL subjects were able to break their pre-test ranking and move up several places in the ranking in contrast to the NON-VISL subjects, who remained at the low end of the scale of achievement. This new information could be the one factor which might explain why the VISL treatment measured in the quantitative study could raise the low-achievers to a higher degree than the other treatments.

The perspectives of the two studies comprised in this thesis is that the VISL treatment is an efficient way of learning the metalinguistic knowledge required in the two study programmes of English and Cand. Negot. From a student-based perspective the VISL treatment is better for the lower-achieving students than the other treatments although the NON-VISL treatment is only slightly less efficient than VISL on a general scale. This needs to be looked into further. There was a difference in the entrance exams and the attendance levels of the two cohorts.

The think-aloud protocols revealed that the cognitive processes which helped students adopt the analytic and systematic approach to the subject-matter were better promoted by the VISL treatment than the other treatments and especially the low-middle achievement students benefited from the VISL treatment.

The protocols also revealed that full explicitness is achieved by few students. This may help explain why the treatments hardly affected the non-syllabus item results. The non-syllabus item results are indicative of the general linguistic proficiency of the students. However, the study did not contain a delayed post-test, which may be needed to detect true improvements in this area. The raised awareness which was effected by the

treatments is not immediately translated into improved proficiency. Further research is need into other mechanisms, routes and processes, but the results for the tense/aspect items indicate that there may be interesting perspectives in this pursuit. The positive VISL results also point to its usefulness at the tertiary level. There are few CALL resources available at this level for the advanced language learner, and the study can be interpreted to mean that a further development of VISL and other tailor-made courseware is a worthwhile effort.

Danish summary

Afhandlingen dækker tre relaterede og interagerende områder:

Fremmedsprogstilegnelse, computerstøttet sproglæring og en undersøgelse af relaterede læreprocesser. I et felt mellem computeren som læringsinstrument, det metalingvistiske aspekt af fremmedsprogslæring, og de involverede kognitive processer undersøger afhandlingen læringens dynamik og resultater.

Afhandlingen omfatter to dele: en kvantitativ del og en kvalitativ del. I den kvantitative del undersøges tre eksperimental-grupper for effekt af forskellige undervisningsmetoder. I den kvalitative del undersøges læreprocessernes kognitive aspekter med henblik på understøttelse og belysning af den kvantitative del. Strukturelt består afhandlingen af fire dele. Første del er introduktion og problemstillinger i form af de fem forskningsspørgsmål, deres begrundelse og baggrund. Anden del indeholder dels oversigter over tidligere forskningslitteratur, som specielt er tilknyttet den kvantitative undersøgelse, d.v.s. fremmedsprogstilegnelse (SLA) og computerstøttet sproglæring (CALL), dels selve den kvantitative undersøgelse inklusive resultaterne fra de tre undersøgte grupper, nemlig Visual Interactive Syntax Learning gruppen (VISL), NON-VISL gruppen (den ikke-computerstøttede gruppe), og BASIS gruppen (d.v.s. gruppen med den almindelige skema-lagte undervisning). Tredje del består dels af forskningslitteraturen for de emner, der berører den kvalitative undersøgelse, nemlig implicit og eksplicit læring og viden, og introspektion, dels selve den kvalitative undersøgelse og resultaterne deraf. Fjerde del er en diskussion af de samlede resultater af afhandlingens undersøgelser relateret til de opstillede forskningsspørgsmål, en konklusion og de perspektiver afhandlingens resultater kan opstille.

Fokus i del 1 er forskningsspørgsmålene og deres relevans for projektet. De fem forskningsspørgsmål er:

1. Er VISL lige så god som traditionel undervisning?
2. Er VISL lige god for Engelsk studerende som for Cand. Negot. studerende?
3. Er VISL lige god for alle studerende, d.v.s. såvel for toppen, midten og bunden.
4. Opnår alle studerende at erhverve sig en viden, der er fuldt eksplicit?

5. Kan undervisningen i pensum påvirke den almene sprogfærdighed?

De tre første spørgsmål vedrører effektiviteten af den undervisning de tre grupper modtager; først og fremmest den computerstøttede undervisning i forhold til den traditionelle klassebaserede undervisning, og i denne ramme undersøges, om der er forskellige læringsmønstre i de forskellige grupper.

De sidste to forskningsspørgsmål vedrører kvaliteten af læringen, og hvilken rolle det implicite og det eksplicite spiller i læring og viden. VISL er udviklet med henblik på at fremme den metalingvistiske del af pensum i engelskstudiet og Cand. Negot.-studiet. Spørgsmålene fire og fem vedrører arten af de kognitive processer, og hvilken forbindelse der eventuelt måtte bestå mellem den metalingvistiske viden og de studerendes sprogfærdighed, for så vidt som det kommer til udtryk i testen af henholdsvis pensum og den intuitive viden. Undersøgelsens dele forbinder således afhandlingens tre hovedområder: CALL, SLA og kognitive processer.

Del to omhandler mål og metode for den kvantitative undersøgelse indenfor SLA og CALL. Tidligere har forskningen i computerstøttet sproglæring været meget fokuseret på teknologi og de nye muligheder, men vendingen mod mere pædagogiske mål har betydet, at interessen samler sig om spørgsmål, der er fælles med fremmedsprogstilegnelsesforskningen. VISL er centralt placeret i forhold til disse problemstillinger. VISL er et grammatisk redskab og netop spørgsmålet om grammatikkens rolle i sprogtilegnelsen er central og fremtrædende i den teoretiske diskussion inden for SLA. Forskningsoversigten på dette område belyser de tre underlæggende og forskellige opfattelser af forbindelsen mellem den metalingvistiske, eksplicite læring og viden, og den implicite læring og viden. Den fremherskende hypotese består i, at der er en mulig forbindelse mellem de to typer viden; dette er dog afhængig af en række forhold, der skal opfyldes, før det er muligt at skabe denne forbindelse. Et sådant forhold er bevidst opmærksomhed hos læreren i læringsprocessen.

VISL har potentialet til at facilitere læring gennem den fokusering af opmærksomheden, der understøttes af brugerfladens opsætning. Programmet har et fint farveprogram, der understøtter og supplerer den indbyggede trædiagramstruktur, som de studerende arbejder med at opbygge i deres sætningsanalyse. Endnu en facilitet, der hjælper de studerende til at fokusere deres opmærksomhed er den umiddelbare feedback, som er en integreret del af programmets pædagogiske værktøjskasse.

Den kvantitative del undersøger forskellen mellem de tre eksperimentelle gruppers resultater for både engelskstuderende og Cand. Negot.-studerende. Undersøgelsen er baseret på et pre-test, post-test design, og løb over ti uger og omfattede 107 studerende fordelt mellem VISL, NON-VISL og BASIS. Såvel pre-test som post-test indeholdt to kategorier af test-spørgsmål, nemlig pensum relaterede og ikke-pensum relaterede spørgsmål. Niveauet af pensumviden var lavt i pre-testen, hvorimod der var et relativt højt niveau af ikke-pensum baseret kunnen, der kan beskrives som et udtryk for sprogfærdighed.

Post-testen viste, at undervisningen i alle tre grupper havde størst effekt på den pensumbaserede viden, hvilket var forventeligt. Der kunne påvises ringe eller ingen effekt efter de ti ugers undervisning på den ikke-pensum relaterede viden.

Undervisningen i pensum viste størst effekt i VISL gruppen, især i Cand. Negot. VISL-gruppen, men NON-VISL gruppen opnåede næsten samme niveau. Mindst effekt viste sig i BASIS gruppens resultater. Her må det dog bemærkes, at engelsk BASIS-gruppe fra starten havde et atypisk højt niveau, hvilket gør, at den vanskeligt kan bruges til sammenligning. Generelt var effekten af undervisningen fire gange større for pensum-delen end for ikke-pensum-delen.

Del tre består af den kvalitative undersøgelse, der supplerer og underbygger den kvantitative undersøgelse, idet den anvendte metode forsøger at belyse de kognitive læringsprocesser. Dette opnås gennem højtænktningsprotokoller, der analyseres ud fra Dienes og Perners (1999) teori om implicit og eksplicit viden. Den relaterede forskningslitteratur om implicit og eksplicit viden diskuteres i lyset af definitionsproblematikken omkring bevidst og ubevidst viden.

I henhold til Dienes og Perner's teori består fuldt eksplicit viden af tre strukturlag: *content* (indhold), *attitude* (opfattelse), og *self* (selvet). Hvert af disse tre lag har egne interne strukturer, som er indikatorer for niveauet af eksplicitet. Det centrale er *factuality* (faktualitet), og hvorvidt en given (del)viden opfattes som viden (kendsgerning). Dienes og Perner knytter eksplicitet til verbalisering og deklarativ viden; ligeledes knyttes den implicite viden til tavs og procedural viden.

Højtækningsprotokollerne i denne undersøgelse er at ligne med bevidsthedstrømme, idet tanker udtales i samme øjeblik, som de kommer til respondentens bevidsthed. Ericsson og Simon (1993) opstiller taksonomier for forskellige typer introspektion, og på baggrund af deres arbejde er den samtidige protokoltype udvalgt på grund af den ringe tidsforskydning mellem tanke og tale. Deltagerne i undersøgelsen blev udvalgt på grundlag af deres pre-test resultater, således at alle præstationsniveauer kunne være repræsenteret. Derpå fulgte en læringsperiode af sammenlængde og med samme indhold som i den kvantitative del, men uden BASIS grupper. Protokollernes indhold blev analyseret i henhold til det eksplicite hierarki, som Dienes og Perner opstiller i deres teori. Formålet med denne analyse var at få indblik i graden af den opnåede eksplicite viden, som derefter kunne sammenlignes med resultaterne af den kvantitative del og den beskrivende statistiske behandling af resultaterne.

Det viste sig, at få studerende erhvervede fuld eksplicit viden i pensum. En nærmere undersøgelse af resultaterne viste, at de studerende i den høje ende af skalaen i pre-testen vedblev at opnå høje resultater. Det viste sig også, at deres viden kunne aktiveres, når de stødte på problemer. Deres viden var kontrolleret, og graden af eksplicit viden høj. De studerende, der i pre-testen opnåede de laveste resultater vedblev at score lavt. Et interessant resultat var at finde i den laveste del af midtergruppen, idet der her var en effekt i henhold til undervisningsmetoden. VISL-gruppens studerende kunne bryde deres relative pre-test placering og bevæge sig op i hierarket med adskillige pladser. Dette var ikke tilfældet for de studerende i NON-VISL-gruppen. Højtækningsprotokollerne kunne vise, at disse studerende i løbet af læringsperioden ændrede deres tilgang til stoffet til en mere analytisk og struktureret fremgangsmåde.

Denne observation kan måske forklare, hvorfor den kvantitative del kunne påvise, at netop VISL-gruppens nedre kvartil lå højere end i de andre eksperimentelle grupper.

Perspektiverne af de to undersøgelser i denne afhandling er, at VISL fremmer en effektiv måde at lære det metalingvistiske pensum på. Overordnet betragtet er VISL kun lidt bedre end NON-VISL, men VISL er bedre end andre metoder for de studerende, der befinder sig i den nederste del af midten.

Protokollerne kunne påvise, at meget få studerende opnå fuldt eksplicit viden i pensum. Denne kendsgerning kan måske være forklaringen på, at niveauet i den generelle sprogfærdighed nærmest ikke blev påvirket i løbet af de ti uger, den eksperimentelle undervisning foregik. Eksperimentet indholdt ikke nogen forsinket post-test, hvilket måske ville være et instrument til at måle denne effekt, hvis den er til stede. Den eksperimentelle undervisning har skabt en opmærksomhed på pensum og på sproget, som ikke direkte overgår til målelig effekt, men måske er en proces sat i gang, hvis effekt først kan måles senere. Dette vil være målet for en senere undersøgelse. Det positive resultat er, at hvis programmet, som tilfældet er med VISL, er målrettet og integreret i curriculum, kan computerstøttet sproglæring være nyttig og motiverende også på universitetsniveau. På visse punkter og for visse studerende har computeren som redskab vist sig en bedre løsning end den traditionelle undervisning alene.

Bibliography

- Alanen, Riika. 1995. Input enhancement and rule presentation in second language acquisition. In R. Schmidt (ed.), *Attention and awareness in foreign language learning* (pp.261-309). Honolulu, Hawai'i: Second Language Teaching & Curriculum Center.
- Alderson, J. Charles, Caroline Clapham & David Steel. 1997. Metalinguistic knowledge, language aptitude and language proficiency. *Language Teaching Research* 1 (2):93-121.
- Andersen, Roger W. 1978. An implication model for second language research. *Language Learning* 28 (2):221-282.
- Anderson, J. R. 1980. *Cognitive Psychology and its Implications*. San Francisco: Freeman.
- Anderson, J. R. 1983. *The Architecture of Cognition*. Cambridge, MA & London: Harvard University Press.
- Anderson, J. R. 1992. Automaticity and the ACT theory. *American Journal of Psychology* 105:165-180.
- Anderson, J.R. 1993. *Rules of the Mind*. Hillsdale, NJ: Erlbaum.
- Anderson, Neil J. and Laurens Vandergift. 1996. Increasing Metacognitive Awareness in the L2 Classroom by Using Think-Aloud Protocols and Other Verbal Report Formats. In R. L. Oxford (ed.), *Language Learning Strategies around the World: Cross-cultural Perspectives* (pp.3-19). University of Hawai'i at Manoa.
- Arnold, Nike. 2006. CALL: Where Are We and Where Do We Go from Here. In Lara ducate & Nike Arnold (eds.), *Calling on CALL: From Theory and Research to New Directions in Foreign Language Teaching* (pp.1-20). CALICO.
- Bache, Carl, Davenport, Mike, Dienhart, John & Larsen, Fritz. 1999. *An Introduction to English Sentence Analysis*. Copenhagen: Gyldendal.
- Bache, Carl & Davidsen-Nielsen, Niels. 1997. *Mastering English Grammar*. Berlin & New York: Mouton de Gruyter.
- Baddeley, Alan. 1995; first published 1986. *Working Memory*. Oxford: Clarendon Press.
- Bardovi-Harlig, Kathleen. 2000. Tense and Aspect in Second Language Acquisition: Form, Meaning, and Use. *Language Learning* 50 (Supplement 1).
- Basturkmen, Helen, Loewen, Shawn & Ellis, Rod. 2002. Metalanguage in focus on form in the communicative classroom. *Language Awareness* 11 (1):1-13.

- Bell, T. 2005. Behaviors and attitudes of effective language teachers: Results of a questionnaire study. *Foreign Language Annals* 38 (2):260-270.
- Berry, Roger. 2000. 'User-friendly' metalanguage: What effect does it have on learners of English? *International Review of Applied Linguistics in Language Teaching* 38 (3/4):195-211.
- Bialystok, Ellen. 1978. A Theoretical Model of Second Language Learning. *Language Learning* 28 (1):69-83.
- Bialystok, E. 1985. The compatibility of teaching and learning strategies. *Applied Linguistics* 3:155-262.
- Bialystok, Ellen. 1994. Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition* 16:157-168.
- Bick, Eckhardt. 2001. The VISL System: Research and applicative aspects of IT-based learning. Paper at NoDaLiDa 2001, at Uppsala.
- Birdsong, D. 2006. Age and second language acquisition and processing: A selective overview. *Language Learning* 56:9-49.
- Borg, Simon. 1999. The use of grammatical terminology in the second language classroom: A qualitative study of teachers' practices and cognitions. *Applied Linguistics* 20 (1):95-126.
- Bracewell, Robert J. & Breuleux, Alain. 1994. Substance and Romance in Analyzing Think-Aloud Protocols. In P. Smagorinsky (ed.), *Speaking about Writing* (pp.55-88). Thousand Oaks: Sage Publications.
- Cadierno, Teresa. 1995. Formal instruction from a processing perspective: A comparison of input-based and output-based instruction in second language acquisition. *Modern Language Journal* 7 (2):179-193.
- Caplan, David and Gloria S. Waters. 1999. Verbal Working Memory and Sentence Comprehension. *Behavioral and Brain Sciences* 22 (1):77-94.
- Carruthers, P. 1992. Consciousness and concepts. *Proceedings of Aristotlian Society*. Supplementary volume, LXVI:42-59.
- Chambers, Angela & Davies, Graham, eds. 2001. *ICT and Language Learning: A European Perspective*. Lisse: Swets & Zeitlinger.
- Chapelle, Carol A. 1994. Theoretical Bases for Human-Computer Interaction Research in CALL. In F. L. Borchardt (ed.), *CALICO Annual Symposium 1994* (pp.53-57). Flagstaff, AZ: CALICO.

- Chapelle, Carol A. 1997. CALL in the year 2000: Still in search of research paradigms? *Language Learning & Technology* 1 (1):19-43.
- Chapelle, Carol A. 1999. Research questions for a CALL research agenda: A reply to Rafael Salaberry. *Language Learning & Technology* 3 (1):108-113.
- Chapelle, Carol A. 2001. *Computer Applications in Second language Acquisition*. Cambridge: Cambridge University Press.
- Chapelle, Carol A. 2003. *English Language Learning and Technology*. Amsterdam & Philadelphia: John Benjamins.
- Chapelle, Carol & Green, P. 1992. Field Independence/Dependence in Second-Language Acquisition Research. *Language Learning* 42 (1):47-83.
- Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.
- Chomsky, Noam. 1964. *Current Issues in Linguistic theory*. The Hague: Mouton.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, Noam. 2000. *New Horizons in the Study of Language and the Mind*. Cambridge: Cambridge University Press.
- Clahsen, H., Meisel, J., & Pienemann, M. 1983. *Deutsch als Zweitsprache*. Tübingen: Gunter Narr.
- Cleeremans, A. & Jiménez, L. 2002. Implicit learning and consciousness: A graded, dynamic perspective. In R. M. French & A. Cleeremans (eds.), *Implicit Learning and Consciousness: An Empirical, Computational and Philosophical Consensus in the Making?* (pp.1-40). Hove, UK: Psychology Press.
- Cohen, Andrew D. 1996. Verbal Reports as a Source of Insights into Second Language Learner Strategies. *Applied Language Learning* 7 (1& 2):5-24.
- Collins. 1990. *Collins Cobuild Grammar*. London: Collins.
- Davies, Graham. 2001. New Technologies and language learning: A suitable subject for research? In G. Davies & A. Chambers (eds.), *ICT and Language Learning: A European Perspective* (pp. 13-27). Lisse: Swets & Zeitlinger Publishers.
- de Graaff, Rick. 1997a. *Differential effects of explicit instruction on second language acquisition*. Leiden: Holland Institute of Generative Linguistics.
- de Graaff, Rick. 1997b. The eXperanto experiment. *Studies in Second Language Acquisition* 19 (2):249-276.

DeKeyser, Robert. 2003. Implicit and Explicit Learning. In C. J. Doughty & M. H. Long (eds.), *The Handbook of Second Language Acquisition* (pp.313-349). Oxford: Blackwell.

DeKeyser, Robert M. 1995. Learning Second Language Grammar Rules: An Experiment With a Miniature Linguistic System. *Studies in Second Language Acquisition* 17:379-410.

DeKeyser, Robert M. 1996. Exploring Automatization Processes. *TESOL Quarterly* 30:349-357.

DeKeyser, Robert M. 1997. Beyond Explicit Rule Learning: Automatizing Second Language Morphosyntax. *Studies in Second Language Acquisition* 19:195-221.

DeKeyser, Robert M. 2001. Automaticity and automatization. In P. Robinson (ed.), *Cognition and Second Language Instruction* (pp.125-151). Cambridge: Cambridge University Press.

Despain, Bruce. 2006. *Diagramming English*. URL:
http://userpages.burgoyne.com/bdespain/grammar/rk_gram/r_k_toc.htm

Dienes, Zoltan & Perner, Josef. 1999. A Theory of Implicit and Explicit Knowledge. *Behavioral and Brain Sciences* 22 (5).

Dienes, Zoltan & Berry, Diane. 1997. Implicit learning: Below the subjective threshold. *Psychonomic Bulletin and Review* 4:3-23.

Dienes, Zoltan & Perner, Josef. 2002. A Theory of Implicit and Explicit Knowledge. In R. M. French, & A. Cleeremans (eds.), *Implicit Learning and Consciousness* (pp.68-92). Hove and New York: Taylor & Francis.

Dietz, Gunther. 2002. On rule complexity: A structural approach. *EUROSLA Yearbook* 2:263-286.

Doughty, Catherine and Jessica Williams, eds. 1998. *Focus on Form in Classroom Second Language Acquisition*. Cambridge: Cambridge University Press.

Ducate, Lara & Arnold, Nike, eds. 2006. *Calling on CALL: From Theory and Research to New Directions in Foreign Language Teaching*. San Marcos, Texas: CALICO Monograph Series.

Ellis, Nick C. 1993. Rules and Instances in Foreign Language Learning: Interactions of Explicit and Implicit Knowledge. *European Journal of Cognitive Psychology* 5 (3):289-318.

Ellis, Nick C., ed. 1994a. *Implicit and explicit learning of languages*. London: Academic Press.

