



HAL
open science

**Companion report to the Ph.D. dissertation:
"Generative Probabilistic Alignment Models for Words
and Subwords: a Systematic Exploration of the Limits
and Potentials of Neural Parametrizations"**

Anh Khoa Ngo Ho

► **To cite this version:**

Anh Khoa Ngo Ho. Companion report to the Ph.D. dissertation: "Generative Probabilistic Alignment Models for Words and Subwords: a Systematic Exploration of the Limits and Potentials of Neural Parametrizations". [Technical Report] Université Paris Saclay; Laboratoire Interdisciplinaire des Sciences du Numérique. 2021. hal-03153752

HAL Id: hal-03153752

<https://hal.science/hal-03153752>

Submitted on 1 Apr 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Companion report to the Ph.D. dissertation:
"Generative Probabilistic Alignment Models
for Words and Subwords: a Systematic
Exploration of the Limits and Potentials of
Neural Parametrizations"

Anh Khoa NGO HO

09/02/2021

Abstract: This is a companion document to the Ph.D dissertation "Generative Probabilistic Alignment Models for Words and Subwords: a Systematic Exploration of the Limits and Potentials of Neural Parametrizations" [Ngo Ho, 2021]. This document contains an exhaustive collection of graphs and tables related the analysis of various aspects of automatic word alignment, such as for instance the aligned/unaligned words, rare/unknown words, function/content words, word order divergences, etc; and for six language pairs: English with French, German, Romanian, Czech, Japanese and Vietnamese. We mostly analyze statistical word alignment models (**Giza++** and **Fastalign**) as well as several variants based on neural models: IBM style word alignment models including context-independent models, contextual models, and character-based models; variants of a fully generative neural model based on variational autoencoders. We also document a deep analysis for Byte-Pair-Encoding, a subword tokenization algorithm. For information regarding these various methods, please refer to the thesis.

Outline: Appendix A displays the performance of the baselines, namely **Giza++** and **Fastalign**, related to aligned/unaligned words, rare/unknown words, function/content words, word order divergences, alignment types, symmetrization and sequence lengths. Appendix B is for IBM style word alignment models including context-independent models, contextual models, and character-based models. Appendix C presents an analysis of word alignment based on attention models. Appendix D details all results for the variants of a fully generative neural model based on variational autoencoders. The subword tokenization algorithm (Byte-Pair-Encoding) is analyzed in Appendix E. We compare our best results with other published numbers for English vs French, German, Romanian and Japanese in Appendix F. Note that all descriptions about the figures and tables are presented in Ngo Ho [2021].

Contents

Contents	5
List of Figures	30
List of Tables	37
Appendices	39
A Appendix: Evaluation methods	41
A.1 Baselines: Scores of the baselines	41
A.2 Baselines: Correct and Incorrect alignment links	44
A.3 Baselines: Unaligned words	47
A.4 Alignment types	53
A.4.1 Baselines: Alignment types	53
A.4.2 Baselines: Correct/incorrect links as a function of alignment types	55
A.5 Jump width statistics	57
A.5.1 Baselines: Jump width statistics	63
A.5.2 Baselines: Correct and incorrect jumps	69
A.6 Word length as a function of word occurrence	82
A.7 Baselines: Rare words	84
A.8 Baselines: Unknown words	85
A.9 Baselines: Part-of-Speech	88
A.9.1 Baselines: Alignment links for content and function words	88
A.9.2 Baselines: Unaligned words for content and function words	94
A.10 Baselines: Symmetrization	99
A.10.1 Baselines: Intersection	99
A.10.2 Baselines: Grow-diag-final (GDF)	100
A.11 Word-based sentence length statistics	102
A.11.1 Sentence length statistics	102
A.11.2 Length difference statistics between English and foreign sentence	105
A.11.3 Sentence length and unknown/rare words	107
A.11.4 Sentence length and word repetition	109
A.11.5 Baselines: AER as a function of sentence length	111
A.11.6 Baselines: AER as a function of sentence length difference	123
B Appendix: Neural word alignment models	133
B.1 Scores	133
B.2 Correct and incorrect alignment links	137
B.3 Unaligned words	139
B.4 Alignment types	145
B.5 Jump widths	147
B.5.1 Jump width statistics	147
B.5.2 Correct and incorrect jumps	153

B.6	Rare words	186
B.7	Unknown words	189
	B.7.1 Unknown words in the case of a full vocabulary	189
	B.7.2 Unknown words in the case of 50K word vocabulary	192
B.8	Part-of-Speech	196
	B.8.1 Alignment links for content and function words	196
	B.8.2 Unaligned words for content and function words	202
B.9	Symmetrization	208
	B.9.1 Intersection	208
	B.9.2 Grow-diag-final (GDF)	211
B.10	AER as a function of sentence length	214
B.11	AER as a function of sentence length difference	242
C	Appendix: Word alignment by attention	271
	C.1 Scores	271
	C.2 Correct and incorrect alignment links	272
	C.3 Unaligned words	274
	C.4 Alignment types	278
	C.5 AER as a function of sentence length	279
	C.6 Symmetrization	285
	C.6.1 Intersection	285
	C.6.2 Grow-diag-final	286
D	Appendix: Generative latent neural alignment models	289
	D.1 Scores	289
	D.2 Correct and incorrect alignment links	293
	D.3 Unaligned words	295
	D.4 Alignment types	299
	D.4.1 Correct/incorrect links as a function of alignment types	301
	D.5 Jump widths	303
	D.5.1 Jump width statistics	303
	D.5.2 Correct and incorrect jumps	309
	D.6 Rare words	333
	D.7 Unknown words	337
	D.8 Part-of-Speech	340
	D.8.1 Alignment links for content and function words	340
	D.8.2 Unaligned words for content and function words	346
	D.9 Symmetrization	352
	D.9.1 Intersection	352
	D.9.2 Grow-diag-final (GDF)	356
	D.10 Training with monolingual data	360
	D.11 AER as a function of sentence length	360
	D.12 AER as a function of sentence length difference	383
E	Appendix: Subwords for word alignment task	405
	E.1 Scores	405
	E.1.1 Fastalign: Matrix of scores	405
	E.1.2 Fastalign: Scores of BPE-based vs word-based alignment	409
	E.1.3 Fastalign with BPE-dropout: Matrix of scores	422
	E.1.4 Eflomal: Matrix of scores	426
	E.1.5 Neural models with recommended BPE configurations	430
	E.2 Correct and incorrect alignment links	433
	E.3 Unaligned words	436

E.4	Alignment types	442
E.5	Rare words	457
E.5.1	Matrix of scores	457
E.5.2	Scores of BPE-based vs word-based alignment	464
E.6	Number of fragments as a function of word occurrence	479
E.7	Number of one-to-many/many-to-one as a function of word occurrence	485
E.8	Part-of-Speech	546
E.8.1	Alignment links for content and function words	546
E.8.2	Unaligned words for content and function words	552
E.9	Symmetrization	558
E.9.1	GDF before the transformation step	558
E.9.2	GDF after the transformation step	561
E.10	Scores as a function of sentence length	585
E.11	AER as a function of sentence length difference	651
F	Appendix: Our models versus models from related works	665

List of Figures

A.1	Results of our baselines: Alignment links for English-French	44
A.2	Results of our baselines: Alignment links for English-German	44
A.3	Results of our baselines: Alignment links for English-Romanian	45
A.4	Results of our baselines: Alignment links for English-Czech	45
A.5	Results of our baselines: Alignment links for English-Japanese	45
A.6	Results of our baselines: Alignment links for English-Vietnamese	46
A.7	Results of our baselines: Unaligned words for English-French	47
A.8	Results of our baselines: Unaligned words for English-German	48
A.9	Results of our baselines: Unaligned words for English-Romanian	49
A.10	Results of our baselines: Unaligned words for English-Czech	50
A.11	Results of our baselines: Unaligned words for English-Japanese	51
A.12	Results of our baselines: Unaligned words for English-Vietnamese	52
A.13	Results of our baselines: Alignment types for English-French	53
A.14	Results of our baselines: Alignment types for English-German	53
A.15	Results of our baselines: Alignment types for English-Romanian	53
A.16	Results of our baselines: Alignment types for English-Czech	54
A.17	Results of our baselines: Alignment types for English-Japanese	54
A.18	Results of our baselines: Alignment types for English-Vietnamese	54
A.19	Results of our baselines: Alignment types for English-French	55
A.20	Results of our baselines: Alignment types for English-German	55
A.21	Results of our baselines: Alignment types for English-Czech	55
A.22	Results of our baselines: Alignment types for English-Romanian	56
A.23	Results of our baselines: Alignment types for English-Japanese	56
A.24	Results of our baselines: Alignment types for English-Vietnamese	56
A.25	Jump patterns for English words in the direction French-English and jump patterns for foreign words in the direction English-French	57
A.26	Jump patterns for English words in the direction German-English and jump patterns for foreign words in the direction English-German	58
A.27	Jump patterns for English words in the direction Romanian-English and jump patterns for foreign words in the direction English-Romanian	59
A.28	Jump patterns for English words in the direction Czech-English and jump patterns for foreign words in the direction English-Czech	60
A.29	Jump patterns for English words in the direction Japanese-English and jump patterns for foreign words in the direction English-Japanese	61
A.30	Jump patterns for English words in the direction Vietnamese-English and jump patterns for foreign words in the direction English-Vietnamese	62
A.31	Baselines: Jump patterns for English words in the direction French-English and jump patterns for foreign words in the direction English-French	63
A.32	Baselines: Jump patterns for English words in the direction German-English and jump patterns for foreign words in the direction English-German	64
A.33	Baselines: Jump patterns for English words in the direction Romanian-English and jump patterns for foreign words in the direction English-Romanian	65

