

Contents

1	Introduction	1
2	Historical Linguistics	9
2.1	Entities	10
2.1.1	Languages	11
2.1.2	Words	15
2.1.3	Forms	18
2.1.4	Meanings	21
2.1.5	Representations	22
2.2	Change	24
2.2.1	Sound Change	26
2.2.2	Semantic Change	34
2.2.3	Lexical Change	37
2.3	Relations	39
2.3.1	Relations between Words	41
2.3.2	Relations between Languages	44
2.4	Resemblances	47
2.4.1	Resemblances in Form	48
2.4.2	Resemblances in Meaning	50
2.5	Proof	51
2.5.1	Laws	54
2.5.2	Evidence	55
2.6	Methods	57
3	Sequence Comparison	61
3.1	Sequences	62
3.1.1	Discrete and Continuous Entities	63
3.1.2	Sequences and Sets	64
3.1.3	A Formal Definition of Sequences	64

3.2	Sequence Comparison	66
3.2.1	Differences between Sequences	67
3.2.2	Modelling Sequence Differences	69
3.2.3	Alignment Analyses	72
3.3	Pairwise Alignment Analyses	76
3.3.1	The Basic Algorithm for Pairwise Alignment	77
3.3.2	Structural Extensions of the Basic Algorithm	82
3.3.3	Substantial Extensions of the Basic Algorithm	91
3.3.4	Summary on Pairwise Alignment Analyses	99
3.4	Multiple Alignment Analyses	99
3.4.1	Progressive Alignment	100
3.4.2	Enhancements for Progressive Alignment	108
3.4.3	Summary on Multiple Alignment Analyses	115
4	Sequence Comparison in Historical Linguistics	117
4.1	Sequence Modelling	119
4.1.1	Paradigmatic Aspects	121
4.1.2	Syntagmatic Aspects	130
4.2	Phonetic Alignment	134
4.2.1	Previous Work	137
4.2.2	SCA – Sound-Class Based Phonetic Alignment	139
4.2.3	Specific Features	142
4.2.4	Evaluation	147
4.2.5	Examples	165
4.3	Automatic Cognate Detection	169
4.3.1	Previous Work	172
4.3.2	LexStat – Multilingual Cognate Detection	173
4.3.3	Specific Features	175
4.3.4	Evaluation	185
4.3.5	Examples	205
5	Conclusion	209
	Bibliography	215
	Sources	229
	Dictionaries and Databases	231

Supplementary Material	235
List of Abbreviations	237
Indices	239
Index of Topics	239
Index of Persons	247
Index of Languages	248
Index of Word Forms	250
Index of Reconstructed Forms	254
Appendix A Language-Specific Resources	255
A.1 Phonetic Transcriptions	255
A.2 Etymological Sources	256
Appendix B Sequence Modelling	257
B.1 The DOLGO Scoring function	257
B.2 The SCA Scoring function	258
B.3 The ASJP Scoring function	259
Appendix C Phonetic Alignment	261
C.1 Covington's Testset	261
C.2 Language-Specific Pairwise Alignments	262
Appendix D Cognate Detection	271
D.1 Comparison of Phonetic Distances	271
D.2 Cognate Detection	274