- Ellis, Nick C. 1994b. Implicit and Explicit Language Learning - An Overview. In N. C. Ellis (ed.), *Implicit and explicit learning of languages* (pp.1-31). London: Academic Press.
- Ellis, Nick C., Laporte, N., O'Dochartaigh, C., Hicks, B., Hughes, E., Quinn, P., & Morgan, M. 1996. Designing and Evaluating a New CALL Program: *Rhugl (Fluent)* for learners of Welsh. In *EuroCALL 1996*.
- Ellis, Nick C. 1998. Emergentism, Connectionism and Language learning. *Language Learning* 48 (4):631-664.
- Ellis, Nick C. 2005. At the Interface: Dynamic Interactions of Explicit and Implicit Language Knowledge. *Studies in Second Language Acquisition* 27:305-352.
- Ellis, Nick C. & Larsen-Freeman, Diane. 2006. Language Emergence: Implications for Applied Linguistics - Introduction to the Special Issue. *Applied Linguistics* 27 (4):558-589.
- Ellis, Rod. 1984. Can syntax be taught? A study of the effects of formal instruction on the acquisition of WH-questions by children. *Applied Linguistics* 5 (2):138-155.
- Ellis, Rod. 1991. Grammaticality Judgements and Second Language Acquisition. *Studies in Second Language Acquisition* 13:161-186.
- Ellis, Rod. 1993. The Structural Syllabus and Second Language Acquisition. *TESOL Quarterly* 27 (1):91-113.
- Ellis, Rod. 2004. The Definition and Measurement of L2 Explicit Knowledge. *Language Learning* 54 (2):227-275.
- Ellis, Rod. 2005. Measuring implicit and explicit knowledge of a second language: A Psychometric Study. *Studies in Second Language Acquisition* 27:141-172.
- Ellis, Rod. 2006. Current Issues in the Teaching of Grammar: An SLA Perspective. *TESOL Quarterly* 40 (1):83-107.
- Ericsson, Anders K. & Simon, Herbert A.. 1980. Verbal Reports as Data. *Psychological Review* 87 (3):215-251.
- Ericsson, Anders K. & Simon, Herbert A. 1984,1993. *Protocol Analysis: Verbal Reports as Data*. Cambridge, MA: MIT Press.
- Flynn, Suzanne. 1991. The relevance of linguistic theory for language pedagogy: Debunking the myths. In J. Alatis (ed.), *Linguistics and language pedagogy: The state of the art* (pp. 547-554). Georgetown University Round Table on Languages and Linguistics 1991. Washinton, D.C.: Georgetown University Press.

Fodor, Jerry A. 1983. *The Modularity of Mind*. MIT Press.

Fotos, Sandra. 2005. Traditional and Grammar Translation Methods for Second Language Teaching. In E. Hinkel (ed.), *Handbook of Research in Second Language Teaching and Learning* (pp.653-670). Mahwah, New Jersey: Lawrence Erlbaum Associates Publishers.

Frensch, Peter A., Haider, Hilde, Rüniger, Dennis, Uwe Neugebauer, Voigt, Sabine & Werg, Jana. 2003. The route from implicit learning to verbal expression of what has been learned. In L. Jiménez (ed.), *Attention and Implicit Learning* (pp. 335-366). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Færch, Claus. 1985. Meta talk in the FL classroom. *Studies in Second Language Acquisition* 7 (2):184-199.

Færch, Claus, Haastrup, Kirsten & Phillipson, Robert. 1984. *Learner Language and Language Learning*. Copenhagen: Gyldendal.

Færch, Claus, & Kasper, G.. 1986. The Role of Comprehension in Second-language Learning. *Applied Linguistics* 7 (3):257-274.

Færch, C. & Kasper, Gabriele, ed. 1987. *Introspection in second language research*. Clevedon: Multilingual Matters.

Færch, Claus & Kasper, Gabriele. 1987a. From Product to Process - Introspective methods in Second Language Research. In C. Færch & G. Kasper (eds), *Introspection in Second Language Research* (pp.1-23). Clevedon & Philadelphia: Multilingual Matters.

Færch, Claus & Kasper, Gabriele. 1987b. One learner - two languages: Investigating types of interlanguage knowledge. In J. House, & S. Blum-Kilka, (eds.) *Introspection in second language research* (pp. 211-227).Tübingen: Gunter Narr.

Garrett, Nina. 1995. ICALL and Second Language Acquisition. In M. V. Holland, J. D. Kaplan, & M. R. Sams (eds.), *Intelligent Language Tutors* (pp.345-358). Mahwah, NJ: Lawrence Erlbaum Associates.

Garrett, Nina. 1998. Where do research and practice meet? Developing a discipline. *ReCALL* 10 (1):7-12.

Gass, Susan, Svetics, Ildekó, & Lemelin, Sarah. 2003. Differential Effects of Attention. *Language Learning* 53 (3):497-545.

Gordon, R.M. 1995. Simulation without introspection or inference from me to you. In M. Davies, & T. Stone (eds.), *Mental simulations: Evaluations and applications*. Oxford: Blackwell.

- Green, Alison. 1998. *Verbal protocol analysis in language testing research: A handbook*. Cambridge:Cambridge University Press.
- Gregg, Kevin R. 1986. The input hypothesis: Issues and Implications (review). *TESOL Quarterly* 20:116-122.
- Guan, Z., Lee, S., Cuddihy, E., & Ramey, J. 2006. The validity of the stimulated retrospective think-aloud method as measured by eye tracking. Paper at CHI 2006, April 22-27, at Montreal, Canada.
- Han, Youngju & Ellis, Rod. 1998. Implicit knowledge, explicit knowledge and general language proficiency. *Language Teaching Research* 2 (1):1-23.
- Harley, Birgit. 1989. Functional grammar in French immersion: A classroom experiment. *Applied Linguistics* 10 (3):331-359.
- Holland, Melissa V., Kaplan, Jonathan D., & Sams, Michelle R. eds. 1995. *Intelligent Language Tutors: Theory Shaping Technology*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Hulstijn, Jan H. 1989. Implicit and Incidental Second Language Learning: Experiments in the Processing of Natural and Partly Artificial Input. In H. Deckert & M. Raupack (eds.), *Interlingual Processes* (pp.49-73). Tübingen: Günther Narr.
- Hulstijn, Jan H. 1995. Not all grammar rules are equal: Giving grammar instruction its proper place in foreign language teaching. In R. Schmidt (ed.), *Attention and awareness in foreign language learning* (pp. 359-383). Honolulu, Hawai'i: Second Language Teaching & Curriculum Center.
- Hulstijn, Jan H. 1997. Second Language Acquisition Research in the Laboratory. Possibilities and Limitations. *Studies in Second Language Acquisition* 19:131-143.
- Hulstijn, Jan. 2002. Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research* 18 (3):193-223.
- Hulstijn, Jan H. 2003. Incidental and Intentional learning. In Doughty, C. J. & M. H. Long (eds.), *The Handbook of Second Language Acquisition* (pp. 349-382). Oxford: Blackwell.
- Hulstijn, Jan H. 2005. Theoretical and empirical issues in the study of implicit and explicit second-language learning. *Studies in Second Language Acquisition* 27 (2):129-140.
- Hulstijn, Jan H. & Rick de Graaff. 1994. Under what conditions does explicit knowledge of a second language facilitate the acquisition of implicit knowledge? A research proposal. *AILA review* 11:97-112.
- Hyltenstam, K. 1978. Variation in interlanguage syntax. *Lund University Department of Linguistics Working Papers* 18:1-79.

- Haastrup, Kirsten. 1987. Using think-aloud and retrospection to uncover learners' lexical inferencing procedures. In C. Færch & G. Kaper (eds.), *Introspection in second language research* (pp.197-212). Clevedon and Philadelphia: Multilingual matters.
- Haastrup, Kirsten. 1989. Lexical inferencing procedures. Doctoral thesis. Copenhagen: Det Erhvervsproglige Fakultet, Copenhagen Business School.
- Jager, Sake. 2001. From gap-filling to filling the gap. In G. Davies & A. Chambers (eds.), *ICT and Language Learning: A European Perspective* (pp.101-110). Lisse: Swets & Zeitlinger.
- Karlsson, Fred, Voutilainen, Atro, Heikilä, Juka, & Anttila, Arto, eds. 1995. *Constraint Grammar: A Language-Independent System for Parsing Unrestricted Text*. Berlin: Mouton de Gruyter.
- Karmiloff-Smith, Anette. 1986. From meta-processes to conscious access: Evidence from children's metalinguistic and repair data. *Cognition* 23:95-147.
- Kempen, Gerard. 1998. Sentence Parsing. In A. D. Frederici (ed.), *Language comprehension: A biological Perspective* (pp. 213-228). Berlin/Heidelberg: Springer.
- Kempen, Gerard. 1999, 2nd edition. Sentence parsing. In A. D. Frederici (ed.), *Language comprehension: A biological Perspective* (pp. 211-227). Berlin/Heidelberg: Springer.
- Kettemann, Bernhard. 1995. How effective is CALL in ELT? *ReCALL* 7 (1):49-53.
- Kim, Hae-Jin. 2007. Q-matrix construction: Defining the link between constructs and test items in cognitive diagnosis approaches. Paper read at 29th Annual Language Testing Research Colloquium (LTRC), June 7-11, at University of Barcelona, Spain.
- Kleinmann, Howard. 1987. The Effect of Computer Assisted Instruction on ESL Reading Achievement. *The Modern Language Journal* 71 (3):267-276.
- Knowlton, B. J., Ramus, S. J., & Squire, L. R. 1992. Intact artificial grammar learning in amnesia: Dissociation of classification learning and explicit memory for specific instances. *Psychological Science* 3:172-179.
- Kolln, Martha. 1994. *Understanding English Grammar*. New York: MacMillan Publishing Company.
- Kowal, Maria & Swain, Merrill. 1997. From semantic to syntactic processing. In R. K. Johnson & M. Swain (eds.), *Immersion Education: International Perspectives* (pp.284-309). Cambridge: Cambridge University Press.
- Krashen, Stephen D. 1978. Is the "natural order" an artifact of the bilingual syntax measure? *Language Learning* 28 (1):187-191.

- Krashen, Stephen. 1985. *The input hypothesis: Issues and implications*. London: Longman.
- Krashen, Stephen D. 1981. *Second language Acquisition and Second Language Learning*. Oxford: Pergamon.
- Krashen, Stephen D. 1991. The Input Hypothesis: An Update. In J. Alatis (ed.), *Linguistics and language pedagogy: The state of the art* (pp. 409-431). Washington, D. C.: Georgetown University Press.
- Krashen, Stephen D. 1994. The Input Hypothesis and Its Rivals. In Ellis, Nick C. (ed.), *Implicit and Explicit Learning of Languages* (pp.45-77). London: Academic Press.
- Krashen, Stephen & Terrell, Tracy. 1983. *The Natural Approach: Language acquisition in the classroom*. Oxford: Pergamon.
- LaBerge, D. & Samuels, S. J. 1974. Towards a theory of automatic information processing in reading. *Cognitive Psychology* 6:293-323.
- Larsen-Freeman, D., & Long, M. 1991. *An Introduction to Second Language Acquisition*. London: Longman.
- Legenhausen, Lienhard & Wolff, Dieter. 1990. CALL in use - use of CALL; evaluating CALL software. *System* 18 (1):1-14.
- Leow, Ronald P. 1993. To simplify or not to simplify. *Studies in Second Language Acquisition* 15:333-355.
- Leow, Ronald P. & Morgan-Short, Kara. 2004. To Think Aloud or Not to Think Aloud: The Issue of Reactivity in SLA Research Methodology. *Studies in Second Language Acquisition* 26:35-57.
- Liebermann, Matthew D. , Chang, Grace, Chiao, Joan, Bookheimer, Susan V. & Knowlton, Barbara J. 2004. An Event-Related fMRI Study of Artificial Grammar Learning in a Balanced Chunk Strength Design. *Journal of Cognitive Neuroscience* 16:427-438.
- Lightbown, Patsy M. 2000. Anniversary Article: Classroom SLA Research and Second Language Teaching. *Applied Linguistics* 21 (4):431-462.
- Lightbown, Patsy M. 2001. Input filters in second language acquisition. *EUROSLA Yearbook*:79-97.
- Lightbown, Patsy M. & Pienemann, Manfred. 1993. Comments on Stephen Krashen's "Teaching Issues: Formal Grammar Instruction". *TESOL Quarterly* 27 (4):717-725.
- Lightbown, Patsy M. & Spada, Nina. 1990. Focus-on-form and corrective feedback in communicative language teaching. *Studies in Second Language Acquisition* 12:429-448.

Lightbown, Patsy M. & Spada, Nina. 1999. *How Languages Are Learned*: Oxford University Press.

Logan, G. D. 1988. Towards an instance theory of automatization. *Psychological Review* 95:492-527.

Logan, G. D. 1990. Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology* 22:1-35.

Logan, G. D. 1992. Shapes of reaction-time distributions and shapes of learning curves: A test of instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18:883-914.

Long, Michael H. 1991. Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg & C. Kramersch (eds.), *Foreign language research in cross-cultural perspective* (pp.39-52). Amsterdam: John Benjamins.

Long, Michael H. & Doughty, Catherine J. 2003. SLA and Cognitive Science. In C. J. Doughty & M. H. Long (eds.), *Handbook of second language acquisition* (pp. 866-70). Oxford: Blackwell.

Long, Michael H. & Robinson, Peter. 1998. Focus on form: theory, research, and practice. In C. J. Doughty & J. Williams (eds.), *Focus on Form in Classroom Second Language Acquisition* (pp.15-41). Cambridge: Cambridge University Press.

Lund, Karen. 1997. Do all learners follow the natural sequences? *The Clarion* 3 (2):3-7.

Lyster, Roy & Ranta, Leila. 1997. Corrective Feedback and Learner Uptake. Negotiation of Form in Communicative Classrooms. *Studies in Second Language Acquisition* 19:37-66.

Mandell, Paul B. 1999. On the reliability of grammaticality judgement tests in second language acquisition research. *Second Language Research* 15 (1):73-99.

Mandler, G. 1975. Consciousness: Respectable, useful and probably necessary. In Solso, R. (ed.), *Information processing and cognition: the Loyola Symposium* (pp. 229-254). Hillsdale, NJ.: Lawrence Erlbaum Associates.

Marcel, A. J. 1983. Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology* 15:197-237.

Master, P. 1994. The effect of systematic instruction on learning the English article system. In T. Odlin (ed.), *Perspectives on pedagogical grammar* (pp.229-252). Cambridge, UK: Cambridge University Press.

Mathews, Robert C. & Cohran, Barbara P. 1998. Project Grammmarama Revisited: Generativity of Implicitly Acquired Knowledge. In M. A. Stadler & P. A. Frensch

(eds.), *Handbook of Implicit Learning* (pp.223-259). Thousand Oaks, CA; London, New Delhi: Sage Publications.

Matthews, C. 1993. Grammar Frameworks in Intelligent CALL. *The CALICO Journal* 11 (1):5-26.

McBride, Nicole & Seago, Karen. 1997. Bridging the gap: grammar as hypertext. *ReCall* 9 (2):17-25.

Meisel, Jürgen M., Clahsen, Harald & Pienemann, Manfred. 1981. On determining developmental stages in second language acquisition. *Studies in Second Language Acquisition* 3:109-135.

Miller, G. A. 1962. *Psychology: The science of mental life*. New York: Harper & Row.

Nagata, Noriko. 1995. An Effective Application of Natural Language Processing in Second Language Instruction. *CALICO* 13 (1): 47-67.

Nagata, Noriko. 1996. Computer vs. Workbook Instruction in Second Language Acquisition. *CALICO Journal* 14 (1):53-75.

Nagata, Noriko. 1997a. An experimental comparison of deductive and inductive feedback generated by a simple parser. *System* 25 (4):515-534.

Nagata, Noriko. 1997b. Intelligent Computer Feedback for Second Language Instruction. *The Modern Language Journal* 77 (3):330-339.

Neal, Andrew & Hesketh, Beryl. 1997. Episodic knowledge and implicit learning. *Psychonomic Bulletin and Review* 4 (1):24-37.

Neisser, U. 1967. *Cognitive Psychology*. New York: Harper & Row.

Nisbett, Richard E. & Wilson, Timothy DeCamp. 1977. Telling More than We Can Know: Verbal reports on Mental Processes. *Psychological Review* 84 (3):231-259.

Norris, John M. & Ortega, Lourdes. 2000. Effectiveness of L2 Instruction: A Research Synthesis and Quantitative Meta-analysis. *Language Learning* 50 (3):417-528.

Norris, John M. & Ortega, Lourdes. 2001. Does Type of Instruction Make a Difference? Substantive Findings from a Meta-analytic Review. *Language Learning* 51 (Supplement 1):157-213.

Odlin, Terence. 1993. The Introspective Hierarchy: A comparison of intuitions of linguists, teachers, and learners. In T. Odlin (ed.), *Perspectives on Pedagogical grammar* (pp. 271-292). Cambridge: Cambridge University Press.

- Oxford, Rebecca. 1995. Linking theories of learning with computer-assisted language learning. In V. M. Holland, J. D. Kaplan & M. R. Sams (eds.), *Intelligent language tutors: Theory shaping technology* (pp.359-369). Mahwah, NJ: Lawrence Erlbaum Associates.
- Perruchet, Pierre, Vinter, Annie & Gallego, Jorge. 1997. Implicit Learning shapes new conscious percepts and representations. *Psychonomic Bulletin & Review* 4 (1):43-48.
- Perruchet, Pierre and Pacteau, Chantal. 1990. Synthetic Grammar Learning: Implicit Rule Abstraction or Explicit Fragmentary Knowledge. *Journal of Experimental Psychology: General* 119 (3):264-275.
- Perruchet, Pierre & Vinter, Annie. 2002. The self-organizing consciousness. *Behavioral and Brain Sciences* (25):297-388.
- Perruchet, Pierre & Vinter, Annie. 2002b. The Self-organizing Consciousness: A framework for implicit learning. In R. M. French & A. Cleeremans (eds.), *Implicit Learning and Consciousness* (pp.41-67). Hove: Psychology Press.
- Pienemann, Manfred. 1986. Psychological constraints on the teachability of languages. In C. W. Pfaff (ed.), *First and second language acquisition processes*. Rowley, MA: Newbury House.
- Pienemann, Manfred. 1998. Developmental dynamics in L1 and L2 acquisition: Processability theory and generative entrenchment. *Bilingualism: Language and Cognition* 1:1-20.
- Pienemann, Manfred. 1999. *Language Processing and Second Language Development: Processability Theory*. Amsterdam and Philadelphia: John Benjamins.
- Pienemann, M., & Johnston, M. 1987. Factors affecting the development of language proficiency. In Nunan, D. (ed.), *Applying second language acquisition research* (pp. 45-141). Adelaide: National Curriculum resource Centre.
- Pienemann, Manfred, Johnston, M. & Brindley, G. 1988. Constructing an acquisition-based procedure for second language assessment. *Studies in Second Language Acquisition* 10 (2):217-243.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. 1985. *A Comprehensive Grammar of the English Language*. Harlow: Longman.
- Reber, Arthur S. 1967. Implicit Learning of Artificial Grammars. *Journal of Verbal Learning and Verbal Behavior* 6:855-863.
- Reber, Arthur S. 1980. On the relationship between implicit and explicit modes in the learning of a complex rule structure. *Journal of Experimental Psychology: Human Learning and Memory* 6 (5):492-502.

- Reber, A.S. 1985. Syntactical Learning and Judgement, Still Unconscious and Still Abstract: Comment on Dulany, Carlson, and Dewey. *Journal of Experimental Psychology: General* 114 (1):17-24.
- Reber, A.S. 1989. Implicit Learning and Tacit Knowledge. *Journal of Experimental Psychology: General* 18:219-235.
- Reber, A.S. 1992. The cognitive unconscious: An evolutionary perspective. *Consciousness and Cognition* 1:93-133.
- Reber, Arthur S. 1997. How to Differentiate Implicit and Explicit Modes of Acquisition (pp.137-160). In J. D. Cohen & J. W. Schooler (eds.), *Scientific Approaches to Consciousness*. Mahwah, NJ.: Lawrence Erlbaum Associates.
- Reber, Arthur S. 1997. Implicit ruminations. *Psychonomic Bulletin & Review* 4 (1):49-55.
- Reed, Alonzo & Kellogg, Brainerd. 1907. *Higher Lessons in English: A Work on English Grammar and Composition, In which the Science of the Language is made tributary to the Art of Expression*. New York.
- Robinson, Peter. 1996. Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition* 18:27-67.
- Robinson, Peter. 1997. Generalizability and Automaticity of Second Language Learning under Implicit, Incidental, Enhanced and Instructed Conditions. *Studies in Second Language Acquisition* 19 (2):223-247.
- Robinson, Peter. ed. 2001. *Cognition and second language instruction*. New York: Cambridge University Press.
- Robinson, Peter & Ha, Mee Aie. 1993. Instance theory and second language rule learning under explicit conditions. *Studies in Second Language Learning* 15 (4):413-438.
- Rosa, Elena & O'Neill, Michael D. 1999. Explicitness, intake and the issue of awareness. *Studies in Second Language Acquisition* 21:511-556.
- Rosenthal, David M. 1997. A Theory of Consciousness. In N. Block, O. Flanagan & G. Güzeldere (eds.), *The Nature of Consciousness* (pp.729-753). Cambridge, Massachusetts, London, England: MIT Press.
- Rosenthal, Judith W. ed. 2000. *Handbook of Undergraduate Second Language Education*. Mahwah, NJ./London: Lawrence Erlbaum.
- Ruin, Inger. 1996. *Grammar and the Advanced Learner*. Uppsala: Uppsala University.

- Ryle, Gilbert. 1949. *The Concept of Mind*. London: Penguin.
- Salaberry, Rafael. 1999. CALL in the year 2000: Still developing the research agenda. *Language Learning & Technology* 3 (1):104-107.
- Salaberry, Rafael & Shirai, Yasukiro, ed. 2002. *The L2 Acquisition of Tense-Aspect Morphology*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Salaberry, Rafael M. 2001. The Use of Technology for Second Language Learning and Teaching: A Retrospective. *The Modern Language Journal* 85 (1):39-56.
- Schmidt, R., & Frota, S. 1986. Developing basic conversational ability in a second language: A case study of an adult learner. In R. Day (ed.) *Talking to learn* (pp.237-326). Rowley, MA: Newbury House.
- Schmidt, Richard. 1990. The role of consciousness in second language learning. *Applied Linguistics* 11:129-158.
- Schmidt, Richard. 1994. Implicit Learning and the Cognitive Unconscious: Of Artificial Grammars and SLA. In N. C. Ellis (ed.), *Implicit and Explicit Learning of Languages* (pp.165-209). London:Academic Press.
- Schmidt, Richard, ed. 1995a. *Attention and awareness in foreign language learning*. Vol. Technical Report #9. Honolulu, Hawai'i: Second Language Teaching & Curriculum Center.
- Schmidt, Richard. 1995b. Consciousness and foreign language learning: a tutorial on the role of attention and awareness in learning. In R. Schmidt (ed.), *Attention and awareness in foreign language learning* (pp.1-65). Honolulu, Hawai'i: Second Language Teaching & Curriculum Center.
- Schneider, W. & Shiffrin, R. M. 1977. Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review* 84:1-66.
- Schulze, Mathias. 1999. From the developer to the learner:describing grammar - learning grammar. *ReCALL* 11 (1):117-124.
- Scollon, Roy & Scollon, Suzanne Wong. 1995. *Intercultural Communication*. Oxford, UK & Cambridge, USA: Blackwell.
- Segalowitz, Norman. 2003. Automaticity and Second Languages. In C. J. Doughty & M. H. Long (eds.) *The Handbook of Second Language Acquisition* (pp.382-409). Oxford: Blackwell.
- Segalowitz, Norman & Lightbown, Patsy M. 1999. Psycholinguistic approaches to SLA. *Annual Review of Applied Linguistics* 19:43-63.