A.34	Baselines: Jump patterns for English words in the direction Czech-English and jump patterns for foreign words in the direction English-Czech	66
A.35	Baselines: Jump patterns for English words in the direction Japanese-English and jump patterns for foreign words in the direction English-Japanese	67
A.36	Baselines: Jump patterns for English words in the direction Vietnamese-English and jump patterns for foreign words in the direction English-Vietnamese	68
A.37	Baselines for the direction English-French: Correct (TP) and incorrect (FP) jump widths for French words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	70
A.38	Baselines for the direction French-English: Correct (TP) and incorrect (FP) jump widths for English words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	71
A.39	Baselines for the direction English-German: Correct (TP) and incorrect (FP) jump widths for German words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	72
A.40	Baselines for the direction German-English: Correct (TP) and incorrect (FP) jump widths for English words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	73
A.41	Baselines for the direction English-Czech: Correct (TP) and incorrect (FP) jump widths for Czech words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	74
A.42	Baselines for the direction Czech-English: Correct (TP) and incorrect (FP) jump widths for English words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	75
A.43	Baselines for the direction English-Romanian: Correct (TP) and incorrect (FP) jump widths for Romanian words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	76
A.44	Baselines for the direction Romanian-English: Correct (TP) and incorrect (FP) jump widths for English words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	77
A.45	Baselines for the direction English-Japanese: Correct (TP) and incorrect (FP) jump widths for Japanese words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	78
A.46	Baselines for the direction Japanese-English: Correct (TP) and incorrect (FP) jump widths for English words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	79
A.47	Baselines for the direction English-Vietnamese: Correct (TP) and incorrect (FP) jump widths for Vietnamese words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	80
A.48	Baselines for the direction Vietnamese-English: Correct (TP) and incorrect (FP) jump widths for English words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	81
A.49	The direction English-French: Word length as a function of word occurrence. . .	82
A.50	The direction English-German: Word length as a function of word occurrence. . .	82
A.51	The direction English-Czech: Word length as a function of word occurrence. . .	82
A.52	The direction English-Romanian: Word length as a function of word occurrence. . .	83
A.53	The direction English-Japanese: Word length as a function of word occurrence. . .	83
A.54	The direction English-Vietnamese: Word length as a function of word occurrence. . .	83
A.55	Baselines for English-French: The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).	88

A.56	Baselines for English-German: Baselines for English-French: The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).	89
A.57	Baselines for English-Czech: Baselines for English-French: The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).	90
A.58	Baselines for English-Romanian: Baselines for English-French: The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).	91
A.59	Baselines for English-Japanese: Baselines for English-French: The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).	92
A.60	Baselines for English-Vietnamese: Baselines for English-French: The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).	93
A.61	Baselines for English-French: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).	94
A.62	Baselines for English-German: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).	95
A.63	Baselines for English-Czech: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).	95
A.64	Baselines for English-Romanian: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).	96
A.65	Baselines for English-Japanese: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).	97
A.66	Baselines for English-Vietnamese: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).	98
A.67	Sentence lengths in English-French	102
A.68	Sentence lengths in English-German	102
A.69	Sentence lengths in English-Romanian	103
A.70	Sentence lengths in English-Czech	103
A.71	Sentence lengths in English-Japanese	103
A.72	Sentence lengths in English-Vietnamese	104
A.73	Length differences in English-French	105
A.74	Length differences in English-German	105
A.75	Length differences in English-Romanian	105
A.76	Length differences in English-Czech	106
A.77	Length differences in English-Japanese	106
A.78	Length differences in English-Vietnamese	106
A.79	Number of unknown/rare words as a function of sentence length for English-French	107
A.80	Number of unknown/rare words as a function of sentence length for English-German	107
A.81	Number of unknown/rare words as a function of sentence length for English-Czech	107

A.82	Number of unknown/rare words as a function of sentence length for English-Romanian	108
A.83	Number of unknown/rare words as a function of sentence length for English-Japanese	108
A.84	Number of unknown/rare words as a function of sentence length for English-Vietnamese	108
A.85	Number of words that repeat at least twice as a function of sentence length for English-French	109
A.86	Number of words that repeat at least twice as a function of sentence length for English-German	109
A.87	Number of words that repeat at least twice as a function of sentence length for English-Czech	109
A.88	Number of words that repeat at least twice as a function of sentence length for English-Romanian	110
A.89	Number of words that repeat at least twice as a function of sentence length for English-Japanese	110
A.90	Number of words that repeat at least twice as a function of sentence length for English-Vietnamese	110
A.91	Baselines for the direction English-French: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	111
A.92	Baselines for the direction French-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	112
A.93	Baselines for the direction English-German: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	113
A.94	Baselines for the direction German-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	114
A.95	Baselines for the direction English-Czech: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	115
A.96	Baselines for the direction Czech-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	116
A.97	Baselines for the direction English-Romanian: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	117
A.98	Baselines for the direction Romanian-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	118
A.99	Baselines for the direction English-Japanese: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	119
A.100	Baselines for the direction Japanese-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	120
A.101	Baselines for the direction English-Vietnamese: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	121

A.10	Baselines for the direction Vietnamese-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	122
A.10	Baselines for the direction English-French: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	123
A.10	Baselines for the direction French-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	124
A.10	Baselines for the direction English-German: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	124
A.10	Baselines for the direction German-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	125
A.10	Baselines for the direction English-Czech: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	126
A.10	Baselines for the direction Czech-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	127
A.10	Baselines for the direction English-Romanian: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	127
A.11	Baselines for the direction Romanian-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	128
A.11	Baselines for the direction English-Japanese: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	129
A.11	Baselines for the direction Japanese-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	129
A.11	Baselines for the direction English-Vietnamese: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	130
A.11	Baselines for the direction Vietnamese-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	131
B.1	Results of our neural models: Alignment links for English-French	137
B.2	Results of our neural models: Alignment links for English-German	137
B.3	Results of our neural models: Alignment links for English-Czech	138
B.4	Results of our neural models: Alignment links for English-Romanian	138
B.5	Results of our neural models: Alignment links for English-Japanese	138
B.6	Results of our neural models: Alignment links for English-Vietnamese	138
B.7	Results of our neural models: Unaligned words for English-French	139
B.8	Results of our neural models: Unaligned words for English-German	140
B.9	Results of our neural models: Unaligned words for English-Czech	141
B.10	Results of our neural models: Unaligned words for English-Romanian	142
B.11	Results of our neural models: Unaligned words for English-Japanese	143
B.12	Results of our neural models: Unaligned words for English-Vietnamese	144
B.13	Results of our neural models: Alignment types for English-French	145

B.14	Results of our neural models: Alignment types for English-German	145
B.15	Results of our neural models: Alignment types for English-Czech	145
B.16	Results of our neural models: Alignment types for English-Romanian	146
B.17	Results of our neural models: Alignment types for English-Japanese	146
B.18	Results of our neural models: Alignment types for English-Vietnamese	146
B.19	Neural models: Jump widths for English words in the direction French-English and jump widths of foreign words in direction English-French	147
B.20	Neural models: Jump widths for English words in the direction German-English and jump widths of foreign words in direction English-German	148
B.21	Neural models: Jump widths for English words in the direction Czech-English and jump widths of foreign words in direction English-Czech	149
B.22	Neural models: Jump widths for English words in the direction Romanian- English and jump widths of foreign words in direction English-Romanian	150
B.23	Neural models: Jump widths for English words in the direction Japanese-English and jump widths of foreign words in direction English-Japanese	151
B.24	Neural models: Jump widths for English words in the direction Vietnamese- English and jump widths of foreign words in direction English-Vietnamese	152
B.25	Neural models for the direction English-French: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	155
B.26	Neural models for the direction French-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	158
B.27	Neural models for the direction English-German: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	161
B.28	Neural models for the direction German-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	163
B.29	Neural models for the direction English-Czech: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	166
B.30	Neural models for the direction Czech-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	169
B.31	Neural models for the direction English-Romanian: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	172
B.32	Neural models for the direction Romanian-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	174
B.33	Neural models for the direction English-Japanese: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	177
B.34	Neural models for the direction Japanese-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	180
B.35	Neural models for the direction English-Vietnamese: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	183
B.36	Neural models for the direction Vietnamese-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.	185

B.37	Neural models for English-French The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).	196
B.38	Neural models for English-German The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).	197
B.39	Neural models for English-Czech The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).	198
B.40	Neural models for English-Romanian The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).	199
B.41	Neural models for English-Japanese: The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).	200
B.42	Neural models for English-Vietnamese: The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).	201
B.43	Neural models for English-French: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).	202
B.44	Neural models for English-German: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).	203
B.45	Neural models for English-Czech: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).	204
B.46	Neural models for English-Romanian: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).	205
B.47	Neural models for English-Japanese: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).	206
B.48	Neural models for English-Vietnamese: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).	207
B.49	Neural models for the direction English-French: AER score as a function of sentence length	216
B.50	Neural models for the direction French-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	218
B.51	Neural models for the direction English-German: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	221
B.52	Neural models for the direction German-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	223
B.53	Neural models for the direction English-Czech: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	225
B.54	Neural models for the direction Czech-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	228

B.55	Neural models for the direction English-Romanian: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	230
B.56	Neural models for the direction Romanian-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	232
B.57	Neural models for the direction English-Japanese: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	235
B.58	Neural models for the direction Japanese-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	237
B.59	Neural models for the direction English-Vietnamese: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	239
B.60	Neural models for the direction Vietnamese-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	242
B.61	Neural models for the direction English-French: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences. . .	244
B.62	Neural models for the direction French-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences. . .	246
B.63	Neural models for the direction English-German: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences. . .	249
B.64	Neural models for the direction German-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences. . .	251
B.65	Neural models for the direction English-Czech: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences. . .	253
B.66	Neural models for the direction Czech-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences. . .	256
B.67	Neural models for the direction English-Romanian: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences. . .	258
B.68	Neural models for the direction Romanian-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences. . .	260
B.69	Neural models for the direction English-Japanese: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences. . .	263
B.70	Neural models for the direction Japanese-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences. . .	265
B.71	Neural models for the direction English-Vietnamese: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences. . .	267

B.72	Neural models for the direction Vietnamese-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	270
C.1	Results of our attention-based models: Alignment links for English-French	273
C.2	Results of our attention-based models: Alignment links for English-German	273
C.3	Results of our attention-based models: Alignment links for English-Czech	273
C.4	Results of our attention-based models: Alignment links for English-Romanian	273
C.5	Results of our attention-based models: Unaligned words for English-French	274
C.6	Results of our attention-based models: Unaligned words for English-German	275
C.7	Results of our attention-based models: Unaligned words for English-Czech	276
C.8	Results of our attention-based models: Unaligned words for English-Romanian	277
C.9	Results of our attention-based models: Alignment types for English-French	278
C.10	Results of our attention-based models: Alignment types for English-German	278
C.11	Results of our attention-based models: Alignment types for English-Czech	278
C.12	Results of our attention-based models: Alignment types for English-Romanian	279
C.13	Attention-based models for the direction English-French: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	279
C.14	Attention-based models for the direction French-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	280
C.15	Attention-based models for the direction English-German: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	281
C.16	Attention-based models for the direction German-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	282
C.17	Attention-based models for the direction English-Czech: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	282
C.18	Attention-based models for the direction Czech-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	283
C.19	Attention-based models for the direction English-Romanian: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	284
C.20	Attention-based models for the direction Romanian-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	285
D.1	Results of our variational models: Alignment links for English-French	293
D.2	Results of our variational models: Alignment links for English-German	293
D.3	Results of our variational models: Alignment links for English-Czech	293
D.4	Results of our variational models: Alignment links for English-Romanian	294
D.5	Results of our variational models: Alignment links for English-Japanese	294
D.6	Results of our variational models: Alignment links for English-Vietnamese	294
D.7	Results of our variational models: Unaligned words for English-French	295
D.8	Results of our variational models: Unaligned words for English-German	296
D.9	Results of our variational models: Unaligned words for English-Czech	296
D.10	Results of our variational models: Unaligned words for English-Romanian	297
D.11	Results of our variational models: Unaligned words for English-Japanese	297
D.12	Results of our variational models: Unaligned words for English-Vietnamese	298

