

Inhaltsverzeichnis

Abbildungsverzeichnis	XI
Tabellenverzeichnis	XV
Quellcode-Verzeichnis	XIX
Abkürzungsverzeichnis	XXIII
1 Einleitung	1
1.1 Gegenstand	4
1.2 Nutzen	5
1.3 Einordnung in die Fachgebiete: Das Verhältnis von Computerlinguistik und Sprachwissenschaft	6
1.4 Ziele	19
1.5 Vorbemerkungen zur Notation	21
1.6 Aufbau	21
2 Stand der Forschung	23
2.1 Überblick	23
2.2 Burton-Hunter (1976)	32
2.3 Eastlack (1977): <i>Iberochange</i>	32
2.4 Hartman (1981–2018): <i>Phono</i>	33
2.5 Schmuki (2000): <i>Etymo</i>	36
2.6 Kritische Zusammenfassung	38
3 Praktische Umsetzung	43
3.1 Allgemeine Betrachtungen	44

3.1.1	Aufbau des Systems	45
3.1.2	Merkmalstrukturen	48
3.2	Wortformerkenung	49
3.2.1	LA-Grammatik und JSLIM	49
3.2.2	Aufbau und Funktionsweise der Wortformerkenungskomponente im <i>CEW</i>	55
3.2.3	Unterschiede zu JSLIM	63
3.2.4	Umsetzung im C#-Code	65
3.2.5	Zusammenfassung	67
3.3	Phonetisch-phonologische Vorverarbeitung und Nachbereitung	68
3.3.1	Transkription	69
3.3.2	Silbentrennung	71
3.3.3	Betonung	74
3.3.4	Umsetzung im C#-Code	75
3.3.5	Zusammenfassung	77
3.4	Lautentwicklungen	78
3.4.1	Regelaufbau und Regelbedingungen	80
3.4.2	Mapping von Lauten und Graphemen	84
3.4.3	Weitere Änderungen	97
3.4.4	Umsetzung im C#-Code	98
3.4.5	Zusammenfassung	99
3.5	Morphologiebasierte Entwicklungen	100
3.6	Abgleich zwischen Quell- und Zielsprache	103
3.7	Graphische Oberfläche	104
3.8	Zusammenfassung	108
4	Die Umsetzung des <i>CEW</i> am Beispiel des Französischen (<i>CEW-fr</i>)	109
4.1	Sprachgeschichtlicher Überblick	109
4.1.1	Externe Sprachgeschichte	110
4.1.2	Definitionen und damit verbundene Schwierigkeiten aus compu- terlinguistischer Sicht	112
4.1.3	Interne Sprachgeschichte	118
4.2	Implementierung der lateinischen Wortformerkenung	122

4.2.1	Ausgangsdaten	124
4.2.2	Substantive	127
4.2.3	Adjektive	133
4.2.4	Verben	136
4.2.5	Unregelmäßige Paradigmen	142
4.2.6	Weitere Wortarten	143
4.2.7	Zusammenfassung	144
4.3	Phonetisch-phonologische Vorverarbeitung der lateinischen Wortform . .	145
4.3.1	Transkription der lateinischen Grapheme	145
4.3.2	Lateinische Silbentrennung	154
4.3.3	Betonung lateinischer Wortformen	156
4.3.4	Zusammenfassung	158
4.4	Implementierung von Lautregeln	158
4.4.1	Die Wahl einer geeigneten Datengrundlage	158
4.4.2	Methodologie und Aufbau von Rheinfelders <i>Altfranzösische Gram-</i> <i>matik</i>	162
4.4.3	Überblick über die Implementierung	165
4.5	Morphologiebasierte Entwicklungen am Beispiel des Deklinationssystems	183
4.5.1	Vulgärlateinische Veränderungen	183
4.5.2	Entwicklungen zum Altfranzösischen	187
4.5.3	Entwicklungen zum Mittel- und Neufranzösischen	196
4.5.4	Zusammenfassung	197
4.6	Implementierung der französischen Wortformerkennung	197
4.6.1	Ausgangsdaten	198
4.6.2	Substantive	200
4.6.3	Adjektive	202
4.6.4	Verben	204
4.6.5	Unregelmäßige Paradigmen	211
4.6.6	Nicht flektierbare Wortarten	211
4.6.7	Zusammenfassung	212
4.7	Phonetisch-phonologische Nachbereitung der französischen Wortform:	
	Transkription	213
4.8	Abgleich der lateinischen und neufranzösischen Wortform	235