Sharwood Smith, Michael. 1981. Consciousness-raising and the second language learner. *Applied Linguistics* 2 (159-168).

Sharwood Smith, Michael. 1991. Speaking to many minds: On the relevance of different types of language information to the L2 learner. *Second Language Research* 7 (2):118-132.

Sharwood Smith, Michael. 1993. Input Enhancement in Instructed SLA. *Studies in Second Language Acquisition* 15 (2):165-179.

Sharwood Smith, Michael. 2000a. Attentional Mechanisms and the Language Acquisition Device. In Birgitta E. Dimitrova & Kenneth Hylltenstam (eds.) *Language processing and simultaneous interpreting: Interdisciplinary perspectives* (pp.25-45). Amsterdam: John Benjamins.

Sharwood Smith, Michael. 2000b. SLA beyond 2000. In C. Riemer (ed.), *Kognitive Aspekte des Lehrens und Lernens von Fremdsprachen: Cognitive Aspects of Foreign Language Learning and Teaching. Festschrift für Willis J. Edmondson zum 60. Geburtstag* (pp.1-11). Tübingen: Günter Naur Verlag.

Sharwood Smith, Michael. 2004. In two minds about grammar: On the interaction of linguistic and metalinguistic knowledge in performance. *Transaction of the Philosophical Society* 102 (2):225-280.

Sharwood Smith, Michael. 2005. More on the meta mode: In search of deeper explanations for the role of consciousness in second language learning. In A. Psaltou-Joyce & M. Mattheoudakis (eds.), *Selected Papers on Theoretical and Applied Linguistics, 11-13 April, 2003* (pp.28-43). Thessaloniki: University Studio Press.

Sharwood Smith, Michael & Rutherford, William E. 1985. Consciousness-raising and universal grammar. *Applied Linguistics* 6 (3):274-282.

Sharwood Smith, Michael & Truscott, John. 2005. Stages or continua in second language acquisition: A MOGUL solution. *Applied Linguistics* 26 (2):219-240.

Skehan, Peter. 1989. *Individual Differences in Second-Language Learning*. London, New York, Melbourne, Auckland: Edward Arnold.

Smagorinsky, Peter. 1987. Graves revisited: A look at the methods and conclusions of the New Hampshire study. *Written Communication* 4 (4):331-342.

Smagorinsky, Peter. 1994a. Think-Aloud Protocol Analysis Beyond the Black Box. In P. Smagorinsky (ed.), *Speaking about Writing* (pp. 3-19).. Thousand Oaks: Sage Publication.

Smagorinsky, Peter, ed. 1994b. *Speaking about Writing*. Thousand Oaks: Sage Publications.

- Squire, L.R. 1992. Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review* 99 (2):195-231.
- Squire, Larry R. 1995. Declarative and Nondeclarative Memory: Multiple Brain Systems Supporting Learning and Memory. In *Memory Systems 1994* (pp. 203-231). Cambridge, Mass. & London, England: The MIT Press.
- Stadler, M. A. 1995. The role of attention in implicit learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 21:674-685.
- Stadler, Michael A. 1997. Distinguishing implicit and explicit learning. *Psychonomic Bulletin & Review* 4 (1):56-62.
- Strawson, P.F. 1959. *Individuals*. Methuen.
- Svartvik, J. & Sager, O. 1978. *Engelsk Universitetsgrammatik*. Stockholm: Esselte Studium.
- Swain, Merrill. 1985. Communicative competence: Some roles of comprehensible input and comprehensible output and its development. In S. M. Gass & C. Madden (eds.) *Input in second language acquisition* (pp.235-253). Rowley, MA: Newbury House.
- Swain, Merrill. 1998. Focus on form through conscious reflection. In C. Doughty, & J. Williams (eds.) *Focus on form in classroom second language acquisition* (pp.64-81). Cambridge: Cambridge University Press.
- Swain, Merrill. 2005. The Output Hypothesis: Theory and Research. In E. Hinkel (ed.), *Handbook of Research in Second Language Teaching and Learning* (pp.471-483). Mahwah, New Jersey: Lawrence Erlbaum Associates Publishers.
- Swain, Merrill & Carroll, Susanne. 1987. The immersion observation study. In B. Harley, P. Allen, J. Cummins, & M. Swain (eds.). *The development of bilingual proficiency: Final report, Vol 2* (pp.190-341). Toronto: OISE.
- Swain, Merrill & Lapkin, S. 1986. Immersion French at the secondary level: The "goods" and the "bads". *Contact* 5:2-9.
- Swan, Michael. 2005. *Practical English Usage*. Oxford: Oxford University Press.
- Teichert, Herman U. 1985. Computer Assisted Instruction in Beginning College German: An Experiment. *CALICO Journal* 3 (2):18-24.
- Towell, Richard. 2000. *Research Assessment Exercise (RAE) and CALL* CILT Research Forum; Linguanet, 1999 [retrieved 17 March 2000]. Available from <http://www.linguanet.org.uk/research/resfor2/towell.htm>.
- Trahey, M. & White, L. 1993. Positive evidence and preemption in the second language classroom. *Studies in Second Language Acquisition* 15 (181-204).

Uddannelsesstyrelsen, Undervisningsministeriet. 2000. *Uddannelsesvejledning for gymnasiet. Engelsk. Fagets forskellige niveauer* Maj 1999 [retrieved 8 August 2000]. Available from:

<http://www.uvm.dk/gymnasie/almen/vejledning/undervisgym/engelsk.htm>.

Undervisningsministeriet. 2000. *Bekendtgørelse om kusus til højere forberedelseksamen og om studieforberevende enkeltfagseksamen* 1995 [retrieved 18 September 2000]. Available from <http://www.uvm.dk/lov/bek/1995/00000605.htm>.

Undervisningsministeriet. 2000. *Gymnasiebekendtgørelsen* 1999 [retrieved 8 August 2000]. Available from <http://uvm.dk/lov/bek/1999/0000411.htm>.

Undervisningsministeriet. 2000. *Bekendtgørelse om den erhvervsfaglige uddannelse til højere handelseksamen* 2000 [retrieved 18 September 2000]. Available from <http://www.uvm/lov/bek/2000/0000249.htm>.

Undervisningsministeriet. 2000. *Bekendtgørelse om den erhvervsgymnasiale uddannelse til højere teknisk eksamen* 2000 [retrieved 18 September 2000]. Available from <http://www.uvm.dk/lov/bek/2000/0000524.htm>.

van Lier, Leo. 1991. Language awareness: The common ground between linguist and language teacher. In J. Alatis (ed.), *Linguistics and language pedagogy: The state of the art* (pp.528-546). Washinton D.C.:Georgetown University Press.

van Someren, Maarten W., Barnard, Yvonne F. & Sandberg, Jacobijn A. C.. 1994. *The Think Aloud Method*. London, San Diego: Academic Press.

VanPatten, Bill. 1990. Attending to form and content in the input: An experiment in consciousness. *Studies in Second Language Acquisition* 12:287-301.

VanPatten, Bill. 1994. Evaluating the role of consciousness in second language acquisition: Terms, linguistic features & research methodology. *AILA Review* 11:27-36.

VanPatten, Bill, ed. 1996. *Input Processing and Grammar Instruction in Second Language Acquisition*. Norwood, NJ: Ablex Publishing Corporation.

VanPatten, Bill & Cadierno, Teresa. 1993. Explicit instruction and input processing. *Studies in Second Language Acquisition* 15 (2):225-261.

Warschauer, M. & Healey, D. 1998. Computers and language learning: An overview. *Language Teaching* 31:57-71.

Waters, Gloria S. & David Caplan. 2004. Verbal Working Memory and on-line syntactic processing: Evidence from self-paced listening. *The Quarterly Journal of Experimental Psychology* 57A (1):129-163.

Wei, Yu Hong & Davies, Graham. 1996. Do grammar checkers work? Paper at EUROCALL 96, Daniel Berszeny College, Hungary.

White, Lydia. 1991. Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research* 7 (2):133-161.

White, Lydia, Spada, Nina, Lightbown, P. & Ranta, L. 1991. Input Enhancement and L2 Question Formation. *Applied Linguistics* 12 (4):416-432.

White, Peter. 1980. Limitations on verbal report of internal events: A refutation of Nisbett and Wilson and of Bern. *Psychological Review* 87 (1):105-112.

Whittlesea, Bruce W.A. & Dorken, Michael D. 1997. Implicit learning: Indirect, not unconscious. *Psychonomic Bulletin & Review* 4 (1):63-67.

Widdowson, H. 1990. *Aspects of Language Teaching*. Oxford: Oxford University Press.

Willingham, Daniel B. & Goedert-Eschmann, Kelly. 1999. The relation between implicit and explicit learning: Evidence for Parallel Development. *Psychological Science* 10 (6):531-534.

Winter, Bill & Reber, Arthur S. 1994. Implicit Learning and the Acquisition of Natural Languages. In N. C. Ellis (ed.). *Implicit and Explicit Learning of Languages* (pp.114-145). London:Academic Press.

Wode, Henning. 1976. Developmental Sequences in Naturalistic L2 Acquisition. *Working Papers on Bilingualism (OISE)* 11:1-31.

Wundt, William. 1896. *Grundriss der Psychologie*. Leipzig: Engelmann.

Zähner, Christopher. 1995. Second language acquisition and the computer: Variation in second language acquisition. *ReCall* 7 (1):34-48.

Appendix I Analysed and quantified sentence corpus (A : Sentences ; B : Quantification)

A : Sentences

Anne wrote a letter

STA:cl

S:n Anne

P:v wrote

Od:g

=D:art a

=H:n letter

Anne wrote a letter for her sister

STA:cl

S:n Anne

P:v wrote

Od:g

=D:art a

=H:n letter

A:g

=H:prp for

=D:g

==D:pron her

==H:n sister

The students were busy

STA:cl

S:g

=D:art The

=H:n students

P:v were

Cs:adj busy

The students were busy with their studies every afternoon

STA:cl

S:g

=D:art	The
=H:n	students
P:v	were
Cs:g	
=H:adj	busy
=D:g	
==H:prp	with
==D:g	
===D:pron	their
===H:n	studies
A:g	
=D:pron	every
=H:n	afternoon

Afterwards they had a drink

STA:cl	
A:adv	Afterwards
S:pron	they
P:v	had
Od:g	
=D:art	a
=H:n	drink

The autumn air is so enjoyable

STA:cl	
S:g	
=D:art	The
=D:n	autumn
=H:n	air
P:v	is
Cs:g	
=D:adv	so
=H:adj	enjoyable

Outside, the concert-goers were heading towards the tube stations

STA:cl
 A:adv Outside
 S:g
 =D:art the
 =H:n concert-goers
 P:g
 =D:v were
 =H:v heading
 A:g
 =H:prp towards
 =D:g
 ==D:art the
 ==D:n tube
 ==H:n stations

We enjoy good wine

STA:cl
 S:pron We
 P:v enjoy
 Od:g
 =D:adj good
 =H:n wine

I have always regarded him as my friend

STA:cl
 S:pron I
 P:g-
 =D:v have
 A:adv always
 -P:g
 =H:v regarded
 Od:pron him
 A:g
 =H:prp as

=D:g
==D:pron my
==H:n friend

Where did you come from?

QUE:cl

A:g-

=D:adv Where

P:g-

=D:v did

S:pron you

-P:g

=H:v come

-A:g

=H:prp from

punc:punc\$?

Have you been here before?

QUE:cl

P:g-

=D:v Have

S:pron you

-P:g

=H:v been

A:adv here

A:adv before

punc:punc\$?

Are you giving him enough attention at this stage?

QUE:cl

P:g-

=D:v Are

S:pron you

-P:g

=H:v giving

Oi:pron him

Od:g

=D:pron enough
=H:n attention
A:g
=H:prp at
=D:g
==D:pron this
==H:n stage
punc:punc\$?

I have always accepted him unconditionally

STA:cl
S:pron I
P:g-
=D:v have
A:adv always
-P:g
=H:v accepted
Od:pron him
A:adv unconditionally

I made the house clean

STA:cl
S:pron I
P:v made
Od:g
=D:art the
=H:n house
Co:adj clean

His youngest daughter gave him an unusual present for his birthday

STA:cl
S:g
=D:pron His
=D:adj youngest

=H:n	daughter
P:v	gave
Oi:pron	him
Od:g	
=D:art	an
=D:adj	unusual
=H:n	present
A:g	
=H:prp	for
=D:g	
==D:pron	his
==H:n	birthday

It was good to see you

STA:cl	
Sf:pron	It
P:v	was
Cs:adj	good
S:cl	
=P:g	
==D:infm	to
==H:v	see
=Od:pron	you

Generally, it is a question of money

STA:cl	
A:adv	Generally
S:pron	it
P:v	is
Cs:g	
=D:art	a
=H:n	question
=D:g	
==H:prp	of
==D:n	money

He plays the violin exceptionally well

STA:cl
S:pron He
P:v plays
Od:g
=D:art the
=H:n violin
A:g
=D:adv exceptionally
=H:adv well

To me it seemed like a difficult decision

A:g
=H:prp To
=D:pron me
S:pron it
P:v seemed
A:g
=H:prp like
=D:g
==D:art a
==D:adj difficult
==H:n decision

To me it seemed a difficult decision

A:g
=H:prp To
=D:pron me
S:pron it
P:v seemed
Cs:g
=D:art a
=D:adj difficult
==H:n decision

I get angry only occasionally

STA:cl

S:pron I
P:v get
Cs:adj angry
A:g
=D:adv only
=H:adv occasionally

London has everything

STA:cl
S:n London
P:v has
Od:pron everything

London is everything

STA:cl
S:n London
P:v is
Cs:pron everything

The experience made him a bitter man

STA:cl
S:g
=D:art The
=H:n experience
P:v made
Od:pron him
Co:g
=D:art a
=D:adj bitter
=H:n man

Judge Miller's place, it was called

STA:cl
Cs:g
=D:g

==D:n	Judge
==H:n	Miller's
=H:n	place
S:pron	it
P:g	
=D:v	was
=H:v	called

The whole realm was his

STA:cl	
S:g	
=D:art	The
=D:adj	whole
=H:n	realm
P:v	was
Cs:pron	his

Manuel had one besetting sin

STA:cl	
S:n	Manuel
P:v	had
Od:g	
=D:num	one
=D:adj	besetting
=H:n	sin

He loved to play Chinese lottery

STA:cl	
S:pron	He
P:v	loved
Od:cl	

=P:g	
==D:infm	to
==H:v	play
=Od:g	
==D:adj	Chinese
==H:n	lottery

It was true, there were other dogs

STA:cl	
Sf:pron	It
P:v	was
Cs:adj	true
S:cl	
=Sf:pron	there
=P:v	were
=S:g	
==D:pron	other
==H:n	dogs

"How much did the other mug get?" the saloon-keeper demanded

STA:cl	
Od:cl	
=Od:g	
==D:adv	How
==H:pron	much
=P:g-	
==D:v	did
=S:g	
==D:art	the
==D:pron	other
==H:n	mug
=-P:g	
==H:v	get
S:g	
=D:art	the
=H:n	saloon-keeper

P:v demanded

There even remained in his remoter crannies some relics of the boy

STA:cl

Sf:pron There

A:adv even

P:v remained

A:g

=H:prp in

=D:g

==D:pron his

==D:adj remoter

==H:n crannies

S:g

=D:pron some

=H:n relics

=D:g

==H:prp of

==D:g

===D:art the

===H:n boy

The hoarse shriek of a locomotive whistling a crossing told him

where he was

STA:cl

S:g

=D:art The

=D:adj hoarse

=H:n shriek

=D:g

==H:prp of

==D:g

===D:art a

===H:n locomotive

===D:cl

====P:v whistling

====Od:g

====D:art	a
====H:n	crossing
P:v	told
Oi:pron	him
Od:cl	
=A:adv	where
=S:pron	he
=P:v	was

I have always been convinced that he was innocent

STA:cl	
S:pron	I
P:g-	
=D:v	have
A:adv	always
-P:g	
=D:v	been
=H:v	convinced
Od:cl	
=SUB:conj	that
=S:pron	he
=P:v	was
=Cs:adj	innocent

Since you are the expert, you might stay to help.

STA:cl	
A:cl	
=SUB:conj	Since
=S:pron	you
=P:v	are
=Cs:g	
==D:art	the
==H:n	expert
S:pron	you
P:g	
=D:v	might
=H:v	stay

A:g
=D:infm to
=H:v help

I know it will be enjoyable

STA:cl
S:pron I
P:v know
Od:cl
=S:pron it
=P:g
==D:v will
==H:v be
=Cs:adj enjoyable

To create new design demands much inspiration

STA:cl
S:cl
=P:g
==D:infm To
==H:v create
=Od:g
==D:adj new
==H:n design
P:v demands
Od:g
=D:pron much
=H:n inspiration

It is no use denying the truth

STA:cl
Sf:pron It
P:v is
Cs:g
=D:pron no

=H:n	use
S:cl	
=P:v	denying
=Od:g	
==D:art	the
==H:n	truth

Because it is so obvious, nobody has thought about it

STA:cl	
A:cl	
=SUB:conj	Because
=S:pron	it
=P:v	is
=Cs:g	
==D:adv	so
==H:adj	obvious
S:pron	nobody
P:g	
=D:v	has
=H:v	thought
A:g	
=H:prp	about
=D:pron	it

I only want to see you work

STA:cl	
S:pron	I
A:adv	only
P:v	want
Od:cl	
=P:g	
==D:infm	to
==H:v	see
=Od:cl	
==S:pron	you
==P:v	work

Have you been standing here for a long time?

QUE:cl

P:g-

=D:v Have

S:pron you

-P:g

=D:v been

=H:v standing

A:adv here

A:g

=H:prp for

=D:g

==D:art a

==D:adj long

==H:n time

punc:punc \$?

I have an annoying habit of asking questions

STA:cl

S:pron I

P:v have

Od:g

=D:art an

=D:adj annoying

=H:n habit

=D:g

==H:prp of

==D:cl

===P:v asking

===Od:n questions

Everything considered, the Germans were very clever to lose that war

STA:cl

A:cl

=S:pron	Everything
=P:v	considered
S:g	
=D:art	the
=H:n	Germans
P:v	were
Cs:g	
=H:g	
==D:adv	very
==H:adj	clever
=D:cl	
==P:g	
===D:infn	to
===H:v	lose
==Od:g	
===D:pron	that
===H:n	war

If you're asking me whether we should proceed with this, I think not

A:cl	
=SUB:conj	If
=S:pron	you
=P:g	
==D:v	're
==H:v	asking
=Oi:pron	me
=Od:cl	
==SUB:conj	whether
==S:pron	we
==P:g	
===D:v	should
===H:v	proceed
==A:g	
===H:prp	with
===D:pron	this
S:pron	I

P:v think
Od:adv not

Most perceptive people consider Alice better at practical tasks

STA:cl
S:g
=D:pron Most
=D:adj perceptive
=H:n people
P:v consider
Od:n Alice
Co:g
=H:adj better
=D:g
==H:prp at
==D:g
===D:adj practical
===H:n tasks

I don't know if you'll go to prison, but I hope so

STA:par
CJT:cl
=S:pron I
=P:g-
==D:v do
=A:adv n't
=-P:g
==H:v know
=Od:cl
==SUB:conj if
==S:pron you
==P:g
===D:v 'll
===H:v go
==A:g
===H:prp to

===D:n	prison
CO:conj	but
CJT:cl	
=S:pron	I
=P:v	hope
=Od:pron	so

The Americans appear to be planning to resume negotiations soon

STA:cl	
S:g	
=D:art	The
=H:n	Americans
P:v	appear
Cs:cl	
=P:g	
==D:infm	to
==D:v	be
==H:v	planning
=Od:cl	
==P:g	
===D:infm	to
===H:v	resume
==Od:n	negotiations
==A:adv	soon

I cheered the end of the war and the coming of peace, as did everyone

else	
STA:cl	
S:pron	I
P:v	cheered
Od:par	
=CJT:g	
==D:art	the
==H:n	end
==D:g	
===H:prp	of

===D:g
 ====D:art the
 ====H:n war
 =CO:conj and
 =CJT:g
 ==D:art the
 ==H:n coming
 ==D:g
 ===H:prp of
 ===D:n peace
 A:cl
 =SUB:conj as
 =P:v did
 =S:g
 ==H:pron everyone
 ==D:adj else

But Buck was neither house-dog nor kennel-dog

STA:cl
 A:adv But
 S:n Buck
 P:v was
 Cs:par
 =CO:conj neither
 =CJT:n house-dog
 =CO:conj nor
 =CJT:n kennel-dog

Claude was a successful producer who lived in a big house that was an exact reproduction of the old Dupuy mansion near Biloxi

STA:cl
 S:n Claude
 P:v was
 Cs:g

=D:art a
 =D:adj successful
 =H:n producer
 =D:cl
 ==S:pron who
 ==P:v lived
 ==A:g
 ===H:prp in
 ===D:g
 ====D:art a
 ====D:adj big
 ====H:n house
 ====D:cl
 =====S:pron that
 =====P:v was
 =====Cs:g
 =====D:art an
 =====D:adj exact
 =====H:n reproduction
 =====D:g
 =====H:prp of
 =====D:g
 =====D:art the
 =====D:adj old
 =====H:g
 =====D:n Dupuy
 =====H:n mansion
 =====D:g
 =====H:prp near
 =====D:n Biloxi