D.13 Results of our variational models: Alignment types of English-French	299
D.14 Results of our variational models: Alignment types of English-German	299
D.15 Results of our variational models: Alignment types of English-Czech	299
D.16 Results of our variational models: Alignment types of English-Romanian	300
D.17 Results of our variational models: Alignment types of English-Japanese	300
D.18 Results of our variational models: Alignment types of English-Vietnamese	300
D.19 Results of our variational models: Alignment types for English-French	301
D.20 Results of our variational models: Alignment types for English-German	301
D.21 Results of our variational models: Alignment types for English-Czech	301
D.22 Results of our variational models: Alignment types for English-Romanian	302
D.23 Results of our variational models: Alignment types for English-Japanese	302
D.24 Results of our variational models: Alignment types for English-Vietnamese	302
D.25 Variational models: Jump widths for English words in the direction French- English and jump widths for foreign words in the direction English-French	303
D.26 Variational models: Jump widths for English words in the direction German- English and jump widths for foreign words in the direction English-German	304
D.27 Variational models: Jump widths for English words in the direction Czech- English and jump widths for foreign words in the direction English-Czech	305
D.28 Variational models: Jump widths for English words in the direction Romanian- English and jump widths for foreign words in the direction English-Romanian	306
D.29 Variational models: Jump widths for English words in the direction Japanese- English and jump widths for foreign words in the direction English-Japanese	307
D.30 Variational models: Jump widths for English words in the direction Vietnamese- English and jump widths for foreign words in the direction English-Vietnamese	308
D.31 Variational models for the direction English-French: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the right graph: The darker the cell, the greater the number of confusions.	310
D.32 Variational models for the direction French-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the right graph: The darker the cell, the greater the number of confusions.	312
D.33 Variational models for the direction English-German: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the right graph: The darker the cell, the greater the number of confusions.	314
D.34 Variational models for the direction German-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the right graph: The darker the cell, the greater the number of confusions.	316
D.35 Variational models for the direction English-Czech: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the right graph: The darker the cell, the greater the number of confusions.	318
D.36 Variational models for the direction Czech-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the right graph: The darker the cell, the greater the number of confusions.	320
D.37 Variational models for the direction English-Romanian: Correct (TP) and incor- rect (FP) jump widths for source words on the left graph. The confusion matrix is on the right graph: The darker the cell, the greater the number of confusions.	322
D.38 Variational models for the direction Romanian-English: Correct (TP) and incor- rect (FP) jump widths for source words on the left graph. The confusion matrix is on the right graph: The darker the cell, the greater the number of confusions.	324
D.39 Variational models for the direction English-Japanese: Correct (TP) and incor- rect (FP) jump widths for source words on the left graph. The confusion matrix is on the right graph: The darker the cell, the greater the number of confusions.	326

D.40	Variational models for the direction Japanese-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the right graph: The darker the cell, the greater the number of confusions.	328
D.41	Variational models for the direction English-Vietnamese: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the right graph: The darker the cell, the greater the number of confusions.	330
D.42	Variational models for the direction Vietnamese-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the right graph: The darker the cell, the greater the number of confusions.	332
D.43	Models for English-French: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).	340
D.44	Models for English-German: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).	341
D.45	Models for English-Czech: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).	342
D.46	Models for English-Romanian: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).	343
D.47	Models for English-Japanese: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).	344
D.48	Models for English-Vietnamese: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).	345
D.49	Models for English-French: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).	346
D.50	Models for English-German: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).	347
D.51	Models for English-Czech: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).	348
D.52	Models for English-Romanian: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).	349
D.53	Models for English-Japanese: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).	350
D.54	Models for English-Vietnamese: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).	351
D.55	Variational models for the direction English-French: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	362
D.56	Variational models for the direction French-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	364

D.57 Variational models for the direction English-German: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	366
D.58 Variational models for the direction German-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	368
D.59 Variational models for the direction English-Czech: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	370
D.60 Variational models for the direction Czech-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	372
D.61 Variational models for the direction English-Romanian: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	374
D.62 Variational models for the direction Romanian-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	376
D.63 Variational models for the direction English-Japanese: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	377
D.64 Variational models for the direction Japanese-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	379
D.65 Variational models for the direction English-Vietnamese: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	381
D.66 Variational models for the direction Vietnamese-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.	383
D.67 Models for the direction English-French: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	385
D.68 Models for the direction French-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	386
D.69 Models for the direction English-German: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	388
D.70 Models for the direction German-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	390
D.71 Models for the direction English-Czech: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	392
D.72 Models for the direction Czech-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	393
D.73 Models for the direction English-Romanian: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	395

D.74	Models for the direction Romanian-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	397
D.75	Models for the direction English-Japanese: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	399
D.76	Models for the direction Japanese-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	400
D.77	Models for the direction English-Vietnamese: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	402
D.78	Models for the direction Vietnamese-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.	404
E.1	BPE-based <code>Fastalign</code> for English-French: For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).	406
E.2	BPE-based <code>Fastalign</code> for English-German: For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).	406
E.3	BPE-based <code>Fastalign</code> for English-Czech: For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).	407
E.4	BPE-based <code>Fastalign</code> for English-Romanian: For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).	407
E.5	BPE-based <code>Fastalign</code> for English-Japanese: For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).	408
E.6	BPE-based <code>Fastalign</code> for English-Vietnamese: For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).	408
E.7	BPE-based <code>Fastalign</code> for the direction English-French: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	410
E.8	BPE-based <code>Fastalign</code> for the direction French-English: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	411
E.9	BPE-based <code>Fastalign</code> for the direction English-German: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	412
E.10	BPE-based <code>Fastalign</code> for the direction German-English: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	413
E.11	BPE-based <code>Fastalign</code> for the direction English-Czech: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	414
E.12	BPE-based <code>Fastalign</code> for the direction Czech-English: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	415

E.13 BPE-based <code>Fastalign</code> for the direction English-Romanian: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	416
E.14 BPE-based <code>Fastalign</code> for the direction Romanian-English: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	418
E.15 BPE-based <code>Fastalign</code> for the direction English-Japanese: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	419
E.16 BPE-based <code>Fastalign</code> for the direction Japanese-English: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	420
E.17 BPE-based <code>Fastalign</code> for the direction English-Vietnamese: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	421
E.18 BPE-based <code>Fastalign</code> for the direction Vietnamese-English: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	421
E.19 BPE-based <code>Fastalign</code> with BPE-dropout for English-French: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).	422
E.20 BPE-based <code>Fastalign</code> with BPE-dropout for English-German: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).	423
E.21 BPE-based <code>Fastalign</code> with BPE-dropout for English-Czech: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).	423
E.22 BPE-based <code>Fastalign</code> with BPE-dropout for English-Romanian: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).	424
E.23 BPE-based <code>Fastalign</code> with BPE-dropout for English-Japanese: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).	424
E.24 BPE-based <code>Fastalign</code> with BPE-dropout for English-Vietnamese: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).	425
E.25 BPE-based <code>Eflomal</code> for English-French: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).	426
E.26 BPE-based <code>Eflomal</code> for English-German: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).	427
E.27 BPE-based <code>Eflomal</code> for English-Czech: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).	427
E.28 BPE-based <code>Eflomal</code> for English-Romanian: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).	428
E.29 BPE-based <code>Eflomal</code> for English-Japanese: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).	428

E.30 BPE-based <code>Eflomal</code> for English-Vietnamese: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).	429
E.31 BPE-based <code>Fastalign</code> for English-French: We observe correct and incorrect alignment links.	433
E.32 BPE-based <code>Fastalign</code> for English-German: We observe correct and incorrect alignment links.	433
E.33 BPE-based <code>Fastalign</code> for English-Czech: We observe correct and incorrect alignment links.	434
E.34 BPE-based <code>Fastalign</code> for English-Romanian: We observe correct and incorrect alignment links.	434
E.35 BPE-based <code>Fastalign</code> for English-Japanese: We observe correct and incorrect alignment links.	434
E.36 BPE-based <code>Fastalign</code> for English-Vietnamese: We observe correct and incorrect alignment links.	435
E.37 BPE-based <code>Fastalign</code> : Unaligned words for English-French	436
E.38 BPE-based <code>Fastalign</code> : Unaligned words for English-German	437
E.39 BPE-based <code>Fastalign</code> : Unaligned words for English-Czech	438
E.40 BPE-based <code>Fastalign</code> : Unaligned words for English-Romanian	439
E.41 BPE-based <code>Fastalign</code> : Unaligned words for English-Japanese	440
E.42 BPE-based <code>Fastalign</code> : Unaligned words for English-Vietnamese	441
E.43 BPE-based <code>Fastalign</code> for the direction English-French: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).	443
E.44 BPE-based <code>Fastalign</code> for the direction French-English: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).	444
E.45 BPE-based <code>Fastalign</code> for the direction English-German: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).	445
E.46 BPE-based <code>Fastalign</code> for the direction German-English: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).	447
E.47 BPE-based <code>Fastalign</code> for the direction English-Czech: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).	448
E.48 BPE-based <code>Fastalign</code> for the direction Czech-English: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).	449
E.49 BPE-based <code>Fastalign</code> for the direction English-Romanian: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).	451
E.50 BPE-based <code>Fastalign</code> for the direction Romanian-English: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).	452

E.51 BPE-based <code>Fastalign</code> for the direction English-Japanese: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).	453
E.52 BPE-based <code>Fastalign</code> for the direction Japanese-English: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).	454
E.53 BPE-based <code>Fastalign</code> for the direction English-Vietnamese: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).	455
E.54 BPE-based <code>Fastalign</code> for the direction Vietnamese-English: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).	456
E.55 BPE-based <code>Fastalign</code> for the direction English-French: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.	457
E.56 BPE-based <code>Fastalign</code> for the direction French-English: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.	458
E.57 BPE-based <code>Fastalign</code> for the direction English-German: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.	458
E.58 BPE-based <code>Fastalign</code> for the direction German-English: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.	459
E.59 BPE-based <code>Fastalign</code> for the direction English-Czech: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.	459
E.60 BPE-based <code>Fastalign</code> for the direction Czech-English: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.	460
E.61 BPE-based <code>Fastalign</code> for the direction English-Romanian: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.	460
E.62 BPE-based <code>Fastalign</code> for the direction Romanian-English: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.	461

