

Inhaltsverzeichnis

1	Einleitung	5
1.1	Motivation	5
1.2	Aufgabenstellung	5
2	Lëtzebuergesch	6
2.1	Geschichte Luxemburgs	6
2.2	Luxemburgisch	7
2.3	Dialekte	8
2.4	Koiné	8
2.5	Phonetik der luxemburgischen Koiné	8
2.6	Phonetik des Deutschen	9
2.7	Gegenüberstellung des Deutschen und des Lëtzebuergesch	11
3	Phonetische Alignierung als Erkennungsproblem	12
3.1	Toolgestützte Spracherkennung mit dem Aligner	12
3.2	Funktionsweise	12
3.3	Phoneminventar und Aussprachelexikon	13
3.4	Abbildung der lëtzebuergesch Phoneme auf deutsche Sprachlaute	14
4	Aufbereitung der Sprachdaten	14
4.1	Vorhandene Sprachdaten	14
4.2	Aufteilung der Datensätze	16
4.3	Vorbereitung der Daten - Gilles-Korpus	16
4.3.1	Orthographische Transkription	17
4.3.2	Erstellung eines Aussprachelexikons	18
4.3.3	Generierung der phonetischen Transkriptionen	19
4.3.4	Korpusstruktur	21
4.4	Vorbereitung der Daten - das 6000-Wierder-Korpus	22
4.4.1	Bereinigung des Datensatzes	22
4.4.2	Aufteilung für Test und Training	23
4.5	Zusammenfassung	23

5 Implementierung	23
5.1 Alignieren der phonetischen Annotation mithilfe des deutschen Aligners	24
5.2 Aligierte phonetische Annotation - Ersetzung der deutschen Phonemmarkierungen mit den lëtzebuergeschen	24
5.3 Erstellung der Hidden-Markov-Modelle für lëtzebuergesche Phoneme	24
5.4 Integration der neuen Modelle in den Aligner.	26
5.5 Anpassung der Vorverarbeitungskomponenten des Aligners	26
5.6 Zusammenfassung	27
6 Evaluierung	27
6.1 Vergleich der Aligner	27
6.2 Programm zur automatischen Evaluierung der Ergebnisse	27
6.3 Korrektur und zweiter Vergleich	28
6.4 Manuelle und automatische Alignierung	29
6.5 Ergebnisse	29
7 Zusammenfassung und zukünftige Forschung	30
7.1 Erweiterung der Funktionalität des Aligners	31
7.2 Entwicklung eines lëtzebuergschen Aligners	32
7.3 Entwicklung eines Sprachkorpus für Lëtzebuergesch	32
A Gilles-Korpus: Fragebuch Lëtzebuergesch	36
B Die Perl-Funktionssammlung ConversionLibrary.pm	38