Todd wondered if it might not be true that actors try too hard to please their critics

STA:cl
 S:n Todd
 P:v wondered

Od:cl	
=SUB:conj	if
=Sf:pron	it
=P:g-	
==D:v	might
=A:adv	not
--P:g	
==H:v	be
=Cs:adj	true
=S:cl	
==SUB:conj	that
==S:n	actors
==P:v	try
==A:g	
===D:adv	too
===H:adv	hard
==Od:cl	
===P:g	
====D:infm	to
====H:v	please
===Od:g	
====D:pron	their
====H:n	critics

While he shaved he thought about the fact that since the moment when he had brought her the news of Alun's death Gwen had not mentioned him in any way

STA:cl	
A:cl	
=A:adv	While
=S:pron	he
=P:v	shaved
S:pron	he
P:v	thought
A:g	
=H:prp	about

=D:g
 ==D:art the
 ==H:n fact
 ==D:cl
 ===SUB:conj that
 ===A:g
 =====H:prp since
 =====D:g
 =====D:art the
 =====H:n moment
 =====D:cl
 =====A:adv when
 =====S:pron he
 =====P:g
 =====D:v had
 =====H:v brought
 =====Oi:pron her
 =====Od:g
 =====D:art the
 =====H:n news
 =====D:g
 =====H:prp of
 =====D:g
 =====D:n Alun's
 =====H:n death
 ===S:n Gwen
 ===P:g-
 =====D:v had
 ===A:adv not
 ===-P:g
 =====H:v mentioned
 ===Od:pron him
 ===A:g
 =====H:prp in
 =====D:g
 =====D:pron any
 =====H:n way

Having grown up in Ireland, a country whose history is stained with blood, I have always needed to believe in the force of forgiveness

STA:cl
A:cl
=P:g
==D:v Having
==H:g
===H:v grown
===D:adv up
=A:g
==H:prp in
==D:g
===H:n Ireland
===D:g
====D:art a
====H:n country
====D:cl
=====S:g
=====D:pron whose
=====H:n history
=====P:g
=====D:v is
=====H:v stained
=====A:g
=====H:prp with
=====D:n blood
S:pron I
P:g-
=D:v have
A:adv always
-P:g
=H:v needed
Od:cl
=P:g
==D:infn to
==H:v believe
=A:g

==H:prp in
 ==D:g
 ===D:art the
 ===H:n force
 ===D:g
 ====H:prp of
 ====D:n forgiveness

Having received the encouraging test results, he had started telling people who asked him how he was feeling that he had felt all right for the last few days

STA:cl
 A:cl
 =P:g
 ==D:v Having
 ==H:v received
 =Od:g
 ==D:art the
 ==D:adj encouraging
 ==H:g
 ===D:n test
 ===H:n results
 S:pron he
 P:g
 =D:v had
 =H:v started
 Od:cl
 =P:v telling
 =Oi:g
 ==H:n people
 ==D:cl
 ===S:pron who
 ===P:v asked
 ===Oi:pron him
 ===Od:cl
 ====A:adv how
 ====S:pron he

=====P:g
 =====D:v was
 =====H:v feeling
 =Od:cl
 ==SUB:conj that
 ==S:pron he
 ==P:g
 ===D:v had
 ===H:v felt
 ==Cs:g
 ===D:pron all
 ===H:adj right
 ==A:g
 ===H:prp for
 ===D:g
 =====D:art the
 =====D:adj last
 =====D:pron few
 =====H:n days

A heroin addict is fully aware that he or she is often acting in a manner which is in direct contradiction to rational judgement

STA:cl
 S:g
 =D:art A
 =D:n heroin
 =H:n addict
 P:v is
 Cs:g
 =D:adv fully
 =H:adj aware
 =D:cl
 ==SUB:conj that
 ==S:par
 ===CJT:pron he
 ===CO:conj or
 ===CJT:pron she

==P:g-
 ===D:v is
 ==A:adv often
 ==-P:g
 ===H:v acting
 ==A:g
 ===H:prp in
 ===D:g
 =====D:art a
 =====H:n manner
 =====D:cl
 =====S:pron which
 =====P:v is
 =====A:g
 =====H:prp in
 =====D:g
 =====D:adj direct
 =====H:n contradiction
 =====D:g
 =====H:prp to
 =====D:g
 =====D:adj rational
 =====H:n judgement

Dad liked to eat fish and chips whenever he visited England, and I grew fond of it too, although the fearsomely acid pickled onion was something I still denied myself

STA:par
 CJT:cl
 =S:n Dad
 =P:v liked
 =Od:cl
 ==P:g
 ===D:infm to
 ===H:v eat
 ==Od:par
 ===CJT:n fish

===CO:conj and
 ===CJT:n chips
 =A:cl
 ==SUB:conj whenever
 ==S:pron he
 ==P:v visited
 ==Od:n England
 CO:conj and
 CJT:cl
 =S:pron I
 =P:v grew
 =Cs:g
 ==H:adj fond
 ==D:g
 ===H:prp of
 ===D:pron it
 =A:adv too
 =A:cl
 ==SUB:conj although
 ==S:g
 ===D:art the
 ===D:g
 ====D:adv fearsomely
 ====H:adj acid
 ===D:adj pickled
 ===H:n onion
 ==P:v was
 =Cs:g
 ===H:pron something
 ===D:cl
 =====Od:0 [which]
 =====S:pron I
 =====A:adv still
 =====P:v denied
 =====Oi:pron myself

Had his failure in school been caused by the fact that most of the subjects were of no interest to him, he would have worked hard in those that did interest him

STA:cl
 A:cl
 =P:g-
 ==D:v Had
 =S:g
 ==D:pron his
 ==H:n failure
 ==D:g
 ===H:prp in
 ===D:n school
 =-P:g
 ==D:v been
 ==H:v caused
 =A:g
 ==H:prp by
 ==D:g
 ===D:art the
 ===H:n fact
 ===D:cl
 ====SUB:conj that
 =====S:g
 =====H:pron most
 =====D:g
 =====H:prp of
 =====D:g
 =====D:art the
 =====H:n subjects
 =====P:v were
 =====Cs:g
 =====H:prp of
 =====D:g
 =====D:pron no
 =====H:n interest
 =====D:g
 =====H:prp to
 =====D:pron him
 S:pron he
 P:g

=D:v would
 =D:v have
 =H:v worked
 A:adv hard
 A:g
 =H:prp in
 =D:g
 ==H:pron those
 ==D:cl
 ===S:pron that
 ===P:g
 ====D:v did
 ====H:v interest
 ===Od:pron him

Jeremy told Mathew he knew Sophie, but omitted to mention that she had been called a traitor by several delegates and that she had come to see him

STA:par
 CJT:cl
 =S:n Jeremy
 =P:v told
 =Oi:n Matthew
 =Od:cl
 ==S:pron he
 ==P:v knew
 ==Od:n Sophie
 CO:conj but
 CJT:cl
 =S:0 [he]
 =P:v omitted
 =Od:cl
 ==P:g
 ===D:infrm to
 ===H:v mention
 ==Od:par
 ===CJT:cl
 ====SUB:conj that

=====S:pron she
 =====P:g
 =====D:v had
 =====D:v been
 =====H:v called
 =====Cs:g
 =====D:art a
 =====H:n traitor
 =====A:g
 =====H:prp by
 =====D:g
 =====H:pron several
 =====D:g
 =====H:prp of
 =====D:g
 =====D:art the
 =====H:n delegates
 =====CO:conj and
 =====CJT:cl
 =====SUB:conj that
 =====S:pron she
 =====P:g
 =====D:v had
 =====H:v come
 =====A:cl
 =====P:g
 =====D:infm to
 =====H:v see
 =====Od:pron him

B. Quantified sentence corpus

The 57 sentences in the sentence corpus contain the following number of functions and forms

STA:cl 50
 STA:par 3
 QUE:cl 4
 CJT: 16

CO:	9
SUB:	17
H:	196
D:	306
S:	123
Sf:	6
P:	127 (of these –P:14; i.e. discontinuous predicator)
Oi:	9
Od:	58
Cs:	28
Co:	12
A:	125 (of these –A:1; i.e. discontinuous adverbial)
:cl	59 (not including STA:cl)
:g	211 (of these :g- 15; i.e. discontinuous group)
:par	5 (not including STA:par)
:n	119
:v	153
:adj	43
:adv	42
:pron	115
:prp	46
:conj	26
:art	50
:infm	13
:num	1

Appendix II Key to VISL symbol set¹²²

Functions (specified with CAPITAL letters)

Utterance (or discourse) Functions: UTT, STA, QUE, COM, EXC, PER.

Basic Clause Functions: S, P, O, A, C.

Additional Clause Functions:

Types of Subject: S, Sf.

Types of Object: Of, Od, Oi.

Types of Complement: Cs, Co.

Subordinate Clauses: SUB.

Compound Unit Functions (paratagma): CJT, CO

Forms (specified with small letters)

Clause (cl)

Group (g)

Compound unit (cu)/Paratagma (par)

Word classes (n, v, adj, adv, art, pron, prp, conj, infm. num, intj)

¹²² Source: www.beta.visl.sdu.dk

Appendix III Pre-test

PRE-TEST

Denne test er del af et ph.d.-projekt, og den vil udelukkende blive anvendt til forsknings- formål. Den vil til enhver tid blive behandlet som fortrolig information.

I testen er ordene "acceptable" og "unacceptable" anvendt i henhold til, hvad der i standard skriftsprog anses for værende acceptabelt/uacceptabelt, og som traditionelt er udtrykt i de grammatiske regler, studerende af engelsk som fremmedsprog stifter bekendtskab med i løbet af studiet.

Det er meget vigtigt, at testen besvares samvittighedsfuldt. Du må gerne vælge "Don't know", hvis du er i tvivl. Følg i øvrigt instruktionen til hvert enkelt spørgsmål, da spørgsmålene varierer.

Hvis du har brug for det, er der en ordliste på sidste side.

NAVN: _____

CPR.NUMMER (de første 6 cifre er nok): _____

1. A) The boy and his father like to eat at Macdonald's.
B) The boy and his father likes to eat at Macdonald's.

Choose one of the following possibilities:

- a) Sentence A is acceptable
- b) Sentence B is acceptable
- c) Both A and B are acceptable
- d) Neither A nor B is acceptable
- e) Don't know

2. A) I understood he might be in some kind of trouble.
B) I understood that he might be in some kind of trouble.

Choose one of the following possibilities:

- a) Sentence A is acceptable
- b) Sentence B is acceptable
- c) Both A and B are acceptable
- d) Neither A nor B is acceptable
- e) Don't know

3. A) The ice was more thick than I expected it to be.
B) The ice was thicker than I expected it to be.

Choose one of the following possibilities:

- a) Sentence A is acceptable
- b) Sentence B is acceptable
- c) Both A and B are acceptable
- d) Neither A nor B is acceptable
- c) Don't know

4. Do the words SPEAK, COMPANION, BELIEVED, HAPPENED belong to the same word class ?

- a) Yes
- b) No
- c) Don't know

5. One of the sentences below contains a complement to the subject. Choose one of the following possibilities:

- a) "I don't know the answer to that very relevant question," said the teacher.
- b) This engine lets you access the various Webster's dictionary services on the internet.
- c) A society armed to the teeth but with clear lungs may be a worthy aim.
- d) Don't know

6. A) Graham Greene, the late novelist, was imprisoned for spying in a South American Country.
B) Graham Greene, the late novelist, has been imprisoned for spying in a South American country.

Choose one of the following possibilities:

- a) Sentence A is acceptable
- b) Sentence B is acceptable
- c) Both A and B are acceptable
- d) Neither A nor B is acceptable
- e) Don't know

7. The sentence below contains a direct object. Underline this object.

"He realised he was mistaken, but he didn't apologise".

8. Do the words UNLIKELY, MOTHERLY, BELIEVABLE, INTERESTING belong to the same word class ?

- a) Yes
- b) No
- c) Don't know

9. A) The boys, who makes a great team, are all great ball-players.
B) The boys, who make a great team, are all great ball-players.

Choose one of the following possibilities:

- a) Sentence A is acceptable
- b) Sentence B is acceptable
- c) Both A and B are acceptable
- d) Neither A nor B is acceptable
- e) Don't know

10. A) I have in Magasin bought a nice jacket.
B) I have bought in Magasin a nice jacket.

Choose one of the following possibilities:

- a) Sentence A is acceptable
- b) Sentence B is acceptable
- c) Both A and B are acceptable
- d) Neither A nor B is acceptable
- e) Don't know

11. Underline all the nouns in the text below:

"Where they live there are no mountains. But there are cold days when there is ice on the trees and it is very cold. On cold afternoons, as he walks with his father, the boy's nose and ears are red with the cold and he thinks about high mountains, white with ice and snow".

12. Which word class do FROM, WITH, TOWARDS belong to ?

- a) Verb
- b) Noun
- c) Adjective
- d) Adverb
- e) Pronoun
- f) Preposition
- g) Conjunction
- h) Article
- e) Don't know

13. One of the sentences below contains a direct object. Choose from the following possibilities:

- a) The boy's father is a quiet man
- b) A lot of the people have children
- c) When his father talks, the boy listens
- e) Don't know

14. In the text below underline all the co-ordinating conjunctions:

"A long train journey on a late December evening, in this new version of peace, is a dreary experience. I suppose that my fellow traveller and I could consider ourselves lucky to have a compartment to ourselves, even though the heating apparatus was not working, even though the lights went out entirely in the frequent Pennine tunnels and were too dim anyway for us to read our books without straining the eyes, and though there was no restaurant car to give at least a change of scene".

15. Only one of the sentences in the text below contains an indirect object. Choose from the following possibilities:

- a) He gave the exercise a lot of thought.
- b) The eclipse of the sun caused quite a commotion.
- c) I never liked the way she writes her novels.
- d) Don't know

16. Underline the subject of the main clause in the sentence below:

"I know people who live in the same house and have been in the same job for twenty, thirty, forty years, and who would hate to pull up their roots and change to something new".

17. Underline all the adjectives in the text below:

"At one o'clock they went down to the restaurant for their lunch.
'It's raining harder now,' said Bill as he looked out of the window.
'Good,' said John. 'It's going to be warmer this afternoon'.
'Perhaps', said Bill. 'It's raining, but there's still ice on the ground, and that's dangerous'.
'That's true', said John. 'Drink your coffee. We must get back now'.
'I'm ready'".

18. Which word class do NOW, HARDLY, HERE belong to ?

- a) Verb
- b) Noun
- c) Adjective
- d) Adverb
- e) Pronoun
- f) Preposition
- g) Conjunction
- h) Article
- i) Don't know

19. Underline all the subjects in the sentences below:

"Dickens's extreme energy was not exhausted by his unique success as a novelist. His weekly journalism made heavy demands on his time after 1850, and he constantly turned to the stage; first in many amateur theatricals, given privately or for charity; later, in his public readings".

20. Underline all the pronouns in the text below:

"WAITER (taking up a large spoon): 'Why, a plum pudding is my favourite pudding. Isn't that lucky. Come on , boy, let's see who gets most'. (They both eat, DAVID with his small spoon and rather slowly; the WAITER with his tablespoon and very fast.) 'Come on, you're getting behind'".

21. Only one of the sentences below is acceptable. Choose one of the following possibilities:

- a) Johnson the ball kicked quick.
- b) Johnson kicked the ball quick.
- c) The ball quickly Johnson kicked.
- d) The ball Johnson kicked quick.
- e) Quickly Johnson kicked the ball.
- f) Don't know

22. Underline all the adverbials in the text below:

"The boy and his parents live in New Street. The street runs down a long hill from Station Road. There are always lots of people in Station Road, but New Street is usually a quiet Street and the people living there are quiet people. There are gardens with trees in them and on warm evenings the people of New Street talk to the people in the houses next to theirs".

23. Only one of the sentences below is correct. Choose one of the following possibilities:

- a) I live in Paris till Christmas because I want to improve my French accent.
- b) My cat is licking its paws every day to keep itself clean.
- c) The old man is taking a nap in his chair every afternoon.
- d) The girl spoke to the people who were waiting at the airport.
- e) My grandparents are talking to my mother on the phone whenever they feel lonely.
- f) Don't know

24. One of the sentences below contains an adjective used as a noun. Underline the word in question.

"The evolution of man seems to be based on natural selection, or as Darwin called it: the survival of the fittest. Nowadays, however, modern science has almost eliminated that principle".

25. A) He knows her real well
B) He knows her real good

Choose one of the following possibilities:

- a) Sentence A is acceptable
- b) Sentence B is acceptable
- c) Both A and B are acceptable
- d) Neither A nor B is acceptable
- e) Don't know

Word list:

Adjective/adjektiv/tillægsord

Adverb/adverbium/biord

Adverbial/adverbialled/ adverbielt sætningsled

Article/artikel/kendeord

Complement to subject/subjektsprædikat/omsagnsled til grundled

Conjunction/konjunktion/bindeord

Co-ordinating conjunction/sideordnende bindeord

Direct object/direkte objekt/genstandsled

Indirect object/indirekte objekt/hensynsled

Main clause/hovedsætning

Noun/substantiv/navneord

Preposition/præposition/forholdsord

Pronoun/pronomen/stedord

Subject/subjekt/grundled

Verb/verbum/udsagnsord

Appendix IV Post-test

POST -TEST

Denne test er del af et ph.d.-projekt, og den vil udelukkende blive anvendt til forsknings- formål. Den vil til enhver tid blive behandlet som fortrolig information.

I testen er ordene "**acceptable**" og "**unacceptable**" anvendt i henhold til, hvad der i standard skriftsprog anses for værende acceptabelt/uacceptabelt, og som traditionelt er udtrykt i de grammatiske regler, studerende af engelsk som fremmedsprog stifter bekendtskab med i løbet af studiet.

Det er meget vigtigt, at testen besvares samvittighedsfuldt. Du må gerne vælge "Don't know", hvis du er i tvivl. Følg i øvrigt instruktionen til hvert enkelt spørgsmål, da spørgsmålene varierer.

Som opfølgning på min egen undersøgelse vil jeg gerne kunne følge jer helt til eksamen. Det vil være af stor betydning for vurderingen af mine resultater, hvis de kunne bekræftes eller afkræftes af jeres eksamensresultater. Derfor vil jeg bede jer oplyse oplyse jeres eksamensnummer. Jeg kan forsikre jer, at det kun vil blive brugt til forskning. Og det vil kun være mig, der ser det. Jeg håber på jeres forståelse.

Hvis du har brug for det, er der en ordliste på sidste side.

NAVN: _____

CPR.NUMMER (de første 6 cifre er nok): _____

EKSAMENSNUMMER: _____

1. A) The book and its author are well-known by the public.
B) The book and its author is well-known by the public.

Choose one of the following possibilities:

- a) Sentence A is acceptable
- b) Sentence B is acceptable
- c) Both A and B are acceptable
- d) Neither A nor B is acceptable
- e) Don't know

2. A) Most people believe that it is wrong to lie.
B) Most people believe it is wrong to lie.

Choose one of the following possibilities:

- a) Sentence A is acceptable
- b) Sentence B is acceptable
- c) Both A and B are acceptable
- d) Neither A nor B is acceptable
- e) Don't know

3. A) The Mississippi river is more wide than the Potomac.
B) The Mississippi river is wider than the Potomac.

Choose one of the following possibilities:

- a) Sentence A is acceptable
- b) Sentence B is acceptable
- c) Both A and B are acceptable
- d) Neither A nor B is acceptable
- e) Don't know

4. Do the words ENQUIRE, ETIQUETTE, FLEW, DIED belong to the same word class ?

- a) Yes
- b) No
- c) Don't know

5. One of the sentences below contains a complement to the subject. Choose one of the following possibilities:

- a) After having recovered from his long illness, he became a different person.
- b) He smiled as he interrupted her.
- c) When the palace clock struck twelve, Lady Bothwell resolutely got up and left.
- d) Don't know

6. A) In the Battle of Waterloo the British forces have defeated Napoleon.
B) In the Battle of Waterloo the British forces defeated Napoleon.

Choose one of the following possibilities:

- a) Sentence A is acceptable
- b) Sentence B is acceptable
- c) Both A and B are acceptable
- d) Neither A nor B is acceptable
- e) Don't know

7. The sentence below contains a direct object. Underline this object.

"He knew he had to act, but he was not ready yet ".

8. Do the words UNFRIENDLY, HOLY, FASCINATING, IMPOSSIBLE belong to the same word class ?

- a) Yes
- b) No
- c) Don't know

9. A) The two features of life in Britain that gives visitors a bad impression are the weather and the food.
B) The two features of life in Britain that give visitors a bad impression are the weather and the food.

Choose one of the following possibilities:

- a) Sentence A is acceptable
- b) Sentence B is acceptable
- c) Both A and B are acceptable
- d) Neither A nor B is acceptable
- e) Don't know

10. A) I have in the bank deposited a large sum of money.
B) I have deposited in the bank a large sum of money.

Choose one of the following possibilities:

- a) Sentence A is acceptable
- b) Sentence B is acceptable
- c) Both A and B are acceptable
- d) Neither A nor B is acceptable
- e) Don't know

11. Underline all the nouns in the text below:

"The man who intends to buy a car will also have to take a good look at its general appearance, the lines of the body, mudguards, and fenders. His wife and children are likely to demand a certain elegance of design if not streamlined smartness".

12. Which word class do IN, OF, AT, AGAINST belong to ?

- a) Verb
- b) Noun
- c) Adjective
- d) Adverb
- e) Pronoun
- f) Preposition
- g) Conjunction
- h) Article
- i) Don't know

13. One of the sentences below contains a direct object. Choose from the following possibilities:

- a) The Englishman's home is his castle.
- b) Americans love change.
- c) I listened patiently to his story.
- d) Don't know

14. In the text below underline all the co-ordinating conjunctions:

"I'll admit that there is not a prettier sight than a cricket ground – with the grass so very green and well rolled, and the players in their spotless white, but I'd never have believed that a sporting nation like the English would be content with anything so slow for their national game".