E.63 BPE-based <code>Fastalign</code> for the direction English-Japanese: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.	461
E.64 BPE-based <code>Fastalign</code> for the direction Japanese-English: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.	462
E.65 BPE-based <code>Fastalign</code> for the direction English-Vietnamese: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.	462
E.66 BPE-based <code>Fastalign</code> for the direction Vietnamese-English: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.	463
E.67 BPE-based <code>Fastalign</code> for the direction English-French: We observe the scores for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	465
E.68 BPE-based <code>Fastalign</code> for the direction French-English: We observe the scores for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	466
E.69 BPE-based <code>Fastalign</code> for the direction English-German: We observe the scores for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	467
E.70 BPE-based <code>Fastalign</code> for the direction German-English: We observe the scores for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	469
E.71 BPE-based <code>Fastalign</code> for the direction English-Czech: We observe the scores for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	470
E.72 BPE-based <code>Fastalign</code> for the direction Czech-English: We observe the scores for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	472
E.73 BPE-based <code>Fastalign</code> for the direction English-Romanian: We observe the scores for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	473
E.74 BPE-based <code>Fastalign</code> for the direction Romanian-English: We observe the scores for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	474

E.75 BPE-based <code>Fastalign</code> for the direction English-Japanese: We observe the scores for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	476
E.76 BPE-based <code>Fastalign</code> for the direction Japanese-English: We observe the scores for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	477
E.77 BPE-based <code>Fastalign</code> for the direction English-Vietnamese: We observe the scores for rare source words. For each source vocabulary size, we report accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	478
E.78 BPE-based <code>Fastalign</code> for the direction Vietnamese-English: We observe the scores for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.	479
E.79 The direction English-French: Average number of BPE-based fragments as a function of word occurrence.	480
E.80 The direction English-German: Average number of BPE-based fragments as a function of word occurrence.	481
E.81 The direction English-Czech: Average number of BPE-based fragments as a function of word occurrence.	482
E.82 The direction English-Romanian: Average number of BPE-based fragments as a function of word occurrence.	483
E.83 The direction English-Japanese: Average number of BPE-based fragments as a function of word occurrence.	484
E.84 The direction English-Vietnamese: Average number of BPE-based fragments as a function of word occurrence.	485
E.85 The direction English-French: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.	491
E.86 The direction French-English: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.	496
E.87 The direction English-German: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.	502
E.88 The direction German-English: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.	508
E.89 The direction English-Czech: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.	513
E.90 The direction Czech-English: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.	519
E.91 The direction English-Romanian: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.	524
E.92 The direction Romanian-English: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.	530
E.93 The direction English-Japanese: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.	535
E.94 The direction Japanese-English: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.	539
E.95 The direction English-Vietnamese: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.	542
E.96 The direction Vietnamese-English: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.	545

E.97 BPE-based Fastalign for English-French: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).	546
E.98 BPE-based Fastalign for English-German: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).	547
E.99 BPE-based Fastalign for English-Czech: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).	548
E.100 BPE-based Fastalign for English-Romanian: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).	549
E.101 BPE-based Fastalign for English-Japanese: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).	550
E.102 BPE-based Fastalign for English-Vietnamese: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).	551
E.103 BPE-based Fastalign for English-French: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).	552
E.104 BPE-based Fastalign for English-German: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).	553
E.105 BPE-based Fastalign for English-Czech: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).	554
E.106 BPE-based Fastalign for English-Romanian: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).	555
E.107 BPE-based Fastalign for English-Japanese: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).	556
E.108 BPE-based Fastalign for English-Vietnamese: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).	557
E.109 GDF-before: BPE-based Fastalign for English-French. For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) of symmetrical alignment. Note that we first apply GDF to asymmetrical BPE-based alignments of both directions for obtaining a symmetrical alignment, and then transform it into the word-based alignment.	558
E.110 GDF-before: BPE-based Fastalign for English-German. For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) of symmetrical alignment. Note that we first apply GDF to asymmetrical BPE-based alignments of both directions for obtaining a symmetrical alignment, and then transform it into the word-based alignment.	558

- E.111 GDF-before: BPE-based **Fastalign** for English-Czech. For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) of symmetrical alignment. Note that we first apply GDF to asymmetrical BPE-based alignments of both directions for obtaining a symmetrical alignment, and then transform it into the word-based alignment. 559
- E.112 GDF-before: BPE-based **Fastalign** for English-Romanian. For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) of symmetrical alignment. Note that we first apply GDF to asymmetrical BPE-based alignments of both directions for obtaining a symmetrical alignment, and then transform it into the word-based alignment. 559
- E.113 GDF-before: BPE-based **Fastalign** for English-Japanese. For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) of symmetrical alignment. Note that we first apply GDF to asymmetrical BPE-based alignments of both directions for obtaining a symmetrical alignment, and then transform it into the word-based alignment. 559
- E.114 GDF-before: BPE-based **Fastalign** for English-Vietnamese. For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) of symmetrical alignment. Note that we first apply GDF to asymmetrical BPE-based alignments of both directions for obtaining a symmetrical alignment, and then transform it into the word-based alignment. 560
- E.115 GDF-after: BPE-based **Fastalign** for English-French. we display Alignment error rate (AER) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments. 561
- E.116 GDF-after: BPE-based **Fastalign** for English-French. We show F-score (F1) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments. 562
- E.117 GDF-after: BPE-based **Fastalign** for English-French. We show precision (PRE) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments. 563
- E.118 GDF-after: BPE-based **Fastalign** for English-French. We show recall (REC) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments. 564
- E.119 GDF-after: BPE-based **Fastalign** for English-German. we display Alignment error rate (AER) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments. 565
- E.120 GDF-after: BPE-based **Fastalign** for English-German. We show F-score (F1) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments. 566
- E.121 GDF-after: BPE-based **Fastalign** for English-German. We show precision (PRE) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments. 567

E.122	GDF-after: BPE-based <code>Fastalign</code> for English-German. We show recall (REC) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.	568
E.123	GDF-after: BPE-based <code>Fastalign</code> for English-Czech. we display Alignment error rate (AER) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.	569
E.124	GDF-after: BPE-based <code>Fastalign</code> for English-Czech. We show F-score (F1) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.	570
E.125	GDF-after: BPE-based <code>Fastalign</code> for English-Czech. We show precision (PRE) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.	571
E.126	GDF-after: BPE-based <code>Fastalign</code> for English-Czech. We show recall (REC) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.	572
E.127	GDF-after: BPE-based <code>Fastalign</code> for English-Romanian. we display Alignment error rate (AER) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.	573
E.128	GDF-after: BPE-based <code>Fastalign</code> for English-Romanian. We show F-score (F1) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.	574
E.129	GDF-after: BPE-based <code>Fastalign</code> for English-Romanian. We show precision (PRE) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.	575
E.130	GDF-after: BPE-based <code>Fastalign</code> for English-Romanian. We show recall (REC) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.	576
E.131	GDF-after: BPE-based <code>Fastalign</code> for English-Japanese. we display Alignment error rate (AER) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.	577
E.132	GDF-after: BPE-based <code>Fastalign</code> for English-Japanese. We show F-score (F1) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.	578
E.133	GDF-after: BPE-based <code>Fastalign</code> for English-Japanese. We show precision (PRE) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.	579
E.134	GDF-after: BPE-based <code>Fastalign</code> for English-Japanese. We show recall (REC) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.	580

E.135	GDF-after: BPE-based Fastalign for English-Vietnamese. we display Alignment error rate (AER) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.	581
E.136	GDF-after: BPE-based Fastalign for English-Vietnamese. We show F-score (F1) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.	582
E.137	GDF-after: BPE-based Fastalign for English-Vietnamese. We show precision (PRE) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.	583
E.138	GDF-after: BPE-based Fastalign for English-Vietnamese. We show recall (REC) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.	584
E.139	BPE-based Fastalign for the direction English-French: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.	591
E.140	BPE-based Fastalign for the direction French-English: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.	597
E.141	BPE-based Fastalign for the direction English-German: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.	603
E.142	BPE-based Fastalign for the direction German-English: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.	609
E.143	BPE-based Fastalign for the direction English-Czech: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.	615
E.144	BPE-based Fastalign for the direction Czech-English: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.	622
E.145	BPE-based Fastalign for the direction English-Romanian: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.	628
E.146	BPE-based Fastalign for the direction Romanian-English: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.	634
E.147	BPE-based Fastalign for the direction English-Japanese: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.	639

E.148	BPE-based <code>Fastalign</code> for the direction Japanese-English: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.	644
E.149	BPE-based <code>Fastalign</code> for the direction English-Vietnamese: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.	648
E.150	BPE-based <code>Fastalign</code> for the direction Vietnamese-English: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.	651
E.151	The direction English-French: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.	652
E.152	The direction French-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.	653
E.153	The direction English-German: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.	654
E.154	The direction German-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.	655
E.155	The direction English-Czech: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.	656
E.156	The direction Czech-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.	657
E.157	The direction English-Romanian: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.	658
E.158	The direction Romanian-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.	659
E.159	The direction English-Japanese: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.	660
E.160	The direction Japanese-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.	661

- E.161 The direction English-Vietnamese: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence. 662
- E.162 The direction Vietnamese-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence. 663

List of Tables

A.1	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-French	42
A.2	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-German	42
A.3	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Czech	42
A.4	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Romanian	43
A.5	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Japanese	43
A.6	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Vietnamese	43
A.7	Baselines for English-French: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction French-English and in the direction English-French	84
A.8	Baselines for English-German: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction German-English and in the direction English-German	84
A.9	Baselines for English-Czech: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Czech-English and in the direction English-Czech	84
A.10	Baselines for English-Romanian: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Romanian-English and in the direction English-Romanian	84
A.11	Baselines for English-Japanese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Japanese-English and in the direction English-Japanese	85
A.12	Baselines for English-Vietnamese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Vietnamese-English and in the direction English-Vietnamese	85
A.13	Baselines for English-French: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction French-English and in the direction English-French.	85
A.14	Baselines for English-German: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction German-English and in the direction English-German.	86
A.15	Baselines for English-Czech: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Czech-English and in the direction English-Czech.	86
A.16	Baselines for English-Romanian: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Romanian-English and in the direction English-Romanian.	86

