

CONTENTS

CONTENTS	III
TABLES	IX
FIGURES	XIII
LIST OF ABBREVIATIONS	XV
 PREFACE	 1
Previous Publications	4
Thesis Overview	5
 CHAPTER I: BACKGROUND & INTRODUCTION	 7
1. The INCOMSLAV Infrastructure	7
1.1. Languages in Focus and INCOMSLAV Publications	9
1.2. Regular Orthographic Correspondences Between PL-CS and BG-RU	9
1.2.1. Hand-crafted correspondences inferred from traditional linguistic assumptions	11
1.2.2. Results of the application of correspondences	11
1.3. Similarity of Linguistic Encoding	14
1.3.1. Lexical distance	16
1.3.2. Orthographic distance	19
1.3.3. Other distance measures in the literature	22
1.4. Asymmetry in Cross-Lingual Intelligibility	23
1.4.1. Conditional entropy	23
1.4.2. Word adaptation surprisal	27
1.5. Surprisal and Context	27
1.6. Context in Intercomprehension	31
2. Thesis Focus: Modelling Linguistic Phenomena of PL for Czech Readers	33

CHAPTER II: COOPERATIVE TRANSLATION EXPERIMENT . . .	35
3. Experimental Setup	35
4. Quantitative Analysis of Written Results and Comparison of Conditions	37
4.1. Hypotheses	37
4.2. Stimuli	39
4.2.1. Linguistic distance of stimuli	41
4.2.2. Surprisal of stimuli	42
4.3. Experimental Conditions: Modification of Stimuli on Different Linguistic Levels	42
4.3.1. Conditions with non-combined modifications	43
4.3.2. Conditions with combined modifications	44
4.4. Results	47
4.4.1. Evaluation of the translations per word	47
4.4.2. Comparison between the conditions	49
4.5. Summary	52
5. Qualitative Analysis	53
5.1. Readers' Strategies	55
5.1.1. Leaving unknown words open and trying to infer them from the context	55
5.1.2. Recognition order as indicator for difficulty	57
5.1.3. Reading again and pronouncing differently	61
5.2. Source of Successful Transfer	61
5.2.1. Inference processes from non-standard CS	62
5.2.2. Inference from languages other than CS	62
5.3. Knowledge of Non-Cognates and Awareness of False Friends	71
5.4. Over-Transfer from Languages Other Than CS	76
5.5. Distrust in Obviously Understandable Words	77
5.6. Revision After Having Already Named the Correct Answer .	85
5.7. Handling Unfamiliar PL Orthography	86
5.7.1. Handling PL diacritics	87
5.7.1.1. Respondents pronounce letters correctly	92
5.7.1.2. Respondents ignore diacritics and pronounce stimulus as if without diacritics	95

5.7.1.3.	Respondents move diacritics to another suitable letter in the word	96
5.7.2.	Handling unfamiliar PL digraphs	97
5.8.	Talking About Grammar	101
5.9.	Problems Caused by Differences in Government Patterns ..	103
5.10.	Problems Caused by Different Prepositions	104
5.11.	Summary	108
CHAPTER III: ON-LINE EXPERIMENTS		111
6.	Hypotheses	111
6.1.	Pronunciation-Based Orthographic Distance	111
6.2.	Surprisal as a Predictor Variable for Context in Intercomprehension	114
7.	Empirical Base	115
7.1.	Online Experiments	115
7.2.	Overview of Experiments and Data Collected	120
7.3.	Participants	121
8.	The Principle of the Closest Possible Translation	122
9.	Measures not Considered	123
10.	Scoring Policy Throughout the Experiments	124
11.	Relevant Statistical Methods in Brief	125
CHAPTER IV: FREE TRANSLATION OF WORDS		
	WITHOUT CONTEXT	127
12.	Cognates with Regular PL-CS Orthographic Correspondences	127
12.1.	Orthographic Distance of the Stimuli	128
12.2.	Results	129
12.3.	Correlations	130
12.4.	Error Analysis	132
12.5.	Summary	135
13.	The 100 Most Frequent PL Nouns	135
13.1.	Results and Correlations	136
13.2.	Error Analysis	138
13.3.	Summary and Outlook	139