15. Only one of the sentences in the text below contains an indirect object. Choose from the following possibilities:

- a) It is hard to prove the connection between free trade and growth.
- b) The President of the United States gave Congress his full support.
- c) Brazil has abandoned government policies that encourage deforestation.
- d) Don't know

16. Underline the subject of the main clause below:

"Everyone knows that the list of accessories that add to the comfort of the passengers of a modern car is almost endless".

17. Underline all the adjectives in the text below:

"He was a corporal, I think, or perhaps a sergeant, a gay, dashing sort of fellow with dark, romantic-looking eyes and black curly hair. He didn't speak much English, but that seemed to be no obstacle to his popularity with the girls of Manchester, and it was soon quite clear that Irene had eyes for no one except Ruperto."

18. Which word class do NOT, SURELY, WHERE, REALLY belong to ?

- a) Verb
- b) Noun
- c) Adjective
- d) Adverb
- e) Pronoun
- f) Preposition
- g) Conjunction
- h) Article
- i) Don't know

19. Underline all the subjects of the sentences below:

" The Johnsons got married just after World War II. The wedding took place at the home of the bride, and for days the friends and family of the bride and groom had been busy decorating the house and preparing the food".

20. Underline all the pronouns in the text below:

HOB: "What is *Pickwick Papers* about ?

MR. PRIESTLY: "You ought to get the book. It's about Mr. Pickwick and his three friends and the absurd situations they get themselves into, and, yet, though we laugh at Mr. Pickwick, we don't think any the worse of him for being a figure of fun".

21. Only one of the sentences below is acceptable. Choose one of the following possibilities:

- a) Extremely badly the manager tackled the problem.
- b) The manager extremely badly tackled the problem.
- c) The manager tackled extremely badly the problem.
- d) The manager tackled the problem extremely badly.
- e) Extremely bad the manager tackled the problem.
- f) Don't know

22. Underline all the adverbials in the text below:

"My grandfather, my mother's father – he was a Scotsman – left me a sum of money in his will, and some useless property, a factory; but until I am twenty-five the money is in trust. However, I get the income from it. It will never support me fully, but it's always a welcome supplement. It's difficult to make ends meet on my low salary".

23. Only one of the sentences below is correct. Choose one of the following possibilities:

- a) Every Christmas we were decorating the Christmas tree in the same way.
- b) Unlike our parents, we are having too many possibilities to choose from.
- c) Whenever I'm in New York, I am making a point of visiting the Village.
- d) When I looked up, the teacher was smiling at me.
- e) Porter, please tell Mr. Longfellow that I wait for him here in the lobby.
- f) Don't know

24. One of the sentences below contains an adjective used as a noun. Underline the word in question.

"The vicar spoke most eloquently about the accomplishments of the dead although he had never actually known any of them. The widows were most grateful and later showed their appreciation in a manner befitting the situation".

25. A) Hardly he spoke to her.
B) He spoke hardly to her.

Choose one of the following possibilities:

- a) Sentence A is acceptable
- b) Sentence B is acceptable
- c) Both A and B are acceptable
- d) Neither A nor B is acceptable
- e) Don't know

Word list:

Adjective/adjektiv/tillægsord

Adverb/adverbium/biord

Adverbial/adverbialled/ adverbielt sætningsled

Article/artikel/kendeord

Complement to subject/subjektsprædikat/omsagnsled til grundled

Conjunction/konjunktion/bindeord

Co-ordinating conjunction/sideordnende bindeord

Direct object/direkte objekt/genstandsled

Indirect object/indirekte objekt/hensynsled

Main clause/hovedsætning

Noun/substantiv/navneord

Preposition/præposition/forholdsord

Pronoun/pronomen/stedord

Subject/subjekt/grundled

Verb/verbum/udsagnsord

Appendix V Overview of entrance levels A and B

Level A

Description	Gymnasium	HF	HTX	HHX
Linguistic awareness	x	x	x	x
Linguistic knowledge		x	x	x
Acquisition of grammar		x	x	x
Sociolects				x
Level of proficiency: command of English:				
fluent	x	x	x	
assured	x	x	x	
varied			x	x
precise			x	
nuanced	x	x		x
idiomatic	x	x		
stylistically correct	x	x		
Word classes:				
understand	x	x		
master	x	x		
use	x	x		
no mention			x	x
Morphology:				
understand	x	x		
master	x	x		
use	x	x		
no mention			x	x
Syntax:				
understand	x	x		
master	x	x		
use	x	x		
no mention			x	x
Lexis:				
understand in text	x	x	x	
use in context	x	x		
acquire ESP terms			x	
acquire/expand	x	x		

					no mention
<hr/>					
Phonology:					
					acquire accepted
					variety of pronunc. x x
					phonemically correct x x
					no mention x
<hr/>					
Exam form:					
Oral		30min	30 min	30 min	30min
					convers.Eng x x x x
					read aloud Eng. x x x
					summarise x x x
					present written
					assignment in Danish in Danish
					translate Da-Eng
					Eng-Da x x x
<hr/>					
Written		5 hours	5 hours	4 hours	5 hours
					comment on text x x
					essay x x x x
					summary of text x
					translation Da-Eng x x x x
					Eng-Da
<hr/>					
Exam evaluation:					
Oral					
					ability to communicate x x x
					linguistic correctness x x x
					general proficiency x x x x
					intonation x x
					diction x x
					fluency x x
					lexis x x
					comprehension of text & context x x x x
<hr/>					
Written					
					degree of correctness x x x
					in...ortography x x
					morphology x x
					syntax x x
					semantics x x x
					idiomatic express. x x
					general proficiency x x
					comprehension of text &

context	x	x
CALL		x

Notes:

1) HF indicates the advanced level (Højniveau) and equals the 'gymnasieniveau';
cf. Gymnasiebekendtgørelsen.

LEVEL B

Description	Gymnasium	HF	HTX	HHX
Linguistic awareness	x	x		x
Linguistic knowledge		x	x	x
Acquisition of grammar		x	x	x
Wordclasses:				
understand	x			
master	x			
use				
no mention		x	x	x
Morphology:				
understand	x			
master	x			
use				
no mention		x	x	x
Syntax:				
understand	x			
master	x			
use	x		x	
no mention		x		x
Lexis:				
understand in text	x			
use in context	x			
acquire ESP terms			x	
acquire/expand	x			x
no mention		x		
Phonology:				
acquire accepted variety of pronunc.	x			
Exam form:	25min	25min	30min	30min
Oral	x	x	x	x
convers.Eng	x	x	x	x
read aloud Eng.	x	x	x	
summarise			in Danish	
present written assignment			in Danish	in Danish
translate	Da-Eng			

	Eng-Da	x		x	
Written		4 hours	4 hours	no	4 hours
	comment on text	x	x		x
	summary of text		x		
	essay				business letter
	translation	Da-Eng	x	x	
		Eng-Da			
Exam evaluation:					
Oral					
	ability to communicate				x
	linguistic correctness	x		minor importance	
	general proficiency	x	x	x	x
	intonation	x			
	diction	x			
	fluency	x			
	lexis	x			
	comprehension of text	x	x		x
Written					
	degree of correctness	x	x		
	in...ortography	x	x		
	morphology	x	x		
	syntax	x	x		
	semantics	x	x		
	idiomatic express.	x	x		
	general proficiency	x	x		
	comprehension of text	x	x		
CALL					
			x		x

Appendix VI English supplement to “Gymnasiebekendtgørelsen”

1 Identity and aims

1.1 English is a subject that comprises skills, knowledge and culture. Among the foreign languages, English has a special position as the first foreign language, partly because of its position in the Danish *folkeskole*, the first nine or ten years of school, partly because such a large part of the culture – music, film and television, IT – is in English, and finally because of the role the language has as an international means of communication. In general there is such comprehensive and close contact with the language that as a means of oral communication, it holds a place somewhere between that of the mother tongue and that of a foreign language.

1.2 The area of study in the subject covers a very wide range, but the main content can be described as concentric circles. The innermost circle, which is the core of the instruction, consists of work with British and American texts. In the next circle there is work with other English-speaking cultures and literatures. The outer ring includes the use of the language as an international means of communication. It is central to the instruction in English that the language is not merely a tool, but that language, literature and culture must be dealt with as closely related phenomena.

LANGUAGE LINE

OBLIGATORY LEVEL (B-level)

2 Aims

The aims of the instruction are:

- The students should gain and be able to demonstrate an understanding and a mastery of the use of the parts of speech, systems of conjugation and declension, and syntax in the English language, as well as gain an extensive vocabulary, both active and passive.
- The students should gain skills in writing and speaking English so that it is intelligible, correct, fluent and precise.
- The students should be able to use a type of pronunciation of the English language that is recognisable and accepted.
- The students should gain solid reading skills and general linguistic awareness, through the intensive reading of texts and through other means. This means a method of reading through which the reader gains a precise understanding of all the details of language and the context they are part of, as well as searching for and gaining knowledge about historical, geographical, cultural and social conditions to an extent that gives depth and perspective to the reading of texts.
- The students should gain skills in the extensive reading of texts. This means that the students should be able to gain a sufficient overview of a text to enable them to retell the main aspects of its content and place it in a larger context, to be able to orientate themselves in textual material of some length, as well as to seek relevant background information.

3 Syllabus

3.1 The syllabus comprises 500-700 pages of standard length. None of the texts may

be adapted or simplified. Primarily British and American texts are to be read. Various periods must be represented. Quality texts are to be chosen; however, to a limited extent texts may be chosen without a view to their intrinsic value, provided that they contribute to the understanding of aspects of the English-speaking world. Texts that deal with topics of global interest and/or use English as an international means of communication may likewise be included to a limited extent.

3.2 A varied selection of imaginative literature must be read. Novels, short stories, poetry and drama must be represented. In addition a number of non-fiction texts of various kinds and degrees of difficulty must be read, e.g. newspaper articles, biographies and interviews. At least one complete work of some length must be included in the syllabus.

Other media than the printed word, e.g. films and videos, must be included in the syllabus. Each lesson spent working with these media is converted to six pages of standard length in the syllabus.

3.3 The syllabus must be structured in such a way as to include at least two topics for special study; each topic will ordinarily comprise from three to six texts. A topic may be, for example, a genre or a mass medium; an important cultural, social or political issue; a period; a literary theme, or the study of an author. One of these topics must include a historical perspective.

3.4 The written work is to be carried out partly as homework and partly as a classroom activity. Fifteen free assignments of two to four pages, and twenty-nine assignments of one to two pages are to be submitted for marking and comments.

4 Examination

4.1 There is an oral and a written examination.

4.2 *The oral examination*

4.2.1 The oral examination consists of a test on one of the texts read and on an unseen text. 25 minutes' preparation time is allotted, including the time spent giving instructions and handing out materials. Including the time spent deciding on a final mark, 2.5 candidates are to be examined per hour.

4.2.2 For use in the examination, 125 standard-length pages of intensively read syllabus are selected, which represent the four literary genres and non-fiction, as well as the topics for special study and one complete work of some length. No single text may be represented by more than twenty standard-length pages. Only texts in printed form may be used for the examination. The selected texts may be accompanied by pictorial material that has been studied.

4.2.3 The examination syllabus for self-taught candidates is approximately 600 pages.

4.2.4 The examiner provides unmarked copies of the texts included in the examination syllabus for the preparation room. For the examination itself, an unmarked copy of the text is used.

4.2.5 For the part of the examination covering a text that has been studied, the starting-point is an extract of an intensively read text of about one half a standard-length page. The extract is read aloud, and during a conversation in English the skills gained through intensive reading are tested. The candidate is tested on skills in spoken English and on the precise understanding of

- Important details of content and language in the extract selected and its function in context,

- The text as a whole and the context to which it belongs.

The candidate begins the conversation with an independent presentation of the text. When about one third of the examination time is left, the candidate draws an unseen text of about half a standard-length page and of an appropriate degree of lexical and syntactical difficulty. After briefly reading through the text, the candidate demonstrates his/her comprehension of the English text by translating it into Danish.

4.3 Assessment criteria (oral examination)

4.3.1 The relative degree of difficulty of the studied text as well as of the unseen text must be taken into consideration in the assessment. The following elements are included in the assessment:

Reading aloud

- Comprehension of the text, as expressed in a coherent reading
- Intonation and diction

Presentation and conversation

- Mastery of the systems of the language
- Vocabulary and fluency
- Ability to read intensively and understanding of details
- General overview and comprehension of the text
- Ability to relate the text to other texts
- Ability to interpret and put into perspective in an independent manner
- Ability to enter into a dialogue about the text

The unseen text

- Precision
- General overview and comprehension

4.3.2 One mark is given for the entire oral examination, based on a general assessment, in which the test on the studied text has more weight than the test on the unseen text.

4.4 The written examination

Four hours are allotted for the written examination. The candidates are presented with a text or texts in English, possibly accompanied by pictorial material, and a number of questions are set that test the candidates' command of the English language, their comprehension of the text material, and their general skills in written English. In addition, a Danish text must be translated.

4.5 Assessment criteria (written examination)

4.5.1 In the assessment emphasis is primarily given to the candidates' command of the English language, i.e. the orthography, morphology, syntax, semantics and idioms of the English language. What is assessed is the extent to which the systems of the language have been mastered.

4.5.2 The summary question is furthermore assessed as to how precisely and with what degree of nuance the candidate is able to reproduce the main points of the set text in a piece of concise, coherent writing.

4.5.3 The essay question is furthermore assessed as to the degree of clarity and nuance the candidate is able to demonstrate in characterising and evaluating the text in a longer, well-organised piece of writing. The essay covers the following aspects: the plot of the text, characterisation, argumentation, the author's attitude and the message of the text.

4.5.4 The other, freer questions are furthermore assessed as to the degree of clarity and nuance the candidate is able to demonstrate in giving an account of and evaluating a problem in the text in a brief, coherent piece of writing, which in some cases may be in fictional form.

4.5.5 The translation question is furthermore assessed as to how precisely the candidate is able to render the Danish text with its nuances into English. The impression given by the answer to this question is included in the assessment as a supplement to the impression given by the candidate's command of the language in the answer as a whole.

4.5.6 Finally the assessment also includes an evaluation of the structure and content of the answer as a whole.

4.5.7 One mark is given, based on a general assessment.

LANGUAGE LINE

HIGH LEVEL (A-level)

5.Aims

In addition to the requirements at obligatory level:

- The students should gain and be able to demonstrate a command of the English language with certainty and nuance, actively as well as passively.
- The students should be able to express themselves with fluency, variety and a high degree of idiomatic correctness and precision, as well as with a sense of register.
- The students' pronunciation should become clear and phonemically correct.
- The students should gain skills in analysing texts of greater difficulty in language as well as content, modern as well as older texts, and be able to employ points of view concerning principle and theory in studying texts and topics.
- The students should be able to use their skills in various extensive reading methods in work with texts of greater difficulty in language and content, including texts of a higher level of abstract thought.

6.Syllabus

6.1 The syllabus comprises 400-600 pages of standard length. None of the texts may be adapted or simplified. The texts must make greater demands on the students than those read at obligatory level. Primarily British and American texts are to be read. Various periods must be represented. Quality texts are to be chosen; however, to a limited extent texts may be chosen without a view to their intrinsic value, provided that they contribute to the understanding of aspects of the English-speaking world. Texts that deal with topics of global interest and/or use English as an international means of communication may likewise be included to a limited extent.

6.2 The syllabus must include the following genres: novel, short story and poetry, as well as various types of non-fiction texts, including newspaper articles and essays. A drama by Shakespeare must be included. In addition a number of texts written before 1900 must be included.

Other media than the printed word, e.g. film and videos, must be included in the syllabus. Each lesson spent working with such media is converted to six pages of standard length for the syllabus. Texts read at obligatory level may not be included.

6.3 The syllabus must be structured in such a way as to include one topic for special study; a topic will ordinarily comprise from three to six texts. The topic must include at least one text of a theoretical nature. A topic may be, for example, a genre or a mass medium; an important cultural, social or political issue; a period; a literary theme, or the study of an author's work.

6.4 The written work is to be carried out partly as homework and partly as a classroom activity. Eight free assignments of four to five pages and fifteen assignments of 1-1½ pages are to be submitted for marking and comments.

7.Examination

7.1 There is an oral and a written examination.

7.2 *The oral examination*

7.2.1 The oral examination consists of a test on one of the texts read and on an unseen text. 30 minutes' preparation time is allotted, including the time spent giving instructions and handing out materials. Including the time spent deciding on a final mark, 2 candidates are to be examined per hour.

7.2.2 Students who have attended classes at obligatory level in the language line, or have taken the examination at that level, present for their examination a selection of 125 standard-length pages from texts read intensively. The following must be represented: the three literary genres, the topic for special study (including the theoretical text or texts) and the texts written before 1900. Fifteen to twenty standard-length pages of the Shakespeare play must be included. No single text may be represented by more than twenty standard-length pages. Only texts in printed form may be used for the examination. Texts that have been studied may be accompanied by pictorial material that has been covered in the instruction.

7.2.3 Students who have not attended classes or taken the examination at obligatory level in the language line present an intensively read selection of 125 standard-length pages for their examination, drawn from a studied syllabus of 900 standard-length pages. This is the sum of the syllabuses studied for obligatory and high levels in the language line.

7.2.4 Self-taught candidates present an examination syllabus of about 500 pages.

7.2.5 The examiner provides unmarked copies of the texts included in the examination syllabus for the preparation room. For the examination itself, an unmarked copy of the text is used.

7.2.6 For the part of the examination covering a text that has been studied, the starting-point is an extract of an intensively read text of about three fourths of a standard-length page. The extract is read aloud, and during a conversation in English the skills gained through intensive reading are tested. The candidate is tested on skills in spoken English and in the precise understanding of

- Important details of content and language in the extract selected and its function in context,
- The text as a whole and the context to which it belongs.

The candidate begins the conversation with an independent presentation of the text. When about one third of the examination time is left, the candidate draws an unseen

text of about three fourths of a standard-length page and of an appropriate degree of lexical and syntactical difficulty. After briefly reading through the text, the candidate demonstrates his/her comprehension of the English text by translating it into Danish.

7.3 Assessment criteria (oral examination)

7.3.1 The relative degree of difficulty of the studied text as well as of the unseen text must be taken into consideration in the assessment. The following elements are included in the assessment:

Reading aloud

- Comprehension of the text, as expressed in a coherent reading
- Intonation and diction

Presentation and conversation

- Mastery of the systems of the language
- Vocabulary and fluency
- Ability to read intensively and understanding of details
- General overview and comprehension of the text
- Ability to relate the text to other texts
- Ability to interpret and put into perspective in an independent manner
- Ability to enter into a dialogue about the text

The unseen text

- Precision
- General overview and comprehension

7.3.2 One mark is given for the entire oral examination, based on a general assessment, in which the test on the studied text has more weight than the test on the unseen text.

7.4 The written examination

Five hours are allotted for the written examination. The candidates are presented with a text or texts in English, possibly accompanied by pictorial material, and a number of questions are set that test the candidates' command of the English language, their comprehension of the text material, and their general skills in written English. In addition, a Danish text must be translated.

7.5 Assessment criteria (written examination)

7.5.1 In the assessment emphasis is primarily given to the candidates' command of the English language, i.e. the orthography, morphology, syntax, semantics and idioms of the English language. What is assessed is the extent to which the systems of the language have been mastered.

7.5.2 The summary question is furthermore assessed as to how precisely and with what degree of nuance the candidate is able to reproduce the main points of the set text in a piece of concise, coherent writing.

7.5.3 The essay question is furthermore assessed as to the degree of clarity and nuance the candidate is able to demonstrate in characterising and evaluating the text in a longer, well-organised piece of writing. The essay covers the following aspects:

the plot of the text, characterisation, argumentation, the author's attitude and the message of the text.

7.5.4 The other, freer questions are furthermore assessed as to the degree of clarity and nuance the candidate is able to demonstrate in giving an account of and evaluating a problem in the text in a brief, coherent piece of writing, which in some cases may be in fictional form.

7.5.5 The translation question is furthermore assessed as to how precisely the candidate is able to render the Danish text with its nuances into English. The impression given by the answer to this question is included in the assessment as a supplement to the impression given by the candidate's command of the language in the answer as a whole.

7.5.6 Finally the assessment also includes an evaluation of the structure and content of the answer as a whole.

7.5.7 One mark is given, based on a general assessment.

8. The major written assignment

8.1 The assignment must deal with a topic that lies within the sphere of anglophone culture, as it is reflected in texts. In working on the assignment, the methods of the subject are to be used.

8.2 The primary literature and parts of the secondary literature are to be read in English, so that nearly all quotations in the assignment are in English.

8.3 One mark is given, based on a general assessment, using the criteria in Supplement 35, and based on the following:

- The extent to which the primary literature has been understood
- Whether quotations are relevant and insightful
- Whether quotations are an integrated part of the assignment
- The extent to which the student is able to progress from the simple reproducing of material to analysing and evaluating it.

MATHEMATICS LINE

OBLIGATORY LEVEL (B-level)

9.Aims

The aims of the instruction are:

- The students should gain and be able to demonstrate an understanding and a mastery of the use of the parts of speech, systems of conjugation and declension, and syntax in the English language, as well as achieve an extensive vocabulary, both active and passive.
- The students should gain skills in writing and speaking English so that it is intelligible, correct, fluent and precise.
- The students should be able to use a type of pronunciation of the English language that is recognisable and accepted.
- The students should gain solid reading skills and general linguistic awareness, through the intensive reading of texts and through other means. This means a method of reading through which the reader gains a precise understanding of all the

details of language and the context they are part of, as well as searching for and gaining knowledge about historical, geographical, cultural and social conditions to an extent that gives depth and perspective to the reading of texts.

- The students should gain skills in the extensive reading of texts. This means that the students should be able to gain a sufficient overview of a text to enable them to retell the main aspects of its content and place it in a larger context, to be able to orientate themselves in textual material of some length, as well as to seek relevant background information.

10.Syllabus

10.1 The syllabus comprises 500-700 pages of standard length. None of the texts may be adapted or simplified. Primarily British and American texts are to be read. Various periods must be represented. Quality texts are to be chosen; however, to a limited extent texts may be chosen without a view to their intrinsic value, provided that they contribute to the understanding of aspects of the English-speaking world. Texts that deal with topics of global interest and/or use English as an international means of communication may likewise be included to a limited extent.