A.17	Baselines for English-Japanese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Japanese-English and in the direction English-Japanese.	87
A.18	Baselines for English-Vietnamese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Vietnamese-English and in the direction English-Vietnamese.	87
A.19	Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-French	99
A.20	Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-German	99
A.21	Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Czech	99
A.22	Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Romanian	100
A.23	Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Japanese	100
A.24	Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Vietnamese	100
A.25	Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-French .	100
A.26	Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-German .	101
A.27	Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Czech . .	101
A.28	Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Romanian	101
A.29	Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Japanese	101
A.30	Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Vietnamese	101
B.1	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-French	134
B.2	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-German	134
B.3	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Czech	135
B.4	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Romanian	135
B.5	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Japanese	136
B.6	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Vietnamese	136
B.19	Models for English-French: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction French-English and in the direction English-French	186

B.20 Models for English-German: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction German-English and in the direction English-German	186
B.21 Models for English-Czech: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Czech-English and in the direction English-Czech	187
B.22 Models for English-Romanian: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Romanian-English and in the direction English-Romanian	187
B.23 Models for English-Japanese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Japanese-English and in the direction English-Japanese	188
B.24 Models for English-Vietnamese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Vietnamese-English and in the direction English-Vietnamese	188
B.25 Models for English-French: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction French-English and in the direction English-French.	189
B.26 Models for English-German: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction German-English and in the direction English-German.	189
B.27 Models for English-Czech: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Czech-English and in the direction English-Czech.	190
B.28 Models for English-Romanian: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Romanian-English and in the direction English-Romanian.	190
B.29 Models for English-Japanese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Japanese-English and in the direction English-Japanese.	191
B.30 Models for English-Vietnamese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Vietnamese-English and in the direction English-Vietnamese.	191
B.31 Models for English-French: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction French-English and in the direction English-French. Note that the training data for for all models including the baselines only has a vocabulary containing the most frequent 50K words.	192
B.32 Models for English-German: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction German-English and in the direction English-German. Note that the training data for for all models including the baselines only has a vocabulary containing the most frequent 50K words.	192
B.33 Models for English-Czech: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Czech-English and in the direction English-Czech. Note that the training data for for all models including the baselines only has a vocabulary containing the most frequent 50K words.	193
B.34 Models for English-Romanian: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Romanian-English and in the direction English-Romanian. Note that the training data for for all models including the baselines only has a vocabulary containing the most frequent 50K words.	193

B.35	Models for English-Japanese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Japanese-English and in the direction English-Japanese. Note that the training data for for all models including the baselines only has a vocabulary containing the most frequent 50K words.	194
B.36	Models for English-Vietnamese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Vietnamese-English and in the direction English-Vietnamese. Note that the training data for for all models including the baselines only has a vocabulary containing the most frequent 50K words.	194
B.37	# links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction German-English and in the direction English-German. Note that the training data for for all models including the baselines only has a vocabulary containing the most frequent 50K words.	195
B.38	Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-French	208
B.39	Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-German	208
B.40	Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Czech	209
B.41	Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Romanian	209
B.42	Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Japanese	210
B.43	Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Vietnamese	210
B.44	Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-French	211
B.45	Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-German	211
B.46	Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Czech	212
B.47	Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Romanian	212
B.48	Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Japanese	213
B.49	Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Vietnamese	213
C.1	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) of English-French	271
C.2	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) of English-German	272
C.3	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) of English-Czech	272
C.4	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) of English-Romanian	272

C.13	Intersection alignment for attention-based models: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-French	285
C.14	Intersection alignment for attention-based models: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-German	286
C.15	Intersection alignment for attention-based models: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Czech	286
C.16	Intersection alignment for attention-based models: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Romanian	286
C.17	Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-French .	286
C.18	Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-German .	287
C.19	Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Czech . .	287
C.20	Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Romanian	287
D.1	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-French	290
D.2	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-German	290
D.3	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Czech	291
D.4	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Romanian	291
D.5	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Japanese	292
D.6	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Vietnamese	292
D.19	Models for English-French: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction French-English and in the direction English-French	333
D.20	Models for English-German: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction German-English and in the direction English-German	334
D.21	Models for English-Czech: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Czech-English and in the direction English-Czech	334
D.22	Models for English-Romanian: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Romanian-English and in the direction English-Romanian	335
D.23	Models for English-Japanese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Japanese-English and in the direction English-Japanese	335
D.24	Models for English-Vietnamese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Vietnamese-English and in the direction English-Vietnamese	336

D.25 Models for English-French: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction French-English and in the direction English-French.	337
D.26 Models for English-German: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction German-English and in the direction English-German.	337
D.27 Models for English-Czech: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Czech-English and in the direction English-Czech.	338
D.28 Models for English-Romanian: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Romanian-English and in the direction English-Romanian.	338
D.29 Models for English-Japanese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Japanese-English and in the direction English-Japanese.	339
D.30 Models for English-Vietnamese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Vietnamese-English and in the direction English-Vietnamese.	339
D.31 Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-French	352
D.32 Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-German	353
D.33 Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Czech	353
D.34 Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Romanian	354
D.35 Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Japanese	354
D.36 Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Vietnamese	355
D.37 Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-French .	356
D.38 Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-German .	357
D.39 Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Czech .	357
D.40 Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Romanian	358
D.41 Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Japanese	358
D.42 Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Vietnamese	359
D.43 Training with a monolingual corpus and the noise model (+Noise) on English-Czech corpus. R-Acc is the accuracy of the reconstruction model.	360
D.44 Training with a monolingual corpus and the noise model (+Noise) on English-Japanese corpus. R-Acc is the accuracy of the reconstruction model.	360

D.45	Training with a monolingual corpus and the noise model (+Noise) on English-Vietnamese corpus. R-Acc is the accuracy of the reconstruction model.	360
E.13	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-French	430
E.14	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-German	430
E.15	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Czech	431
E.16	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Romanian	431
E.17	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Japanese	432
E.18	Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Vietnamese	432
F.1	Alignment error rate (AER) and F-score (F1) for English-French. Note that the training corpus of these experiments is the dataset Hansards which is different from the training corpus used in our analysis. For our models, bi-direction results are generated by using grow-diag-final.	665
F.2	Alignment error rate (AER) and F-score (F1) for English-German. For our models, bi-direction results are generated by using grow-diag-final.	666
F.3	Alignment error rate (AER) and F-score (F1) for English-Romanian. Note that the training corpus of these experiments is the NAACL 2003 corpus [Mihalcea and Pedersen, 2003] which is different from the training corpus used in our analysis. For our models, bi-direction results are generated by using grow-diag-final.	666
F.4	Alignment error rate (AER), F-score (F1) for English-Japanese. For our models, bi-direction results are generated by using grow-diag-final.	666

Appendices

Appendix A

Appendix: Evaluation methods

Contents

A.1	Baselines: Scores of the baselines	41
A.2	Baselines: Correct and Incorrect alignment links	44
A.3	Baselines: Unaligned words	47
A.4	Alignment types	53
A.4.1	Baselines: Alignment types	53
A.4.2	Baselines: Correct/incorrect links as a function of alignment types	55
A.5	Jump width statistics	57
A.5.1	Baselines: Jump width statistics	63
A.5.2	Baselines: Correct and incorrect jumps	69
A.6	Word length as a function of word occurrence	82
A.7	Baselines: Rare words	84
A.8	Baselines: Unknown words	85
A.9	Baselines: Part-of-Speech	88
A.9.1	Baselines: Alignment links for content and function words	88
A.9.2	Baselines: Unaligned words for content and function words	94
A.10	Baselines: Symmetrization	99
A.10.1	Baselines: Intersection	99
A.10.2	Baselines: Grow-diag-final (GDF)	100
A.11	Word-based sentence length statistics	102
A.11.1	Sentence length statistics	102
A.11.2	Length difference statistics between English and foreign sentence	105
A.11.3	Sentence length and unknown/rare words	107
A.11.4	Sentence length and word repetition	109
A.11.5	Baselines: AER as a function of sentence length	111
A.11.6	Baselines: AER as a function of sentence length difference	123

A.1 Baselines: Scores of the baselines

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
Concatenation										
IBM-1 Giza++	40.1	26.7	71.55	16.41	89.01	33.9	36.49	59.24	26.37	88.81
Fastalign	15.19	44.98	82.5	30.92	90.78	16.23	46.32	80.08	32.58	90.79
HMM Giza++	11.99	45.18	86.12	30.62	90.94	11.97	45.98	85.2	31.49	90.98
IBM-4 Giza++	10	44.43	90.61	29.43	91.02	9.64	45.43	89.58	30.43	91.08
Replacing unknown words with the token UNK										
IBM-1 Giza++	30.97	36.89	64.26	25.87	89.21	33.32	36.99	60.06	26.73	88.9
Fastalign	15.33	44.91	82.41	30.86	90.77	16.41	46.21	79.93	32.5	90.77
HMM Giza++	10.83	45.82	87.69	31.01	91.06	11	46.66	86.53	31.94	91.09
IBM-4 Giza++	15.02	41.41	88.94	26.99	90.69	12.44	43.4	88.81	28.71	90.87

Table A.1: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-French

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
Concatenation										
IBM-1 Giza++	39.03	58.76	59.1	58.43	96.4	42.66	55.39	57.02	53.84	96.2
Fastalign	28.98	68.75	71.11	66.54	97.35	31.28	66.47	70.73	62.69	97.23
HMM Giza++	23.92	73.3	79.23	68.2	97.82	26.33	71.04	79.47	64.23	97.7
IBM-4 Giza++	21.46	75.48	85.79	67.39	98.08	23.31	73.63	86.56	64.06	97.99
Replacing unknown words with the token UNK										
IBM-1 Giza++	47.88	49.65	64.7	40.27	96.42	42.49	55.55	57.19	54	96.21
Fastalign	29.05	68.68	71.04	66.47	97.34	31.39	66.37	70.69	62.55	97.22
HMM Giza++	23.26	73.95	80.17	68.63	97.88	26.32	71.05	79.33	64.34	97.7
IBM-4 Giza++	25.9	71.13	83.92	61.73	97.8	27.99	69.09	83.82	58.76	97.69

Table A.2: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-German

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
Concatenation										
IBM-1 Giza++	45.09	46.75	50.4	43.59	95.97	48.47	42.88	49.17	38.02	95.89
Fastalign	25.75	64.09	70.98	58.42	97.34	25.3	62.86	73.13	55.13	97.36
HMM Giza++	27.86	61.22	70.81	53.92	97.23	30.38	57.28	69.26	48.83	97.04
IBM-4 Giza++	20.92	65.7	79.48	56	97.63	26.5	59.81	75.58	49.48	97.3
Replacing unknown words with the token UNK										
IBM-1 Giza++	45.51	46.42	50.05	43.28	95.94	46.08	44.87	51.45	39.79	96.03
Fastalign	26.56	63.48	70.2	57.93	97.29	26.18	62.14	72.29	54.48	97.3
HMM Giza++	27.86	61.23	70.96	53.86	97.23	32.21	56.02	67.32	47.97	96.94
IBM-4 Giza++	28.56	58.94	72.79	49.51	97.2	32.48	54.86	69.59	45.28	96.97