4.9 Zusammenfassung	236
5 Evaluierung des Systems	237
5.1 Quantitative Analyse	237
5.2 Qualitative Analyse	240
5.3 Schwierigkeiten	248
5.3.1 Wortformerkennung	249
5.3.2 Phonetisch-phonologische Vorverarbeitung und Nachbereitung . .	250
5.3.3 Sprachwandel	257
5.3.4 Abgleich der lateinischen und neufranzösischen Wortform	266
5.4 Zusammenfassung	266
6 Zusammenfassung und Ausblick	269
6.1 Umsetzung der Ziele	269
6.2 Ausbaumöglichkeiten	272
6.2.1 <i>CEW-fr</i>	272
6.2.2 <i>CEW</i>	273
Literaturverzeichnis	277
Software und elektronische Datensätze	297
Anhänge	
A Morphologie der lateinischen Verben im <i>CEW-fr</i>	299
A.1 Erste Konjugation (Flexionsklasse <i>cantare</i>)	299
A.2 Zweite Konjugation (Flexionsklasse <i>tenere</i>)	300
A.3 Dritte Konjugation (Flexionsklasse <i>regere</i>)	301
A.4 Vierte Konjugation (Flexionsklasse <i>punire</i>)	302
B Morphologie der französischen Verben im <i>CEW-fr</i>	305
B.1 Flexionsklasse <i>chanter</i>	305
B.2 Flexionsklasse <i>appeler</i>	306
B.3 Flexionsklasse <i>appuyer</i>	306
B.4 Flexionsklasse <i>lever</i>	307

B.5 Flexionsklasse <i>préférer</i>	308
B.6 Flexionsklasse <i>changer</i>	309
B.7 Flexionsklasse <i>commencer</i>	309
B.8 Flexionsklasse <i>assiéger</i>	310
B.9 Flexionsklasse <i>apiécer</i>	311
B.10 Flexionsklasse <i>sortir</i>	312
B.11 Flexionsklasse <i>agir</i>	312
B.12 Flexionsklasse <i>apercevoir</i>	313
B.13 Flexionsklasse <i>ouvrir</i>	314
B.14 Flexionsklasse <i>rompre</i>	315
B.15 Flexionsklasse <i>vendre</i>	315
B.16 Flexionsklasse <i>battre</i>	316
B.17 Flexionsklasse <i>prendre</i>	317
B.18 Flexionsklasse <i>conduire</i>	318
B.19 Flexionsklasse <i>coudre</i>	318
B.20 Flexionsklasse <i>croire</i>	319
B.21 Flexionsklasse <i>accroître</i>	320
B.22 Flexionsklasse <i>atteindre</i>	321
C Liste der verwendeten Symbole im <i>CEW</i>	323
C.1 Symbole in regulären Ausdrücken	323
C.2 Allgemeine Symbole im <i>CEW</i>	324
C.3 Sprachspezifische Symbole	324
C.4 Verwendete X-SAMPA-Symbole	325
D Ausschnitte aus den einzelsprachlichen Textdateien	327
D.1 Lateinische Wortformererkennung	327
D.1.1 Lexikondateien	327
D.1.2 Allomorphiedateien	330
D.1.3 Flexionsdateien	331
D.1.4 Weitere Dateien	336
D.2 Phonetisch-phonologische Vorverarbeitung und Nachbereitung	337
D.3 Sprachwandel	339

D.4	Französische Wortformerkennung	349
D.4.1	Lexikondateien	349
D.4.2	Allomorphiedateien	351
D.4.3	Flexionsdateien	353
D.4.4	Weitere Dateien	358