10.2 A varied selection of imaginative literature must be read. Novels, short stories, poetry and drama must be represented. In addition of number of non-fiction texts of various kinds and degrees of difficulty must be read, e.g. newspaper articles, biographies and interviews. At least one complete work of some length must be included in the syllabus.

Other media than the printed word, e.g. films and videos, must be included in the syllabus. Each lesson spent working with these media is converted to six pages of standard length in the syllabus.

10.3 The syllabus must be structured in such a way as to include at least two topics for special study; each topic will ordinarily comprise from three to six texts. A topic may be, for example, a genre or a mass medium; an important cultural, social or political issue; a period; a literary theme, or the study of an author. One of these topics must include a historical perspective.

10.4 The written work is to be carried out partly as homework and partly as a classroom activity. Four free assignments of two to four pages and eight short assignments of one to two pages are to be submitted for marking and comments.

11.Examination

11.1 The oral examination consists of a test on one of the texts read and on an unseen text. 25 minutes' preparation time is allotted, including the time spent giving instructions and handing out materials. Including the time spent deciding on a final mark, 2.5 candidates are to be examined per hour.

11.2 For use in the examination, 125 standard-length pages of intensively read syllabus are selected, which represent the four literary genres and non-fiction, as well as the topics for special study and one complete work of some length. No single text may be represented by more than twenty standard-length pages. Only texts in printed form may be used for the examination. The selected texts may be accompanied by pictorial material that has been studied.

11.3 The examination syllabus for self-taught candidates is approximately 600 pages.

11.4 The examiner provides unmarked copies of the texts included in the examination syllabus for the preparation room. For the examination itself, an unmarked copy of the text is used.

11.5 For the part of the examination covering a text that has been studied, the starting-point is an extract of an intensively read text of about one half a standard-length page. The extract is read aloud, and during a conversation in English the skills gained through intensive reading are tested. The candidate is tested on skills in spoken English and in the precise understanding of

- Important details of content and language in the extract selected and its function in context,
- The text as a whole and the context to which it belongs.

The candidate begins the conversation with an independent presentation of the text. When about one third of the examination time is left, the candidate draws an unseen text of about half a standard-length page and of an appropriate degree of lexical and syntactical difficulty. After briefly reading through the text, the candidate demonstrates his/her comprehension of the English text by translating it into Danish.

11.6 *Assessment criteria (oral examination)*

11.6.1 The relative degree of difficulty of the studied text as well as of the unseen text must be taken into consideration in the assessment. The following elements are included in the assessment:

Reading aloud

- Comprehension of the text, as expressed in a coherent reading
- Intonation and diction

Presentation and conversation

- Mastery of the systems of the language
- Vocabulary and fluency
- Ability to read intensively and understanding of details
- General overview and comprehension of the text
- Ability to relate the text to other texts
- Ability to interpret and put into perspective in an independent manner
- Ability to enter into a dialogue about the text

The unseen text

- Precision
- General overview and comprehension

11.6.2 One mark is given for the entire oral examination, based on a general assessment, in which the test on the studied text has more weight than the test on the unseen text.

MATHEMATICS LINE HIGH LEVEL (A-level)

12. Aims

In addition to the requirements at obligatory level:

- The students should gain and be able to demonstrate a command of the English language with certainty and nuance, actively as well as passively.
- The students should be able to express themselves with fluency, variety and a high degree of idiomatic correctness and precision, as well as with a sense of register.
- The students' pronunciation should become clear and phonemically correct.
- The students should gain skills in analysing texts of greater difficulty in language as well as content, modern as well as older texts, and be able to employ points of view concerning principle and theory in studying texts and topics.
- The students should be able to use their skills in various extensive reading methods in work with texts of greater difficulty in language and content, including texts of a higher level of abstract thought.

13. Syllabus

13.1 The syllabus comprises 400-600 pages of standard length. None of the texts may be adapted or simplified. The texts must make greater demands on the students than those read at obligatory level. Primarily British and American texts are to be read. Various periods must be represented. Quality texts are to be chosen; however, to a limited extent texts may be chosen without a view to their intrinsic value, provided that they contribute to the understanding of aspects of the English-speaking world. Texts that deal with topics of global interest and/or use English as an international means of communication may likewise be included to a limited extent.

13.2 The syllabus must include the following genres: novel, short story, poetry and drama, as well as various types of non-fiction texts, including newspaper articles and essays. A number of texts written before 1900 must be included. If a drama by Shakespeare is read, it will fulfil the requirements for reading a drama and texts written before 1900. A number of non-fiction texts must be read, including theoretical texts that shed light on one or more significant scientific or technological problems seen from a social and cultural point of view; literary texts may also be included.

Other media than the printed word, e.g. film and videos, must be included in the syllabus. Each lesson spent working with such media is converted to six pages of standard length for the syllabus. Texts read at obligatory level may not be included.

13.3 The syllabus must be structured in such a way as to include one topic for special study; a topic will ordinarily comprise from three to six texts. The topic must include at least one text of a theoretical nature. A topic may be, for example, a genre or a mass medium; an important cultural, social or political issue; a period; a literary theme, or the study of an author. A drama by Shakespeare may be included in place of the topic for special study.

13.4 The written work is to be carried out partly as homework and partly as a classroom activity. Eight free assignments of four to five pages and fifteen assignments of 1-1½ pages are to be submitted for marking and comments.

14 Examination

14.1 There is an oral and a written examination.

14.2 *The oral examination*

14.2.1 The oral examination consists of a test on one of the texts read and on an unseen text. 30 minutes' preparation time is allotted, including the time spent giving instructions and handing out materials. Including the time spent deciding on a final mark, 2 candidates are to be examined per hour.

14.2.2 Students who have attended classes at obligatory level in the mathematics line, or have taken the examination at that level, present for their examination a selection of 125 standard-length pages from texts read intensively. The following must be represented: the four literary genres, the topic for special study (including the theoretical text or texts) and the texts written before 1900. From the texts dealing with scientific or technological problems, between fifteen and twenty standard-length pages are selected. No single text may be represented by more than twenty standard-length pages. Only texts in printed form may be used for the examination. Texts that have been studied may be accompanied by pictorial material that has been covered in the instruction.

14.2.3 Students who have not attended classes or taken the examination at obligatory level in the mathematics line present an intensively read selection of 125 standard-length pages for their examination, drawn from a studied syllabus of 900 standard-length pages. This is the sum of the syllabuses studied for obligatory and high levels in the mathematics line.

14.2.4 Self-taught candidates present an examination syllabus of about 500 pages.

14.2.5 The examiner provides unmarked copies of the texts included in the examination syllabus for the preparation room. For the examination itself, an unmarked copy of the text is used.

14.2.6 For the part of the examination covering a text that has been studied, the starting-point is an extract of an intensively read text of about three fourths of a standard-length page. The extract is read aloud, and during a conversation in English the skills gained through intensive reading are tested. The candidate is tested on skills in spoken English and in the precise understanding of

- Important details of content and language in the extract selected and its function in context,
- The text as a whole and the context to which it belongs.

The candidate begins the conversation with an independent presentation of the text. When about one third of the examination time is left, the candidate draws an unseen text of about three fourths of a standard-length page and of an appropriate degree of lexical and syntactical difficulty. After briefly reading through the text, the candidate demonstrates his/her comprehension of the English text by translating it into Danish.

14.3 *Assessment criteria (oral examination)*

14.3.1 The relative degree of difficulty of the studied text as well as of the unseen text must be taken into consideration in the assessment. The following elements are included in the assessment:

Reading aloud

- Comprehension of the text, as expressed in a coherent reading
- Intonation and diction

Presentation and conversation

- Mastery of the systems of the language
- Vocabulary and fluency
- Ability to read intensively and understanding of details
- General overview and comprehension of the text
- Ability to relate the text to other texts
- Ability to interpret and put into perspective in an independent manner
- Ability to enter into a dialogue about the text

The unseen text

- Precision
- General overview and comprehension

14.3.2 One mark is given for the entire oral examination, based on a general assessment, in which the test on the studied text has more weight than the test on the unseen text.

14.4 *The written examination*

Five hours are allotted for the written examination. The candidates are presented with a text or texts in English, possibly accompanied by pictorial material, and a number of questions are set that test the candidates' command of the English language, their comprehension of the text material, and their general skills in written English. In addition, a Danish text must be translated.

14.5 *Assessment criteria (written examination)*

14.5.1 In the assessment emphasis is primarily given to the candidates' command of the English language, i.e. the orthography, morphology, syntax, semantics and idioms of the English language. What is assessed is the extent to which the systems of the language have been mastered.

14.5.2 The summary question is furthermore assessed as to how precisely and with what degree of nuance the candidate is able to reproduce the main points of the set text in a piece of concise, coherent writing.

14.5.3 The essay question is furthermore assessed as to the degree of clarity and nuance the candidate is able to demonstrate in characterising and evaluating the text in a longer, well-organised piece of writing. The essay covers the following aspects: the plot of the text, characterisation, argumentation, the author's attitude and the message of the text.

14.5.4 The other, freer questions are furthermore assessed as to the degree of clarity and nuance the candidate is able to demonstrate in giving an account of and evaluating a problem in the text in a brief, coherent piece of writing, which in some cases may be in fictional form.

14.5.5 In addition the translation question is assessed as to how precisely the candidate is able to render the Danish text with its nuances into English. The impression given by the answer to this question is included in the assessment as a supplement to the impression given by the candidate's command of the language in the answer as a whole.

14.5.6 Finally the assessment also includes an evaluation of the structure and content of the answer as a whole.

14.6 One mark is given, based on a general assessment.

15. The major written assignment

15.1 The assignment must deal with a topic that lies within the sphere of anglophone culture, as it is reflected in texts. In working on the assignment, the methods of the subject are to be used.

15.2 The primary literature and parts of the secondary literature are to be read in English, so that nearly all quotations in the assignment are in English.

15.3 One mark is given, based on a general assessment, which is based on the general assessment criteria in Supplement 35 and on the following:

- The extent to which the primary literature has been understood
- Whether quotations are relevant and insightful
- Whether quotations are an integrated part of the assignment
- The extent to which the student is able to progress from simple reproducing of material to analysing and evaluating it.

Appendix VII Easiness of items at pre-score and post-score by groups

Pre-score over all groups (n=107):

	item	non_syl	sum	percent
1.	4	syllabus	97	90.65
2.	1	non-syllabus	95	88.79
3.	21	non-syllabus	94	87.85
4.	9	non-syllabus	88	82.24
5.	3	non-syllabus	86	80.37
6.	11	syllabus	77.86	72.77
7.	10	non-syllabus	66	61.68
8.	8	syllabus	65	60.75
9.	2	non-syllabus	63	58.88
10.	12	syllabus	61	57.01
11.	19	syllabus	53.6	50.09
12.	6	non-syllabus	49	45.79
13.	17	syllabus	48	44.86
14.	13	syllabus	46	42.99
15.	23	non-syllabus	46	42.99
16.	25	non-syllabus	40	37.38
17.	18	syllabus	35.6	33.27
18.	24	syllabus	35	32.71
19.	15	syllabus	33	30.84
20.	5	syllabus	32	29.91
21.	16	syllabus	32	29.91
22.	7	syllabus	15	14.02
23.	20	syllabus	11	10.28
24.	14	syllabus	6.33	5.92
25.	22	syllabus	5.95	5.56

Post-score over all groups (n=107):

	item	non_syl	sum	percent
1.	4	syllabus	98	91.59
2.	21	non-syllabus	97	90.65
3.	19	syllabus	94.05	87.9
4.	12	syllabus	94	87.85
5.	3	non-syllabus	91	85.05
6.	18	syllabus	91	85.05
7.	1	non-syllabus	90	84.11
8.	5	syllabus	88	82.24
9.	9	non-syllabus	87	81.31
10.	11	syllabus	86.44	80.79
11.	15	syllabus	86	80.37
12.	8	syllabus	85	79.44
13.	2	non-syllabus	82	76.64
14.	17	syllabus	78.15	73.04
15.	23	non-syllabus	68	63.55
16.	6	non-syllabus	67	62.62
17.	20	syllabus	64.3	60.09
18.	16	syllabus	61	57.01
19.	10	non-syllabus	56	52.34
20.	13	syllabus	55	51.4
21.	14	syllabus	54.69	51.11
22.	25	non-syllabus	54	50.47
23.	7	syllabus	44	41.12
24.	22	syllabus	31.9	29.81
25.	24	syllabus	28	26.17

Easiness of item at pre-test by experimental group

pre-score: NEGOT BASIS group (n=15)

	item	non_syl	sum	percent
1.	1	non-syllabus	14	93.33
2.	4	syllabus	13	86.67
3.	21	non-syllabus	13	86.67
4.	3	non-syllabus	12	80.00
5.	8	syllabus	11	73.33
6.	9	non-syllabus	10	66.67
7.	11	syllabus	9.15	61.03
8.	19	syllabus	9	60.00
9.	10	non-syllabus	8	53.33
10.	23	non-syllabus	7	46.67
11.	17	syllabus	6.20	41.33
12.	6	non-syllabus	6	40.00
13.	12	syllabus	6	40.00
14.	2	non-syllabus	5	33.33
15.	13	syllabus	5	33.33
16.	16	syllabus	5	33.33
17.	24	syllabus	5	33.33
18.	25	non-syllabus	5	33.33
19.	18	syllabus	4	26.67
20.	5	syllabus	3	20.00
21.	7	syllabus	2	13.33
22.	15	syllabus	2	13.33
23.	20	syllabus	1.10	7.33
24.	22	syllabus	0.56	3.70
25.	14	syllabus	0	0.00

pre-score: NEGOT NON-VISL group (n=20)

	item	non_syl	sum	percent
1.	4	syllabus	19	95.00
2.	21	non-syllabus	18	90.00
3.	1	non-syllabus	15	75.00
4.	3	non-syllabus	15	75.00
5.	9	non-syllabus	14	70.00
6.	2	non-syllabus	12	60.00
7.	8	syllabus	12	60.00
8.	11	syllabus	11.54	57.69
9.	12	syllabus	10	50.00
10.	10	non-syllabus	9	45.00
11.	19	syllabus	8.33	41.67
12.	24	syllabus	8	40.00
13.	25	non-syllabus	8	40.00
14.	17	syllabus	7.40	37.00
15.	16	syllabus	7	35.00
16.	23	non-syllabus	7	35.00
17.	5	syllabus	6	30.00
18.	6	non-syllabus	6	30.00
19.	13	syllabus	6	30.00
20.	18	syllabus	5	25.00
21.	15	syllabus	3	15.00
22.	20	syllabus	1.50	7.50
23.	22	syllabus	0.56	2.78
24.	7	syllabus	0	0.00
25.	14	syllabus	0	0.00

pre-score: NEGOT VISL group (n=18)

	item	non_syl	sum	percent
1.	9	non-syllabus	16	88.89
2.	1	non-syllabus	15	83.33
3.	11	syllabus	14.62	81.20
4.	3	non-syllabus	14	77.78
5.	4	syllabus	14	77.78
6.	21	non-syllabus	13	72.22
7.	2	non-syllabus	12	66.67
8.	10	non-syllabus	11	61.11
9.	13	syllabus	10	55.56
10.	15	syllabus	8	44.44
11.	23	non-syllabus	8	44.44
12.	19	syllabus	7.67	42.59
13.	17	syllabus	7.40	41.11
14.	6	non-syllabus	7	38.89
15.	8	syllabus	7	38.89
16.	12	syllabus	7	38.89
17.	25	non-syllabus	6	33.33
18.	16	syllabus	5	27.78
19.	24	syllabus	5	27.78
20.	18	syllabus	4	22.22
21.	5	syllabus	3	16.67
22.	7	syllabus	2	11.11
23.	14	syllabus	2	11.11
24.	20	syllabus	1.40	7.78
25.	22	syllabus	0.22	1.23

pre-score: ENGLISH BASIS group (n=21)

	item	non_syl	sum	percent
1.	1	non-syllabus	20	95.24
2.	4	syllabus	20	95.24
3.	21	non-syllabus	20	95.24
4.	9	non-syllabus	19	90.48
5.	3	non-syllabus	17	80.95
6.	11	syllabus	16.23	77.29
7.	10	non-syllabus	15	71.43
8.	12	syllabus	15	71.43
9.	13	syllabus	13	61.90
10.	18	syllabus	12.6	60.00
11.	19	syllabus	12.33	58.73
12.	2	non-syllabus	12	57.14
13.	8	syllabus	12	57.14
14.	17	syllabus	11.4	54.29
15.	5	syllabus	11	52.38
16.	6	non-syllabus	11	52.38
17.	23	non-syllabus	11	52.38
18.	25	non-syllabus	9	42.86
19.	15	syllabus	8	38.10
20.	24	syllabus	8	38.10
21.	7	syllabus	7	33.33
22.	16	syllabus	7	33.33
23.	20	syllabus	3.9	18.57
24.	14	syllabus	3.33	15.87
25.	22	syllabus	2.56	12.17

pre-score: ENGLISH NON-VISL group (n=15)

	item	non_syl	sum	percent
1.	21	non-syllabus	14	93.33
2.	1	non-syllabus	13	86.67
3.	3	non-syllabus	13	86.67
4.	4	syllabus	13	86.67
5.	11	syllabus	12.23	81.54
6.	6	non-syllabus	11	73.33
7.	9	non-syllabus	11	73.33
8.	8	syllabus	10	66.67
9.	10	non-syllabus	10	66.67
10.	12	syllabus	9	60.00
11.	2	non-syllabus	8	53.33
12.	23	non-syllabus	8	53.33
13.	19	syllabus	6	40.00
14.	24	syllabus	6	40.00
15.	25	non-syllabus	6	40.00
16.	17	syllabus	5.80	38.67
17.	13	syllabus	5	33.33
18.	18	syllabus	5	33.33
19.	5	syllabus	4	26.67
20.	15	syllabus	4	26.67
21.	16	syllabus	2	13.33
22.	7	syllabus	1	6.67
23.	14	syllabus	1	6.67
24.	22	syllabus	0.78	5.19
25.	20	syllabus	0.40	2.67

pre-score: ENGLISH VISL group (n=18)

	item	non_syl	sum	percent
1.	1	non-syllabus	18	100
2.	4	syllabus	18	100
3.	9	non-syllabus	18	100
4.	21	non-syllabus	16	88.89
5.	3	non-syllabus	15	83.33
6.	11	syllabus	14.15	78.63
7.	2	non-syllabus	14	77.78
8.	12	syllabus	14	77.78
9.	8	syllabus	13	72.22
10.	10	non-syllabus	13	72.22
11.	19	syllabus	10.33	57.41
12.	17	syllabus	9.8	54.44
13.	6	non-syllabus	8	44.44
14.	15	syllabus	8	44.44
15.	13	syllabus	7	38.89
16.	16	syllabus	6	33.33
17.	25	non-syllabus	6	33.33
18.	5	syllabus	5	27.78
19.	18	syllabus	5	27.78
20.	23	non-syllabus	5	27.78
21.	7	syllabus	3	16.67
22.	24	syllabus	3	16.67
23.	20	syllabus	2.7	15.00
24.	22	syllabus	1.22	6.79
25.	14	syllabus	0.33	1.85