Table A.3: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Czech

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
Concatenation										
IBM-1 Giza++	56.02	43.99	58.8	35.14	96.66	53.52	46.49	49.92	43.5	96.26
Fastalign	33.36	66.65	72.77	61.49	97.7	32.91	67.1	73.7	61.59	97.75
HMM Giza++	33.36	66.65	75.28	59.8	97.77	36.38	63.64	72.9	56.46	97.59
IBM-4 Giza++	31.04	68.98	79.28	61.04	97.95	32.3	67.72	80.97	58.2	97.93
Replacing unknown words with the token UNK										
IBM-1 Giza++	55.79	44.22	59.11	35.32	96.68	59.4	40.61	55.72	31.95	96.51
Fastalign	33.5	66.52	72.65	61.34	97.7	33.16	66.85	73.46	61.34	97.73
HMM Giza++	33.2	66.82	75.47	59.95	97.78	35.51	64.51	74.05	57.15	97.65
IBM-4 Giza++	35.06	64.96	77.59	55.86	97.75	36.76	63.26	76.76	53.79	97.67

Table A.4: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Romanian

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
Concatenation										
IBM-1 Giza++	63.12	36.88	40.24	34.04	97.22	61.55	38.45	40.13	36.9	97.18
Fastalign	50.67	49.33	55.79	44.21	97.83	49.37	50.63	54.59	47.2	97.8
HMM Giza++	57.01	42.99	52.22	36.53	97.69	54.41	45.59	52.45	40.32	97.7
IBM-4 Giza++	52.52	47.48	61.27	38.76	97.95	49.23	50.77	61.46	43.24	98
Replacing unknown words with the token UNK										
IBM-1 Giza++	63.63	36.37	39.67	33.57	97.2	67.38	32.62	40.81	27.17	97.32
Fastalign	51.55	48.45	54.86	43.38	97.8	50.35	49.65	53.46	46.35	97.76
HMM Giza++	58.29	41.71	50.66	35.45	97.64	54.99	45.01	51.78	39.81	97.68
IBM-4 Giza++	54.63	45.37	59.74	36.57	97.9	52.48	47.52	59.01	39.77	97.9

Table A.5: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Japanese

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
Concatenation										
IBM-1 Giza++	69.43	30.58	43.42	23.59	96.51	46.45	53.56	50.79	56.64	96.8
Fastalign	48.89	51.11	56.87	46.42	97.11	32.82	67.19	64.5	70.11	97.77
HMM Giza++	57.86	42.14	49.18	36.86	96.7	37.57	62.43	60.28	64.74	97.46
IBM-4 Giza++	51.91	48.1	58.38	40.89	97.12	33.19	66.82	67.17	66.47	97.85
Replacing unknown words with the token UNK										
IBM-1 Giza++	68.13	31.87	34.17	29.85	95.13	52.35	47.65	45.54	49.97	95.81
Fastalign	51.4	48.6	53.64	44.42	96.76	35.28	64.72	62.11	67.55	97.46
HMM Giza++	58.84	41.16	47.01	36.61	96.13	41.82	58.18	56.79	59.64	96.83
IBM-4 Giza++	55.19	44.81	54.08	38.25	96.61	39.32	60.68	61.88	59.53	97.23