VI

CHAPTER V: FREE TRANSLATION OF NPS	141
14. Adjectival Modification in PL	142
14.1. Hypothesis	145
14.2. Method	145
14.3. Distance of the Stimuli	145
14.4. Total Difficulty of the Stimuli	146
14.5. Results	148
14.5.1. Intelligibility	148
14.5.2. Processing time	150
14.5.1. Wrong recognition of POS	151
14.6. Summary	152
14.7. Digression: PL NPs Presented to German Readers	153
14.7.1. Hypothesis	153
14.7.2. Stimuli	153
14.7.3. Orthographic distance	154
14.7.4. Surprisal in context	154
14.7.5. Results	155
14.7.5.1. Relation between intelligibility and orthographic distance ...	156
14.7.5.2. Relation between intelligibility and surprisal	157
14.7.5.3. Mean processing time	157
14.7.5.4. Wrong recognition of POS	158
14.7.5.5. Lexical interferences	160
14.7.6. Summary	161
14.8. Comparison of PL NP Results Between Czech and German Readers	162
CHAPTER VI: TRANSLATION OF TARGET WORDS IN CONTEXT	165
15. Highly Predictable Target Words in Cloze Translation Task .	165
15.1. Experiment Design	165
15.2. Stimuli	166
15.2.1. Closest translation	167
15.2.2. Surprisal	169
15.2.3. Linguistic distance	170

15.3.	Scoring of Responses	171
15.4.	Results	172
15.4.1.	Comparison: with vs. without context	172
15.4.2.	Different lexical categories of target words	177
15.4.2.1.	Cognates (C)	178
15.4.2.2.	Cognates in other contexts (C-OC)	180
15.4.2.3.	Non-cognates (NC)	181
15.4.2.4.	False friends (FFs)	182
15.4.3.	Analysis of wrong responses	187
15.4.3.1.	Differences in government pattern	188
15.4.3.2.	Ln interferences	188
15.4.3.3.	(Perceived) morphological mismatches	189
15.5.	Correlations and Model	198
15.6.	Summary and Discussion	200
16.	The Impact of Random Context on the Understanding of Particular Words in Sentences from the Cooperative Translation Task	201
16.1.	Method	202
16.2.	Baseline Experiments: Cloze Probabilities in Monolingual Context	205
16.2.1.	Design	205
16.2.2.	Results	205
16.3.	Scoring of Responses	207
16.4.	Results: Target Words at Random Position	208
16.4.1.	Comparison: Types of errors	210
16.4.2.	Comparison: Target words with vs. without context	214
16.5.	Summary	215
CHAPTER VII: CONCLUSION AND OUTLOOK		217
Conclusion in German Deutsche Zusammenfassung		227

VIII

APPENDICES	239
A 1. Alignment Matrices of the PL and CS Alphabets	239
A 1.1 For the Calculation of Trad LD	239
A 1.2 For the Calculation of Pron LD	241
A 2. Questionnaire on Sociodemographic Data	243
A 3. Instruction for the Participants in the Cooperative Translation Experiments	244
A 4. Intelligibility of Stimuli in the Different Experiments	245
A 4.1. Stimuli with Regular PL-CS Correspondences	245
A 4.2. Most Frequent PL Nouns	259
A 4.3. Free Translation of NPs	262
A 4.3.1. PL NPs for CS readers with the most representative data	262
A 4.3.2. PL NP stimuli for German readers	263
A 4.4. Highly Predictable Target Words	264
A 5. Target Words in Highly Predictive Context Categorised as FFs	269
A 5.1. False Friends that Are also Cognates – FF-C	269
A 5.2. False Friends that Are Cognates in Another Context – FF-OC	270
A 5.3. False Friends that Allow for Correct Associations – FF-A ..	271
A 5.4. False Friends – FF	272
A 6. Monolingual Cloze Tests	273
A 6.1. Task in Monolingual Cloze Tests	273
A 6.2. Stimuli	273
A 7. Correlations and Statistical Models	281
A 7.1. Intelligibility of the 100 Most Frequent PL Ns	281
A 7.2. Intelligibility of NPs for Czech Readers – AN Condition ...	282
A 7.3. Intelligibility of NPs for Czech Readers – NA Condition ...	283
A 7.4. Intelligibility of Target Words in Highly Predictive Context ..	284
A 7.5. Model for Intelligibility of Target Words in Highly Predictive Context	286
A 7.6. Model for Intelligibility of the Target Words Without Context	287
REFERENCES	289