Easiness of item at post-test by experimental group

post-score: NEGOT BASIS group (n=15)

	item	non_syl	sum	percent
1.	19	syllabus	14.01	93.4
2.	3	non-syllabus	14	93.33
3.	5	syllabus	14	93.33
4.	21	non-syllabus	13	86.67
5.	9	non-syllabus	13	86.67
6.	1	non-syllabus	12	80
7.	2	non-syllabus	11	73.33
8.	8	syllabus	10	66.67
9.	4	syllabus	10	66.67
10.	12	syllabus	10	66.67
11.	11	syllabus	9.940001	66.27
12.	17	syllabus	9.15	61
13.	15	syllabus	9	60
14.	25	non-syllabus	8	53.33
15.	18	syllabus	8	53.33
16.	16	syllabus	7	46.67
17.	6	non-syllabus	7	46.67
18.	10	non-syllabus	7	46.67
19.	23	non-syllabus	7	46.67
20.	13	syllabus	6	40
21.	20	syllabus	5.9	39.33
22.	24	syllabus	3	20
23.	22	syllabus	2.39	15.93
24.	7	syllabus	2	13.33
25.	14	syllabus	1.67	11.13

post-score: NEGOT NON-VISL group (n=20)

	item	non_syl	sum	percent
1.	21	non-syllabus	19	95
2.	4	syllabus	19	95
3.	18	syllabus	18	90
4.	12	syllabus	18	90
5.	19	syllabus	17.01	85.05
6.	8	syllabus	17	85
7.	2	non-syllabus	17	85
8.	11	syllabus	15.78	78.9
9.	9	non-syllabus	15	75
10.	1	non-syllabus	15	75
11.	5	syllabus	15	75
12.	20	syllabus	14.2	71
13.	3	non-syllabus	14	70
14.	15	syllabus	14	70
15.	14	syllabus	13.02	65.1
16.	6	non-syllabus	13	65
17.	16	syllabus	13	65
18.	17	syllabus	12.71	63.55
19.	23	non-syllabus	11	55
20.	13	syllabus	9	45
21.	25	non-syllabus	9	45
22.	7	syllabus	9	45
23.	10	non-syllabus	7	35
24.	22	syllabus	6.68	33.4
25.	24	syllabus	2	10

post-score: NEGOT VISL group (n=18)

	item	non_syl	sum	percent
1.	18	syllabus	18	100
2.	4	syllabus	18	100
3.	19	syllabus	17.33	96.28
4.	15	syllabus	17	94.44
5.	21	non-syllabus	17	94.44
6.	12	syllabus	17	94.44
7.	1	non-syllabus	16	88.89
8.	5	syllabus	16	88.89
9.	11	syllabus	15.23	84.61
10.	3	non-syllabus	15	83.33
11.	9	non-syllabus	14	77.78
12.	16	syllabus	13	72.22
13.	8	syllabus	13	72.22
14.	2	non-syllabus	13	72.22
15.	14	syllabus	12.66	70.33
16.	20	syllabus	12.5	69.44
17.	17	syllabus	12.13	67.39
18.	25	non-syllabus	11	61.11
19.	6	non-syllabus	11	61.11
20.	10	non-syllabus	11	61.11
21.	13	syllabus	10	55.56
22.	23	non-syllabus	10	55.56
23.	7	syllabus	7	38.89
24.	22	syllabus	6.92	38.44
25.	24	syllabus	5	27.78

post-score: ENGLISH BASIS group (n=21)

	item	non_syl	sum	percent
1.	9	non-syllabus	21	100
2.	4	syllabus	21	100
3.	3	non-syllabus	21	100
4.	15	syllabus	20	95.24
5.	21	non-syllabus	19	90.48
6.	18	syllabus	19	90.48
7.	11	syllabus	17.4	82.86
8.	17	syllabus	17.15	81.67
9.	19	syllabus	17.01	81
10.	12	syllabus	17	80.95
11.	5	syllabus	16	76.19
12.	2	non-syllabus	16	76.19
13.	1	non-syllabus	16	76.19
14.	23	non-syllabus	15	71.43
15.	8	syllabus	15	71.43
16.	6	non-syllabus	13	61.9
17.	16	syllabus	13	61.9
18.	10	non-syllabus	13	61.9
19.	14	syllabus	12.68	60.38
20.	25	non-syllabus	12	57.14
21.	20	syllabus	11.6	55.24
22.	7	syllabus	11	52.38
23.	24	syllabus	11	52.38
24.	13	syllabus	7	33.33
25.	22	syllabus	6.3	30

post-score: ENGLISH NON-VISL group (n=15)

	item	non_syl	sum	percent
1.	23	non-syllabus	14	93.33
2.	4	syllabus	14	93.33
3.	12	syllabus	14	93.33
4.	1	non-syllabus	14	93.33
5.	18	syllabus	13	86.67
6.	8	syllabus	13	86.67
7.	11	syllabus	12.23	81.53
8.	17	syllabus	12.02	80.13
9.	21	non-syllabus	12	80
10.	3	non-syllabus	12	80
11.	19	syllabus	11.68	77.87
12.	6	non-syllabus	11	73.33
13.	2	non-syllabus	11	73.33
14.	5	syllabus	11	73.33
15.	15	syllabus	11	73.33
16.	13	syllabus	10	66.67
17.	10	non-syllabus	9	60
18.	9	non-syllabus	9	60
19.	20	syllabus	8.900001	59.33
20.	7	syllabus	7	46.67
21.	14	syllabus	7	46.67
22.	25	non-syllabus	6	40
23.	16	syllabus	5	33.33
24.	22	syllabus	4.03	26.87
25.	24	syllabus	4	26.67

post-score: ENGLISH VISL group (n=18)

	item	non_syl	sum	percent
1.	12	syllabus	18	100
2.	19	syllabus	17.01	94.5
3.	8	syllabus	17	94.44
4.	21	non-syllabus	17	94.44
5.	1	non-syllabus	17	94.44
6.	4	syllabus	16	88.89
7.	5	syllabus	16	88.89
8.	11	syllabus	15.86	88.11
9.	18	syllabus	15	83.33
10.	9	non-syllabus	15	83.33
11.	15	syllabus	15	83.33
12.	3	non-syllabus	15	83.33
13.	17	syllabus	14.99	83.28
14.	2	non-syllabus	14	77.78
15.	13	syllabus	13	72.22
16.	6	non-syllabus	12	66.67
17.	20	syllabus	11.2	62.22
18.	23	non-syllabus	11	61.11
19.	16	syllabus	10	55.56
20.	10	non-syllabus	9	50
21.	7	syllabus	8	44.44
22.	25	non-syllabus	8	44.44
23.	14	syllabus	7.66	42.56
24.	22	syllabus	5.58	31
25.	24	syllabus	3	16.67

Appendix VIII Ranking: pre-, post-, gain scores, syllabus and non-syllabus

Pre-score syllabus:

	item	syllabus	sum	percent
1.	4	syllabus	97	90.65
2.	11	syllabus	77.86	72.77
3.	8	syllabus	65	60.75
4.	12	syllabus	61	57.01
5.	19	syllabus	53.6	50.09
6.	17	syllabus	48	44.86
7.	13	syllabus	46	42.99
8.	18	syllabus	35.6	33.27
9.	24	syllabus	35	32.71
10.	15	syllabus	33	30.84
11.	5	syllabus	32	29.91
12.	16	syllabus	32	29.91
13.	7	syllabus	15	14.02
14.	20	syllabus	11	10.28
15.	14	syllabus	6.33	5.92
16.	22	syllabus	5.95	5.56

Pre-score non-syllabus:

	item	non-syllabus	sum	percent
1.	1	non-syllabus	95	88.79
2.	21	non-syllabus	94	87.85
3.	9	non-syllabus	88	82.24
4.	3	non-syllabus	86	80.37
5.	10	non-syllabus	66	61.68
6.	2	non-syllabus	63	58.88
7.	6	non-syllabus	49	45.79
8.	23	non-syllabus	46	42.99
9.	25	non-syllabus	40	37.38

Post-score syllabus :

	item	syllabus	sum	percent
1.	4	syllabus	98	91.59
2.	19	syllabus	94.05	87.90
4.	12	syllabus	94	87.85
3.	18	syllabus	91	85.05
5.	5	syllabus	88	82.24
6.	11	syllabus	86.44	80.79
7.	15	syllabus	86	80.37
8.	8	syllabus	85	79.44
9.	17	syllabus	78.15	73.04
10.	20	syllabus	64.3	60.09
11.	16	syllabus	61	57.01
12.	13	syllabus	55	51.40
13.	14	syllabus	54.69	51.11
14.	7	syllabus	44	41.12
15.	22	syllabus	31.9	29.81
16.	24	syllabus	28	26.17

Post-score non-syllabus :

	item	non-syllabus	sum	percent
1.	21	non-syllabus	97	90.65
2.	3	non-syllabus	91	85.05
3.	1	non-syllabus	90	84.11
4.	9	non-syllabus	87	81.31
5.	2	non-syllabus	82	76.64
6.	23	non-syllabus	68	63.55
7.	6	non-syllabus	67	62.62
8.	10	non-syllabus	56	52.34
9.	25	non-syllabus	54	50.47

Gain score syllabus :

	item	non_syl	points	percent
1.	5	syllabus	56	52.34
2.	18	syllabus	55.4	51.78
3.	20	syllabus	53.3	49.81
4.	15	syllabus	53	49.53
5.	14	syllabus	48.36	45.20
6.	19	syllabus	40.45	37.80
7.	12	syllabus	33	30.84
8.	17	syllabus	30.15	28.18
9.	16	syllabus	29	27.10
10.	7	syllabus	29	27.10
11.	22	syllabus	25.95	24.25
12.	8	syllabus	20	18.69
13.	13	syllabus	9	8.41
14.	11	syllabus	8.58	8.02
15.	4	syllabus	1	0.93
16.	24	syllabus	-7	-6.54

Gain score non-syllabus :

	item	non_syl	points	percent
1.	23	non-syllabus	22	20.56
2.	2	non-syllabus	19	17.76
3.	6	non-syllabus	18	16.82
4.	25	non-syllabus	14	13.08
5.	3	non-syllabus	5	4.67
6.	21	non-syllabus	3	2.80
7.	9	non-syllabus	-1	-0.93
8.	1	non-syllabus	-5	-4.67
9.	10	non-syllabus	-10	-9.35

APPENDIX IX Gains(%) by experimental group

VISL GAIN

	VISL ALL			VISL EBGLISH			VISL NEGOT		
	item	nm_syl	percent	item	nm_syl	percent	item	nm_syl	percent
1.	5	syllabus	66.67	5	syllabus	61.11	18	syllabus	77.78
2.	18	syllabus	66.67	18	syllabus	55.56	5	syllabus	72.22
3.	20	syllabus	54.44	20	syllabus	47.22	20	syllabus	61.67
4.	14	syllabus	50.00	14	syllabus	40.74	14	syllabus	59.26
5.	19	syllabus	45.37	15	syllabus	38.89	12	syllabus	55.56
6.	15	syllabus	44.44	19	syllabus	37.04	19	syllabus	53.70
7.	12	syllabus	38.89	23	nm-syllabus	33.33	15	syllabus	50.00
8.	16	syllabus	33.33	13	syllabus	33.33	16	syllabus	44.44
9.	22	syllabus	30.52	17	syllabus	28.89	22	syllabus	36.96
10.	17	syllabus	27.78	7	syllabus	27.78	8	syllabus	33.33
11.	7	syllabus	27.78	22	syllabus	24.07	25	nm-syllabus	27.78
12.	8	syllabus	27.62	6	nm-syllabus	22.22	7	syllabus	27.78
13.	6	nm-syllabus	22.22	12	syllabus	22.22	17	syllabus	26.35
14.	23	nm-syllabus	22.22	8	syllabus	22.22	4	syllabus	22.22
15.	25	nm-syllabus	19.44	16	syllabus	22.22	6	nm-syllabus	22.22
16.	13	syllabus	16.67	25	nm-syllabus	11.11	21	nm-syllabus	22.22
17.	21	nm-syllabus	13.89	11	syllabus	9.40	23	nm-syllabus	11.11
18.	11	syllabus	6.41	21	nm-syllabus	5.56	1	nm-syllabus	5.56
19.	4	syllabus	5.56	24	syllabus	0	2	nm-syllabus	5.56
20.	2	nm-syllabus	2.78	3	nm-syllabus	0	3	nm-syllabus	5.56
21.	3	nm-syllabus	2.78	2	nm-syllabus	0	11	syllabus	3.42
22.	1	nm-syllabus	0	1	nm-syllabus	-5.56	24	syllabus	0
23.	24	syllabus	0	4	syllabus	-11.11	13	syllabus	0
24.	10	nm-syllabus	-11.11	9	nm-syllabus	-16.67	10	nm-syllabus	0
25.	9	nm-syllabus	-13.89	10	nm-syllabus	-22.22	9	nm-syllabus	-11.11

NONVISL GAIN

	NON-VISL ALL			NON-VISL EBGLISH			NON-VISL NEGOT		
	item	nm_syl	percent	item	nm_syl	percent	item	nm_syl	percent
1.	20	syllabus	60.08	20	syllabus	56.67	18	syllabus	65.00
2.	18	syllabus	59.17	18	syllabus	53.33	14	syllabus	65.00
3.	14	syllabus	52.50	15	syllabus	46.67	20	syllabus	63.50
4.	15	syllabus	50.83	5	syllabus	46.67	15	syllabus	55.00
5.	5	syllabus	45.83	17	syllabus	41.33	7	syllabus	45.00
6.	7	syllabus	42.50	14	syllabus	40.00	5	syllabus	45.00
7.	19	syllabus	40.56	7	syllabus	40.00	19	syllabus	43.33
8.	12	syllabus	36.67	23	nm-syllabus	40.00	12	syllabus	40.00
9.	17	syllabus	33.95	19	syllabus	37.78	6	nm-syllabus	35.00
10.	23	nm-syllabus	30.00	13	syllabus	33.33	16	syllabus	30.35
11.	22	syllabus	25.91	12	syllabus	33.33	22	syllabus	30.00
12.	16	syllabus	25.00	22	syllabus	21.48	17	syllabus	26.57
13.	13	syllabus	24.17	16	syllabus	20.00	8	syllabus	25.00
14.	2	nm-syllabus	22.50	8	syllabus	20.00	2	nm-syllabus	25.00
15.	8	syllabus	22.50	2	nm-syllabus	20.00	11	syllabus	21.15
16.	6	nm-syllabus	17.50	1	nm-syllabus	6.67	23	nm-syllabus	20.00
17.	11	syllabus	10.58	4	syllabus	6.67	13	syllabus	15.00
18.	1	nm-syllabus	3.33	11	syllabus	0.00	21	nm-syllabus	5.00
19.	25	nm-syllabus	3.33	6	nm-syllabus	0.00	25	nm-syllabus	5.00
20.	4	syllabus	2.50	25	nm-syllabus	0.00	9	nm-syllabus	5.00
21.	21	nm-syllabus	-4.17	10	nm-syllabus	-6.67	1	nm-syllabus	0
22.	9	nm-syllabus	-4.17	3	nm-syllabus	-6.67	4	syllabus	0
23.	3	nm-syllabus	-5.83	24	syllabus	-13.33	3	nm-syllabus	-5.00
24.	10	nm-syllabus	-8.33	21	nm-syllabus	-13.33	10	nm-syllabus	-10.00
25.	24	syllabus	-21.67	9	nm-syllabus	-13.33	24	syllabus	-30.00

BASIS GAIN (%)

	BASIS ALL			BASIS ENGLISH			BASIS NEGOT		
	item	non_syl	percent	item	non_syl	percent	item	non_syl	percent
1.	15	syllabus	51.90	15	syllabus	57.14	5	syllabus	73.33
2.	5	syllabus	48.57	14	syllabus	44.45	15	syllabus	46.67
3.	20	syllabus	34.33	20	syllabus	36.67	2 non-syllabus		40.00
4.	2 non-syllabus		29.53	18	syllabus	30.48	19	syllabus	33.33
5.	18	syllabus	28.57	16	syllabus	28.57	20	syllabus	32.00
6.	14	syllabus	27.78	17	syllabus	27.34	12	syllabus	26.67
7.	19	syllabus	27.78	5	syllabus	23.81	18	syllabus	26.66
8.	17	syllabus	23.48	19	syllabus	22.22	25 non-syllabus		20.00
9.	16	syllabus	20.96	7	syllabus	19.05	9 non-syllabus		20.00
10.	12	syllabus	18.09	2 non-syllabus		19.05	17	syllabus	19.62
11.	25 non-syllabus		17.14	3 non-syllabus		19.05	16	syllabus	13.34
12.	3 non-syllabus		16.19	23 non-syllabus		18.05	3 non-syllabus		13.33
13.	22	syllabus	14.86	22	syllabus	17.59	22	syllabus	12.13
14.	9 non-syllabus		14.76	8	syllabus	14.29	14	syllabus	11.11
15.	7	syllabus	9.53	24	syllabus	14.28	13	syllabus	6.67
16.	23 non-syllabus		9.02	25 non-syllabus		14.28	6 non-syllabus		6.67
17.	6 non-syllabus		8.09	12	syllabus	9.52	11	syllabus	5.12
18.	11	syllabus	5.31	9 non-syllabus		9.52	7	syllabus	0
19.	8	syllabus	3.82	6 non-syllabus		9.52	21 non-syllabus		0
20.	24	syllabus	0.48	11	syllabus	5.49	23 non-syllabus		0
21.	21 non-syllabus		-2.38	4	syllabus	4.76	8	syllabus	-6.66
22.	4	syllabus	-7.62	21 non-syllabus		-4.76	10 non-syllabus		-6.66
23.	10 non-syllabus		-8.10	10 non-syllabus		-9.53	24	syllabus	-13.33
24.	13	syllabus	-10.95	1 non-syllabus		-19.05	1 non-syllabus		-13.33
25.	1 non-syllabus		-16.19	13	syllabus	-28.57	4	syllabus	-20.00

Appendix X Think-aloud excerpts with their full protocols

TA Excerpt 7.1

TAS4. POST-Sentence 2.

+-----+
+
POST-Sentence 2: *It is a very good book*
+-----+

Segment	Transcription
01	ahr...this bothers me that this machine is so..... slow
02	hm... (the blue/green boxes did not appear but half-formed words
03	which he has to click on to make the boxes appear) hm...
04	then I have the predicator
05	which is this (points at is but does not click)
06	and subject which is It (does not click)
07	and then I have what it is, and it is <i>a very good book</i>
08	so I'm going to try to put these four words into a group (clicks for the first time)
09	and I'm correct,

+-----+

Full protocol:

And again build the tree...ahr.. this bothers me that this machine is so..... slow, hm.. (the blue/green boxes did not appear but half-finished words which he has to click on to make the boxes appear). Hm...then I have the predicator which is this (points but does not click) and subject which is *It* (does not click) and then I have what it is, and it is a very good book so I'm going to try to put these four words into a group(clicks for the first time), and I'm correct, but then I'm trying to put *very good* into a group.

Hm..so...*Is* must be the predicator, *It* should be the subject noun, and it must be the complement to the subject, it tells me what the subject is so I'm ...hrm (scrolls)...there you go , yes..ah...*a very good book* I have to put this into a group, and I then can put the heads and dependent on. Hmmm..tch..hm...*very good book*..... *very good* must be a dependent because I then I can make a group out of it, yes, and *a* should also be a dependent, and *book* should be the head, yes, hm...a and d...that mmmchh....describing ..somehow guessing that *a* is an article, and I'm correct, and the *book* is a noun and *very good* is a group again, and where I have head and dependent, and then because *good* is the last word I'm guessing that it must the head, and *very* has to be the dependent, and *very* is a verb, and *good* is a...is an adjective because it says something about you something tch..and ...yeah..tch...ahm

PROMPT

Ah.....*very* ...ahr... I can't remember what *very* is now, it tells something about ...the adjective, no.....tch....

PROMPT

Well...(sighs) hrh..I just don't know...hmmm

PROMPT

(Sighs)... Well...ah... if....right now I'm just, I'm tempted to just...to guess, but..... (sighs)I think I'm just gon....

PROMPT

I'm going to reveal the node right now because ..I can't see it, and it's an adverb, hm..
tch..

PROMPT

Well, ahm... I'm actually not much surprised because...but I should be ..and I looked
byI placed it as a verb, adverbs I believe this says something about verbs, but in this
case it obviously says something about the adjective, then there must be some rules that
I have forgotten, so ..I'm going to check that out later in my grammar book 'cause right
now I'm just gonna go on with the sentence. Ahmm..and *is* should be a verb, yes, and *It*
should be... the pronoun for something else, yes, stands in stead of a noun.

TA Excerpt 7.2

TAS7. PRE-Sentence 1a.

Segment	Transcription
01	I know that <i>George</i> is
02
03	the
04 (scrolls)...
PROMPT	
05	...
06	t h e subject.
07	he's doing the paint,
08	and he is of course the noun

TA Excerpt 7.3

TAS7. PRE-Sentence 1b.

Segment	Transcription
09	I'm just trying to put <i>the house</i> as object
10	but I'm not sure if it's an indirect object
11	or what I have to do.
12	I can't
13	...
14	I know

Full protocol:

George is...I've .. can't....where do you start ...hm...*paint* is the...predicor.

Let me see...uhmm...god...why can't we....ahm...hm...let me see... ah..there it is (clicks predicator) and.. ahm..I don't exactly..it's .. it's hmmm...it's..ahm..... ahm... what is it?...let me see...ahm...it'ssss...it's a verb...and..I know that *George* is...the... (scrolls).....t h e subject.

He's doing the paint, and he is of course the noun (removes the box marking and can't get things to work) ahm...there...and we have hm..the..I mean *the house yesterday*...that's..we have to put it together because it's a group. *The house* is one group, and *yesterday* is a group, *the* it's an article

.. mmm.. an article, we have it here, and *the house* is the object the direct object (she has not grouped the two words, and it doesn't work)....No...it...was...wrong..let's say it's an object..hmm

it doesn't work I guess.' Ok.. at least I know that *house* is a noun....and... Let me see, build trees

.. hrmm...(goes to Tools and Build tree)(clicks on the boxes and art disappears from *the* and n from *house*)...what happenedgod ...they'll.. I tried...hmm.....

PROMPT

Ahm... I'm just trying *the house* is the object, and it doesn't work (laughs).....(clicks noun) Ok...it's a noun, and it's an object ...I'm not quite sure about if it's... it must be an..a direct object.... it doesn't work (laughs), let's see (scrolls).....hmm.... it doesn't work anymore

PROMPT

I'm just trying to put ... *house* is the object I mean *the house* is what you group together ..this is the objecthm...and *the* the article (accidentally removes P/v), and it's still..... ok (now has art and n) and it's adverbium (should be dependent)..no..(discovers that P/v has disappeared) I don't know what's going on. Hmmm...den bliver ved med at slette det der. Det har jeg allerede gjort. *Painted* is.....(scrolls) and.....

PROMPT

(Laughs) I'm lost. (Scrolls first Function then form)...

PROMPT

It's a verb. It's a noun word. We have it ..no..it's the predicator¹²³ ...ja,...OK. I'm not quite sure about *yesterday*...let me have the trees. Oh my god what's going on (goes to and clicks on Tools/Build tree) Oh my god..so that's.. Ok fine. ..god.. sorry about that. Hm...the subject and we know it's a noun. So we go to the next which is.... predi...preticator....and.....

PROMPT

I'm just putting it in the right place, and I know that *Painted* is the pretigor....there...like this and the *house*... it's a noun..... and *the* it's an article if I'm...and *the house* is the... object (still hasn't grouped them).... ahm... maybe because it's wrong it doesn't put it or.. indirect object..no.. Pardon... I have to...(Finally groups them). Hmm.. like this. OK. I know that it's the object.

PROMPT

hm..I'm just wondering.. ahm... no... it doesn't work correctly. It must be me..... (laughs)...uhm... suddenll... I don't know what's going on....(laughs)

PROMPT

¹²³ TAS7 has varying pronunciations of the word 'predicator'.

I'm just trying to put *the house* as object but I'm not sure if it's an indirect object or what I have to do. I can't.. I know ... this (*the*) is adverbium (should be dependent).....

PROMPT

It's an adverbium. It doesn't work so it's not....

PROMPT

Ahm..still, if it's wrong or right that *the* it's an adverbium. It's an A. It doesn't work. God. (heavy sigh)...It's not an A. I don't know what's going on....

Ok, fine. Finally Quit. (Decides to quit but has not done anything about *yesterday*)

TA Excerpt 7.4

TAS2. PRE-Sentence 1a.

PRE-Sentence 1: <i>George painted the house yesterday</i>	
Segment	Transcription
01	I'm looking for something
02	...
03	it didn't the first time,
04	but it's the predicator
05	and <i>painted</i> a verb
06	and <i>George</i> is the subject
07	and a noun
08	and

TA Excerpt 7.5

TAS2. PRE-Sentence 1b.