Table A.6: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Vietnamese

A.2 Baselines: Correct and Incorrect alignment links

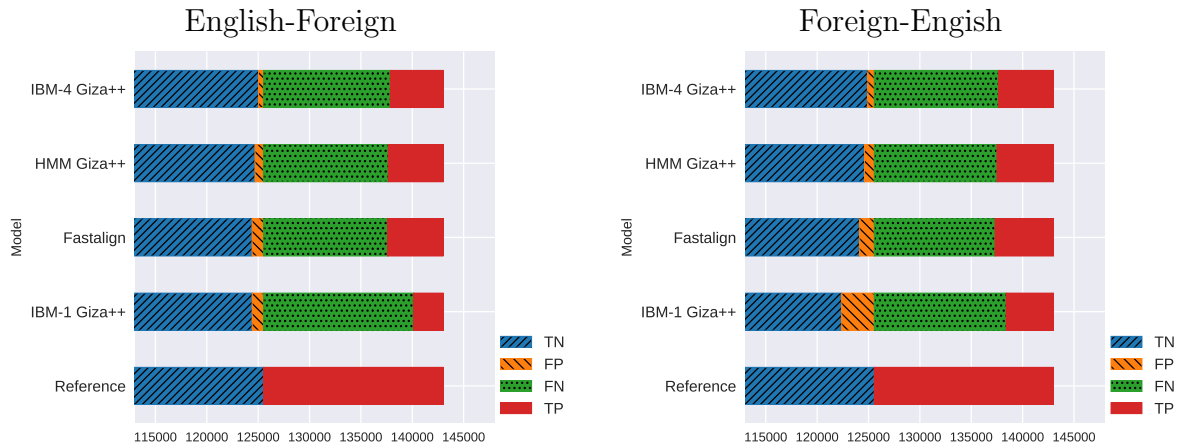


Figure A.1: Results of our baselines: Alignment links for English-French

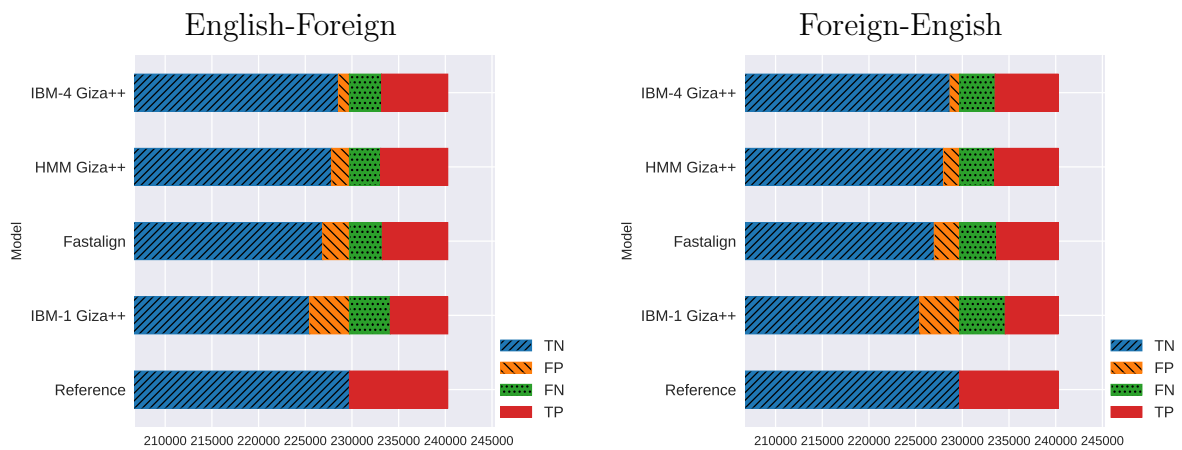


Figure A.2: Results of our baselines: Alignment links for English-German

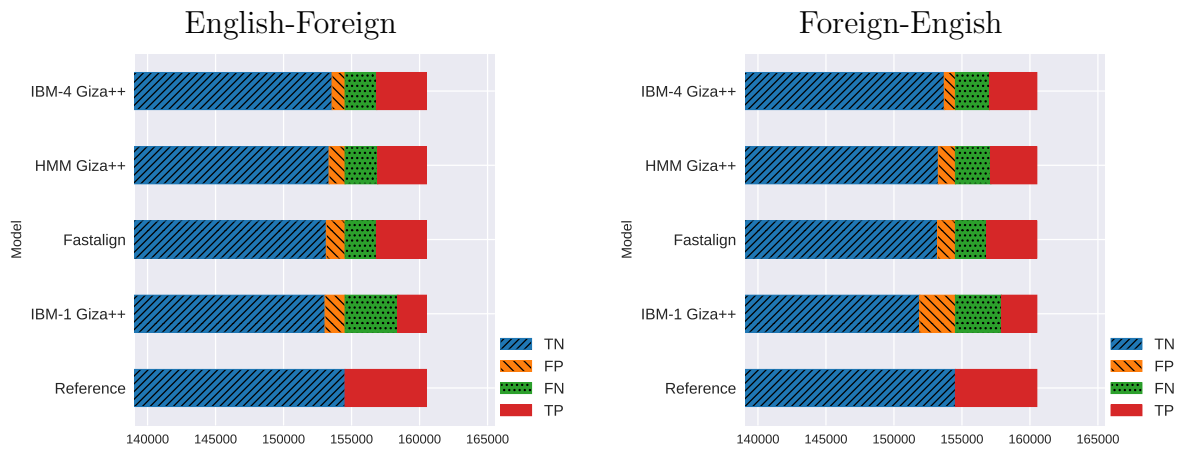


Figure A.3: Results of our baselines: Alignment links for English-Romanian

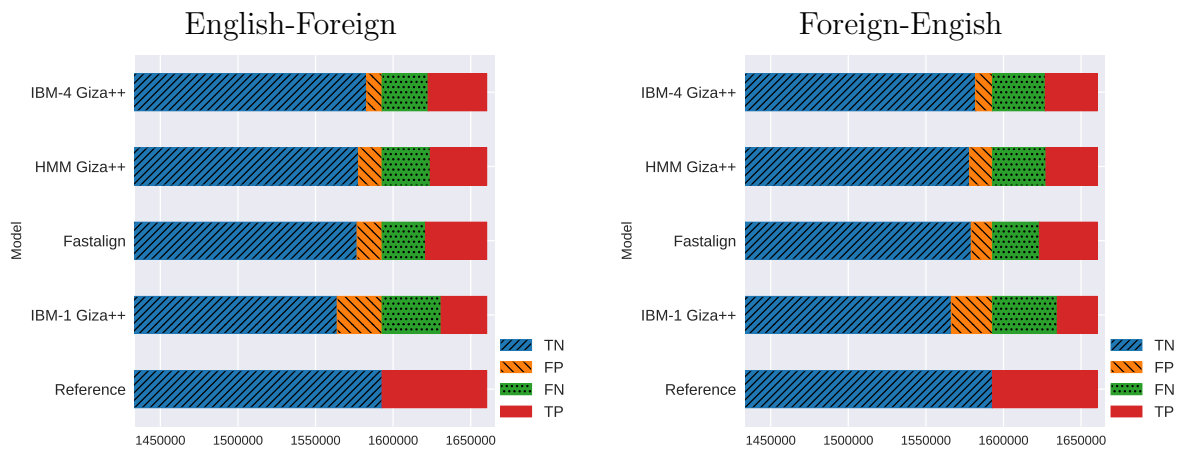


Figure A.4: Results of our baselines: Alignment links for English-Czech

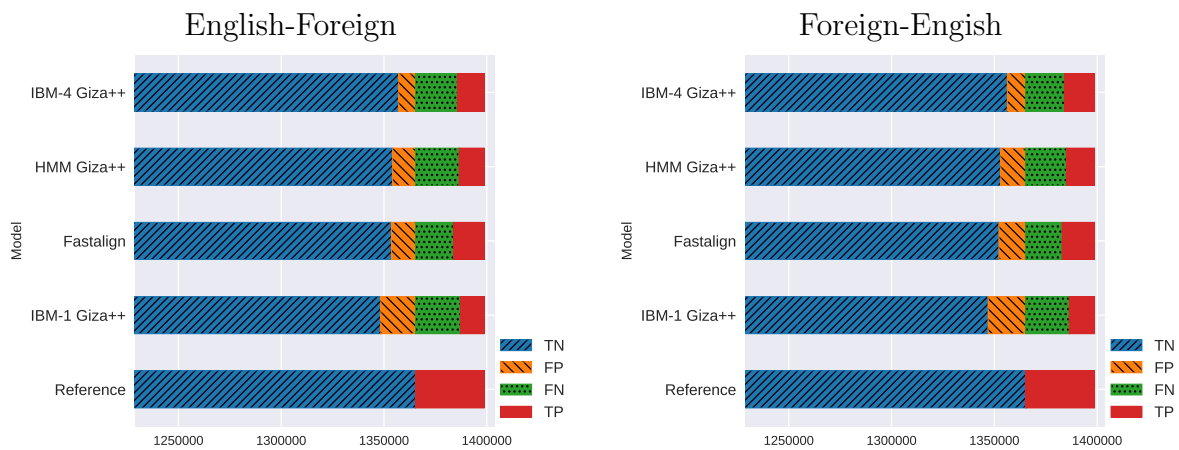


Figure A.5: Results of our baselines: Alignment links for English-Japanese

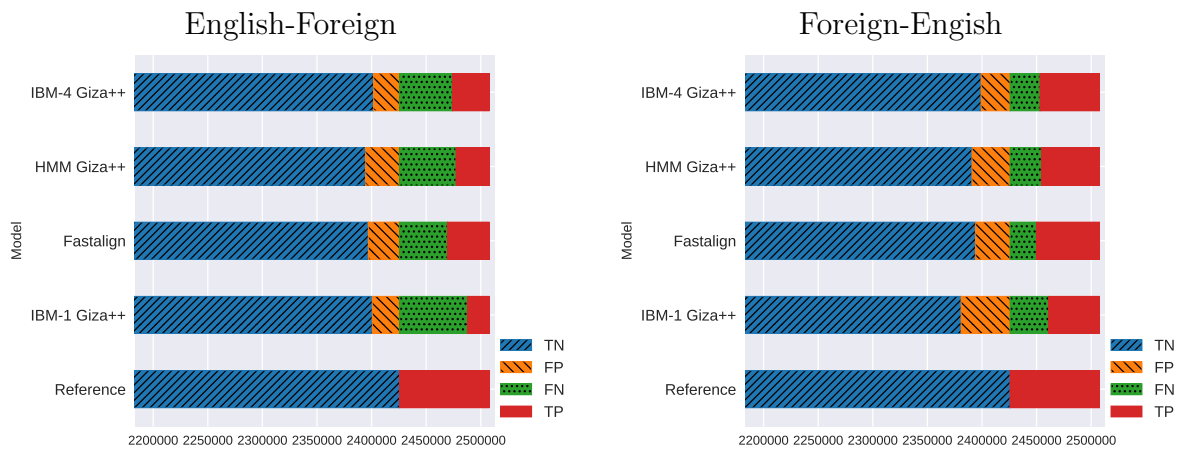


Figure A.6: Results of our baselines: Alignment links for English-Vietnamese

A.3 Baselines: Unaligned words

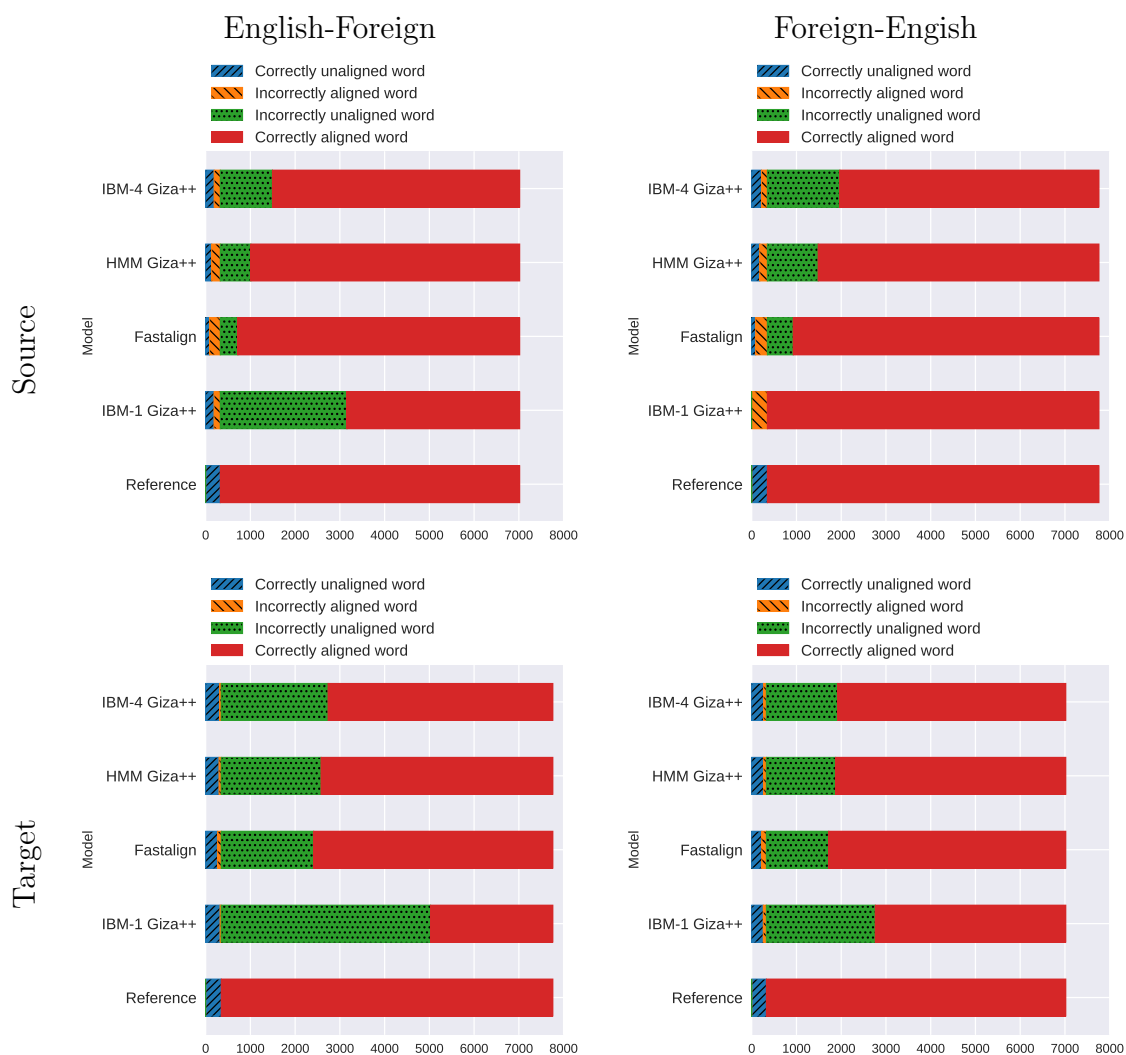