PRE-Sentence 1: <i>George painted the house yesterday</i>	
Segment	Transcription
09	and <i>yesterday</i>
10	...hmmm...
PROMPT	
11	ahmm..
12	I'm just thinking what it is
13	don't know
14	...
15	I'm not sure
16	...adverbial
17	...ah...Ehmm...puh...
PROMPT	
18	I'm wondering what it is yesterday as a form
19	...eh...

PROMPT

20 say something
21 yeah
22 I'm wondering what it is
23 ...
24 is it not adverb
25 yes

+-----+

Full protocol:

George..painted ...the house...yesterday.. hmm.. And Tools, Build tree. Now I begin the top..Function.. where is the.....statement..the form is a clause...so...(uf)..then *painted*, oh, I'm sorry....

PROMPT

Speak out,..yeah, I'm looking for something.....it didn't the first time, but it's the predicator and *painted* a verb and *George* is the subject and a noun and..*the house* is a group...or is it? It is..and function for the house is direct object. Yeah,(small laugh) hm.. and form is a group and then function dependent but form is an article. And *house* is the head of the group. It is... I'm sorry....and *house* is a noun. And *yesterday*..hmmm...

PROMPT

Ahmm.. I'm just thinking what it is. Don't know...I'm not sure. ...an adverbial...h..ehmm....puh...

PROMPT

I'm wondering what it is *yesterday* as a form...Eh...

PROMPT

Say something, yeah, I'm wondering what it is...is it not adverb, yes,..Ok, then I have finished the sentence.

TA Excerpt 7.6

TAS9. POST-Sentence 1a.

+-----+

POST-Sentence 1: *Jill sold the car today*

+-----+

Segment	Transcription
01	And <i>Jill</i> is subject
02	and <i>sold</i> is predicator
03	and <i>the car</i> is object
04	and <i>today</i> is adverbial
05	and then I need to make a tree of this

+-----+

Full protocol:

And *Jill* is subject, and *sold* is predicator, and *the car* is object, and *today* is adverbial..and then I need to make a tree of this. And it's a ...see...statement, clause ...and subject, that's a noun, *sold*, predicator, verb, and object (writes O- g) and *today* (writes A- adv), *the car*, *the* is dependent, and *car* is head and noun. Yeah, that's right.

TA Excerpt 7.7

TAS8. POST-Sentence VISL on paper.

-----+
POST-Sentence VISL on paper: *She visited the hospital yesterday*

Segment	Transcription
01	to start with I just find the predicator and the subject
02	and I ask what did <i>She visit</i>
03	and <i>She visit the hospital</i>
04	and an adverb (has written S P O A above the words)
05	so
06	...

Full protocol:

Ok, to start with I just find the predicator and the subject, and I ask what did *She visit*, and *She visit the hospital*, and an adverb (has written S P O A above the words). So... and we have subject (writes S under *She*) and we predicator (writes P under *visited*)... *She visited*, and we have the object, and it's a group (writes O, but not g under *the hospital*), and we have the head (*hospital*), it's a noun, and we have dependent (*the*), it's an article, and we have *yesterday*, and it's adverb (writes A:adv).Ja. Ok.

TA Excerpt 7.9

TAS11. PRE-Sentence 1a.

-----+
PRE-Sentence 1: *George painted the house yesterday*

Segment	Transcription
01	hvad skete der [what happened there]
02	det må [it must]
03	...
PROMPT	
04	ja [yes]
05	han malede [he painted] (writes o under <i>painted</i>),
06	og hvem malede? [and who painted]
07	det gjorde <i>George</i> [George did] (writes x under <i>George</i>),
08	og hvad malede han? [and what did he paint]
09	han malede huset [he painted the house] (underlines <i>the house</i> , writes Δ)
10	og så ved jeg ikke, hvad den sidste er, [and then I don't know what the last one is]
11	men det er så ...de to der [but this is then...those two] (indicates <i>the house</i>)

TA Excerpt 7.10

TAS11. PRE-Sentence 1b.

-----+
PRE-Sentence 1: *George painted the house yesterday*

Segment	Transcription
12	og så er der en [and then there is one]
13	...
14	her [here]
15	jeg ved ikke, hvad det er for en, [I don't know what it is]
16	jeg må hellere lave en firkant [I'd better make a square] (writes □ under <i>yesterday</i>).

-----+

Full protocol:

OK. *George painted the house yesterday*. Hvad skete der, det må.....

PROMPT

Ja. Han malede (writes o), og hvem malede? Det gjorde *George* (writes x), og hvad malede han? Han malede huset (underlines the house, writes Δ) og så ved jeg ikke, hvad den sidste er, men det er så... de to der (indicates the house), og så er der en... her. Jeg ved ikke, hvad det er for en, jeg må hellere lave en firkant (writes □). *Painted* det er så... hmmm... subject (writes S under *painted*), ... og.....

PROMPT

Ahm... jeg kan ikke huske... om ...*George*, om det er eh...ehm... det er navneord (writes n) og det er *the house* også (writes n under *house*). Og... der er et eller andet deroppe (points at *She*),... det er.. hvad er det det hedder?... det hedder...

PROMPT

Ja, men jeg har ikke lige noget i tankerne lige nu. ...hrmppp. Subject, object... sådan noget der (writes S under *George*) og så tror jeg at ...*the house* det er....

PROMPT

Yesterday det er i hvert fald adverbial (writes A under *yesterday*). Og så *house* det er... det er head (writes H under *house*), og det andet det er det er en SUB(writes SUB under *the*).

PROMPT

Jeg sidder og tænker på, om jeg er færdig, for det synes jeg jeg er, for det er det sidste, jeg lige har gjort mine tanker om.

TA Excerpt 7.11

TAS7. PRE-Sentence 1c.

-----+
PRE-Sentence 1: *George painted the house yesterday*

Segment	Transcription
---------	---------------

15¹²⁴ ahm...there...and we have hm..the..
16 I mean *the house yesterday*...that's..
17 we have to put it together because it's a group *the house* is
one group
18 and *yesterday* is a group,
19 *the* it's an article
20 .. mmm.. an article we have it here
21 and *the house* is the object the direct object (she has not
grouped the two words, and it doesn't work)
22
23 no
24 ...
25 it
26 ...
27 was
28 ...
29 wrong
30 ...
31 let's say it's an object
32 ...hmm
33 it doesn't work I guess
34 ok
35 at least I know that *house* is a noun
36 ...
37 and
38 ...
39 let me see,
40 build trees
41 ... hrmm...(goes to Tools and Build tree)
42 ...(clicks on the boxes and art disappears from *the* and n from
house)
43 ...
44 what happened
45
46 god
47 ...
48 they'll
49 ...
50 I tried
51 ...
52 hmm
53 ...
PROMPT
54 ahm...
55 I'm just trying
56 *the house* is the object,

¹²⁴ See full protocol under TA Excerpt 7.3 above.

57 and it doesn't work
58 (laughs)
59 ... (clicks on noun)
60 ok (the software accepts this input)
61 ...
62 it's a noun,
63 and it's an object
64 ...
65 I'm not quite sure about if it's
66 ...
67 it must be an
68 ...
69 a direct object
70 ...
71 it doesn't work
72 (laughs)
73 let's see (scrolls)
74 ...
75 hmm...
76 it doesn't work anymore
77 ...
PROMPT
78 I'm just trying to put
79 ...
80 *house* is the object
81 I mean *the house* is what you group together
82 ...
83 this is the object
84 ...
85 hm...
86 and *the* the article (accidentally removes P/v),
87 and it's still...
88 ok (now has art and n in place).
89 and it's adverbium (should be dependent)..
90 no...(discovers that P/v has disappeared)
91 I don't know what's going on.
92 hm...
93 den bliver ved med at slette det der [it keeps erasing that]
94 det har jeg allerede gjort [I already did that]
95 *painted* is ...(scrolls)
96 and ...
PROMPT
97 (Laughs)
98 I'm lost. (Scrolls first Function then form)
PROMPT
99 It's a verb.
100 It's a noun word.
101 We have it

102 ...no...
103 it's the predictor
104 ja [yes]
105 ... ok
106 I'm not quite sure about *yesterday*
107 ...
108 let me have the trees
109 oh my god what's going on (goes to Tools/Build tree and
starts over from the top)
110 oh my god...
111 ok fine
112 ... god
113 ...
114 sorry about that
115 hm...
116 the subject
117 and we know it's a noun.
118 so we go to the next which is
119 ...
120 predi...¹²⁵
121 preticator
122 ...
123 and
124 ...
PROMPT
125 I'm just putting it in the right place,
126 and I know that *painted* is the pretigor
127 ...
128 there
129 ...
130 like this
131 and the *house*
132 ...
133 it's a noun
134 ...
135 and *the* it's an article if I'm
136 ...
137 and *the house* is the
138 ...
139 object (still hasn't grouped them)
140 ...
141 ahm...
142 maybe because it's wrong it doesn't put it
143 or
144 ...
145 indirect object

¹²⁵ TAS7 keeps having difficulties in pronouncing 'predicator' correctly.

146	no
147	...
148	pardon
149	...
150	I have to
151	... (finally groups them)
152	hm
153	like this
154	ok
155	I know that it's the object
156	...
PROMPT	
157	hm...
158	I'm just wondering
159	ahm...
160	no
161	...
162	it doesn't work correctly
163	it must be me
164	... (laughs)
165	...uhm...
166	suddenly
167	...
168	I don't know what's going on
169	.. (laughs)
PROMPT	
170	I'm just trying to put <i>the house</i> as object
171	but I'm not sure if it's an indirect object
172	or what I have to do
173	I can't
174	...
175	I know
176	...
177	this (points the cursor at <i>the</i>) is adverbium (should be dependent)
178	...
PROMPT	
179	it's an adverbium
180	it doesn't work so it's not.

+-----+

TA Excerpt 7.12

TAS11. PRE-Sentence 1c.

+-----+
PRE-Sentence 1: *George painted the house yesterday*
+-----+

Segment

Transcription

17 *painted*
18 det er så [that is]
19 ... hmmm...
20 subject (writes S under *painted*),
21 ...
22 og [and]
23 ...
PROMPT
24 ahm...
25 jeg kan ikke huske [I can't remember]
26 ...
27 om [if]
28 ...
29 *George*, om det er [*George* if that is]
30 eh...
31 ehm...
32 det er navneord [it's a noun] (writes n)
33 og det er *the house* også [and *the house* also is] (writes n
under *house*)
34 og [and]
35
36 der er et eller andet deroppe (points at *She* which is in another
sentence), [there is something up here]
37 ...
38 det er [it is]
39 ...
40 hvad er det det hedder? [what is it called]
41 ...
42 det hedder [it is called]
43 ...
PROMPT
44 ja [yes]
45 men jeg har ikke lige noget i tankerne lige nu. [but I have
nothing in my mind right now]
46 ...hrmppp.
47 subject, object
48 ...
49 sådan noget der (writes S under *George*) [something like that]
50 og så tror jeg at [and then I think that]
51 ...
52 *the house*
53 det er [it is]
54
PROMPT
55 *yesterday* det er i hvert fald adverbial [*yesterday* that is
definitely adverbial] (writes A under *yesterday*)
56 og så *house* det er [and then *house* that is]

57 ...
 58 det er head [it is head] (writes H under *house*),
 59 og det andet det er det er en SUB [and the other one that is
 a SUB] (writes SUB under *the*).

TA Excerpt 7.13

TAS3. POST-Sentence on paper.

VISL POST-Sentence on paper: *She visited the hospital yesterday*

Segment	Transcription
[01-15]	
16	og så er det er noget med tidsangivelse [and then there is something involving a definition of time]
17	og det en adverbial [and that's an adverbial]
18	nej [no]
19	adverb og adverbial [adverb and adverbial]
20	eller omvendt var det [or the other way round it was]
21	eh...
22	og det er <i>yesterday</i> [and that is <i>yesterday</i>]
23	sådan [that's it]

Full protocol:

Så starter jeg med at sige, det er et statement og clause. Og *She* det er subject og pronoun, og hvad gjorde hun? Næh,.. *She visited*..det er predicator, verb... og hun besøgte *the hospital*, så det object, ehm.. og det er en group, fordi det er to ord. Og *the hospital*... *hospital* er head, *the* er dependent eh..*the* er article, *hospital* er noun, og så er det er noget med tidsangivelse, og det en adverbial, nej, adverb og adverbial, eller omvendt var det. Eh...Og det er *yesterday*. Sådan.

TA Excerpt 7.14

TAS3. PRE-Sentence 2.

PRE-Sentence 2: *She is a very beautiful woman*

Segment	Transcription
[01-76]	
76	<i>very</i> det må være [<i>very</i> that must be]
77	...
78	jeg prøver og se om det er adverbial [I'll try if it's an adverbial]
79	næh [no]

80 ...
 81 *very*, jeg skal lige prøve og se hvad der er [very, I'll just try
 what there is]
 82 ...
 83 nåh [oh]
 84 det er selvfølgelig dependent igen, fordi det er en gruppe
 [it's of course a dependent again because it is a group]
 85 og så er det [and then it is]
 86 ...
 87 herovre adverb [over here adverb]
 88 ja sådan [yes, that's it]

+-----+

Full protocol:

She is a very beautiful woman (typing). Og så får jeg igen den til at analyse, og igen går jeg igen op i Tools, og får den til at bygge det. Ja, og så kan jeg også se igen, der er sådan nogle grupper der. Jeg tror, det her en en stor gruppe, lad os se hvad den siger, ja, og så er der en til gruppe. Og så starter jeg igen øverst, og det er en eh.. statement, og det er igen en clause, og så starter jeg igen med at finde subject, og det er *she*, og det er et pronoun, og predicator er *is*, se om jeg kan finde det her..... der. Emmm.....og det er en verb,.. er det ikke en verb?Jow, der trykkede jeg lige forkert. Eh,..og så har jeg den her group, som jeg tror er noget.. adverbial. Næh,hmmm.. hun er...nå, så hopper jeg altså lige videre til den her jeg ved hvad.. nåh, det fordi det er helt deroppe, så er hun selvfølgelig et complement. Jeg skal nok ha' det lidt ..subject complement, complement to the subject. Og det er en group. Ehm.. jeg ska' ha' den til at slope, .. vise de der. Ja, .. jeg synes, det er lidt forvirrende. Og så skal jeg have *a* som er article, men det er også dependent, ja. Hvorfor viser den ikke den der igen?.... og *very beautiful*... det er.....ehm....også... dependent, og så er det en group, og det er igen nedenunder. Det kunne være jeg lige skulle tage den først... ikke den der*very* det må være... jeg prøver og se om det er adverbial næh,*very*, jeg skal lige prøve og se hvad der er... nåh, det er selvfølgelig dependent igen, fordi det er en gruppe, og så er det... herovre adverb, ja sådan...og *beautiful* det må så være head i gruppen. Det er den Joh, nu kom den, og det er et adjective, og *woman*.. den skal lige vise den der ..*woman* det må være head i gruppen. Vil den ikke være det? Head..joh, så kom den og *woman* det er et noun. Og så er jeg også færdig med den sætning.

TA Excerpt 7.15

TAS5. PRE-Sentence 1.

+-----+
 PRE-Sentence 1: *George painted the house yesterday*
 +-----+

Segment	Transcription
[01-60]	
61	Then I want to find out what <i>yesterday</i> is
62	I think
63

it's ahm ..it's annn article... ah... Then I want to find out what *yesterday* is. I think.... it's an adverbial. Yes, and ah.... ah ..no.. (laugh)(picks article) ah..... hmmm...

PROMPT

Ah.. I'm not sure, I thought it was an adverbial, maybe I think it's a pronoun. I just gonna try some (picks pronoun)... No, it isn't (laugh). Ahmmmm...

PROMPT

Ya.. ahm... (sigh), it tells me something about time... (sigh.) Ah... It's ...conjunction perhaps.. no. Ah.Uhm. I don't know. (Picks adverb). Oh. Yes. I didn't know what it was so I just ... well, tried some things.

TA Excerpt 7.16

TAS3. POST-Sentence 1.

-----+
 POST-Sentence 1: *Jill sold the car yesterday*

Segment	Transcription
01	<i>Jill</i> , det er subject [<i>Jill</i> is subject]
02	og det er et noun
03	... (n appears a bit delayed, and she doesn't see it)
04	måske et pronoun (sees the n)
05	...
06	og <i>sold</i> det er predicator,
07	og det er verb
08	og hun solgte bilen [and she sold the car]
09	og det er object [and that's object]
10	direct object
11	og det er en group [and it's a group]
12	...
13	og <i>the</i> det er dependent til <i>car</i>
14	og det er article
15	og så er <i>car</i> head i den gruppe [and then car is head in that group]
16	og a noun [and a noun]
17	nej [no]
18	...
PROMPT	
19	[yes]
20	jeg ved bare ikke hvad det skulle være, hvis det ikke er [I don't know what it is supposed to be if it is not]
21	...
22	jeg synes altså at en bil det er et navneord [I do think that car is a noun]
23	...
24	men jeg kan da godt prøve med noget andet [but I can try something else]
25	...(tries noun again)

26	nå det vil den gerne have det [well, that it is willing to accept]
27	ok
28	man skal åbenbart ikke give sig [one should stick to one's guns]
29	og så <i>today</i> [and then <i>today</i>]
30	det har jeg [I have that]
31	...
32	vil jeg sige er adverbial [I put down as adverbial]
33	den skal lige have det et par gange [it just needs to be told more than once]
34	og [and]
35	eh...
36	det er også noun [also it is a noun]
37	nej [no]
38	...
39	det er måske adverb [it could be and adverb]
40	ja [yes]
41	og så er jeg færdig med den

+-----+

Full protocol:

Jill, det er subject, og det er et noun.....(n appears a bit delayed, and she doesn't see it), måske et pronoun (sees the n),.. og *sold* det er predicator, og det er verb, og hun solgte bilen, og det er object, direct object, og det er en group,..... og *the* det er dependent til *car*, og det er article, og så er *car* head i den gruppe, og a noun,nej.....
PROMPT

Ja, jeg ved bare ikke, hvad det skulle være, hvis det ikke er..... Jeg synes altså at en bil, det er et navneord,..... men jeg kan da godt prøve med noget andet,(tries noun again) nå, det vil den gerne have det. OK. Man skal åbenbart ikke give sig. Og så *today*, det har jeg ...vil jeg sige er adverbial. Den skal lige have det et par gange. Og eh...det er også noun, nej.. det er måske adverb. Ja. Og så er jeg færdig med den.

TA Excerpt 7.17

TAS9. POST-Sentence 1b.

+-----+

POST-Sentence 1: *Jill sold the car today*

+-----+

Segment	Transcription
[01-11] ¹²⁶	
12	and subject
13	that's a noun
14	<i>sold</i> predicator verb
15	and object (writes O- g)

¹²⁶ See full protocol under TA Excerpt 7.6 above.

16 and *today* (writes A- adv)
 17 *the car* *the* is dependent
 18 and *car* is head and noun
 19 yeah
 20 that's right

TA Excerpt 7.18

TAS10. POST-Sentence 2.

POST-Sentence 2: *It is a very good book*

Segment	Transcription
01	<i>is</i> det er predicator, verbum [<i>is</i> that is predicator, verb]
02	<i>It</i> er subjektet, som er pronoun [<i>It</i> is subject, which is pronoun]
03	fordi det står i stedet for noget andet [because it is there in stead of something else]
04	ahm...
05	<i>very good book</i> det er en objekt group, hvor <i>very good book</i> [<i>very good book</i> that's an object group where <i>very good book</i> that's]
06	...ahmm....
07	head det må være <i>book</i> , som noun [head that must be <i>book</i>]
08	og dependent er article <i>a</i> [and dependent that is article <i>a</i>]
09	ligemeget (skriver ingenting) [no matter (doesn't write anything)]
10	og dependent er [and dependent that is]
11	...
12	<i>very</i> som er adjektiv [<i>very</i> which is adjective]
13	og dependent igen som er adjektiv, <i>good</i> [and dependent again which is adjective <i>good</i>]
14	ja [yes]

Full protocol:

Is det er predicator, verbum.. *It* er subjektet, som er pronoun, fordi det står i stedet for noget andet ahm...*very good book* det er en objekt group, hvor...ahmm.... head det må være *book*, som noun, og dependent er article *a*, ligemeget (skriver ingenting), og dependent er.. *very* som er adjektiv, og dependent igen som er adjektiv, *good*. Ja.

TA Excerpt 7.19

TAS15. POST-Sentence 1.

POST-Sentence 1: *Jill sold the car today*

Segment	Transcription
01	<i>sold</i> det er predicator [<i>sold</i> that's predicator]
02	<i>Jill</i> , det er subject [<i>Jill</i> that's subject]
03	...
04	(sighs)
05	<i>the car</i>
06	ehh...
07	it's a group
08	but..eh...
09	and object I think
10	<i>today</i> hmm...
11	det kan jeg ikke huske [I don't remember]
12	...eh...
13	nå [well]
14	jeg går lige til den næste [I'll just take the next one] [goes on to do sentence 2 and analyses sentence 2] [returns to sentence 1]
15	jeg kan altså ikke huske <i>today</i> [I really don't remember <i>today</i>]
16	så jeg tror, jeg bliver nødt til at stoppe [so I think I have to quit]

Full protocol:

sold det er predicator. *Jill*, det er subject(sighs) *the car*, eh.. it's a group but..erh..and object, I think, group, *today* hrh..hrh..(goes back to the car/g) that's a DEP, and that's article, and that's head and noun, *today* hmm...det kan jeg ikke huske...ahrh... nå, jeg går lige til den næste.
[goes to sentence 2 and analyses sentence 2] [returns to sentence 1]
Jeg kan altså ikke huske *today*, så jeg tror, jeg bliver nødt til at stoppe.

TA Excerpt 7.20

TAS5. POST-Sentence 1.

POST-Sentence 1: *Jill sold the car today*

Segment	Transcription
[01-03]	
04	and this (<i>sold</i>) is the predicator
05	it's a verb
06	and the subject (also picks noun)
07	and then I group these two (<i>the car</i>)
08	ja
09	and it should be the object
10	and it's a group