Figure A.7: Results of our baselines: Unaligned words for English-French



Figure A.8: Results of our baselines: Unaligned words for English-German

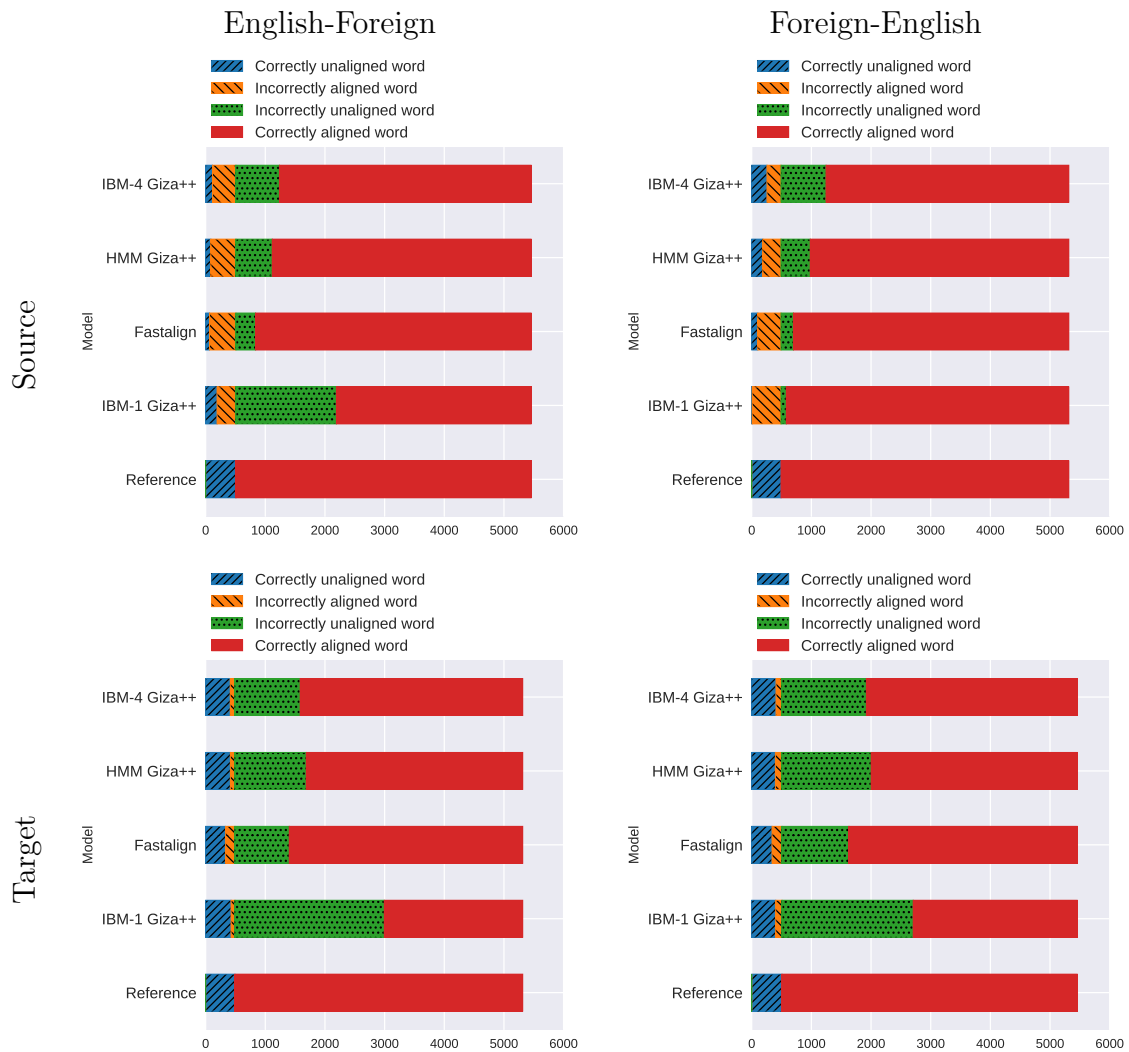


Figure A.9: Results of our baselines: Unaligned words for English-Romanian



Figure A.10: Results of our baselines: Unaligned words for English-Czech



Figure A.11: Results of our baselines: Unaligned words for English-Japanese

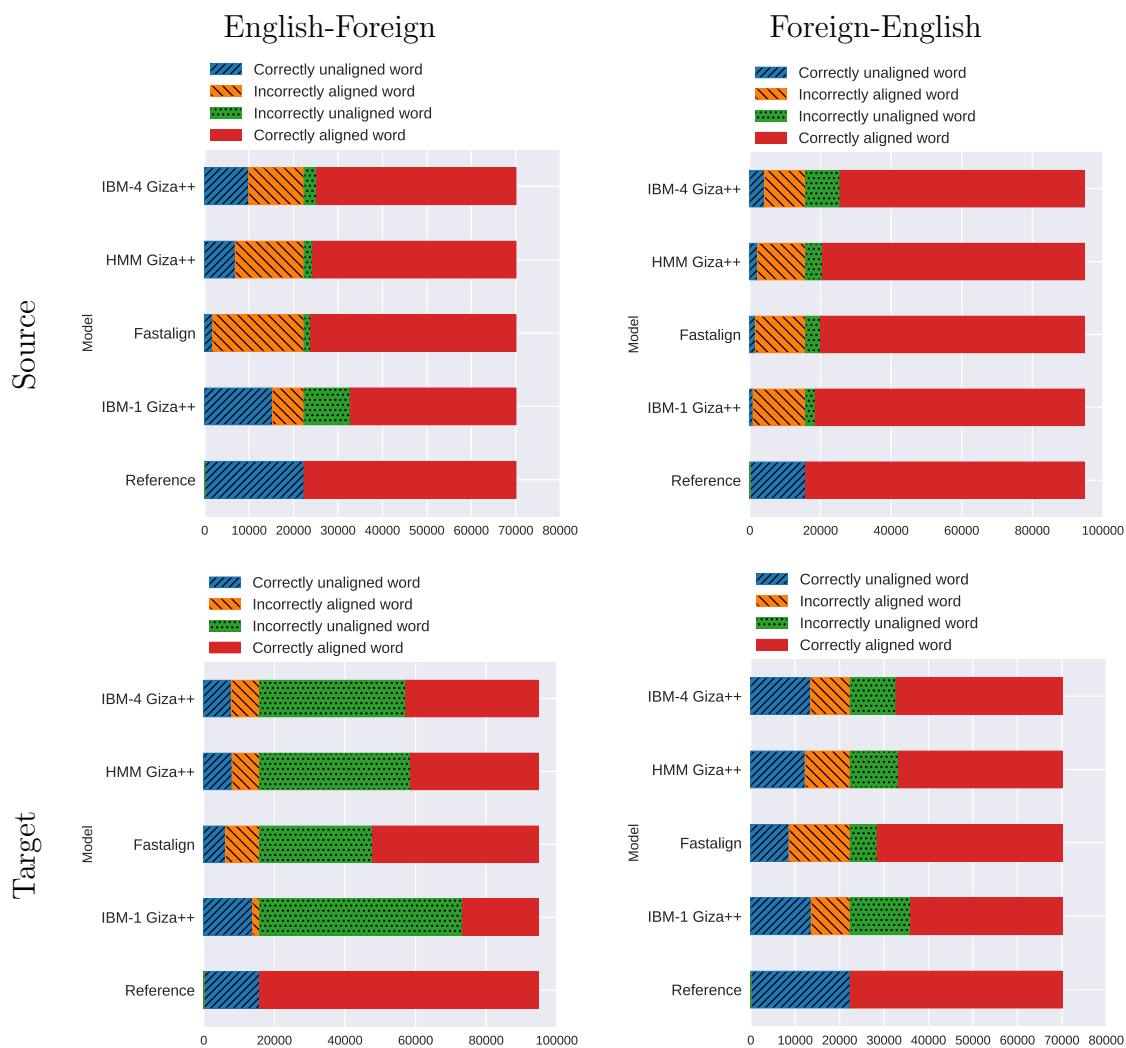


Figure A.12: Results of our baselines: Unaligned words for English-Vietnamese

A.4 Alignment types

A.4.1 Baselines: Alignment types

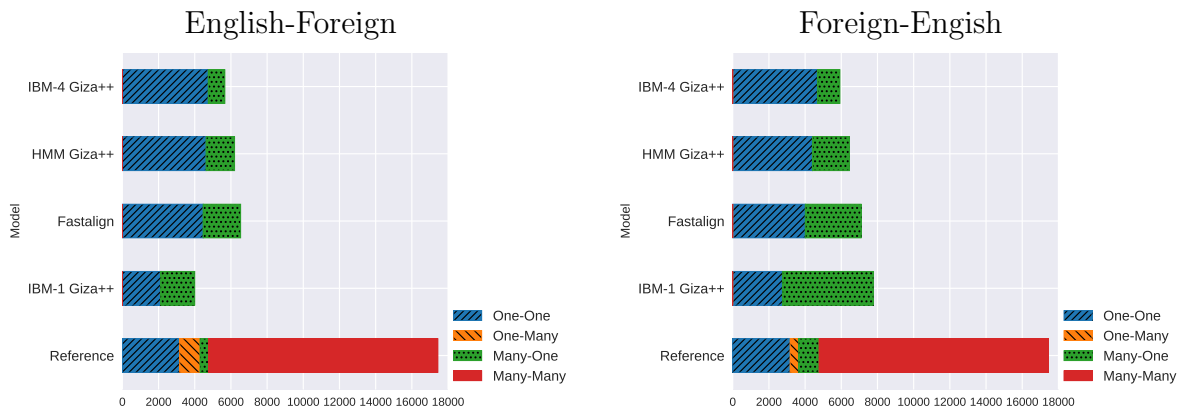


Figure A.13: Results of our baselines: Alignment types for English-French

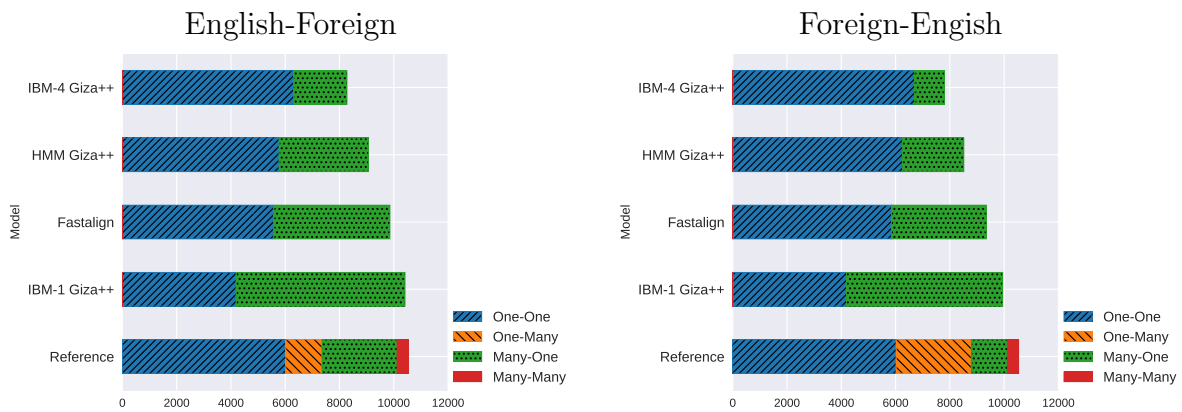


Figure A.14: Results of our baselines: Alignment types for English-German

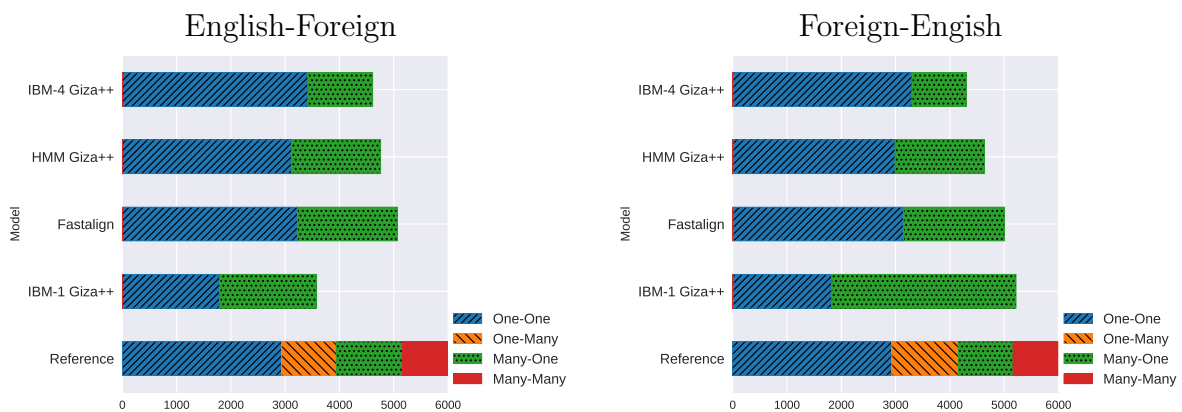


Figure A.15: Results of our baselines: Alignment types for English-Romanian

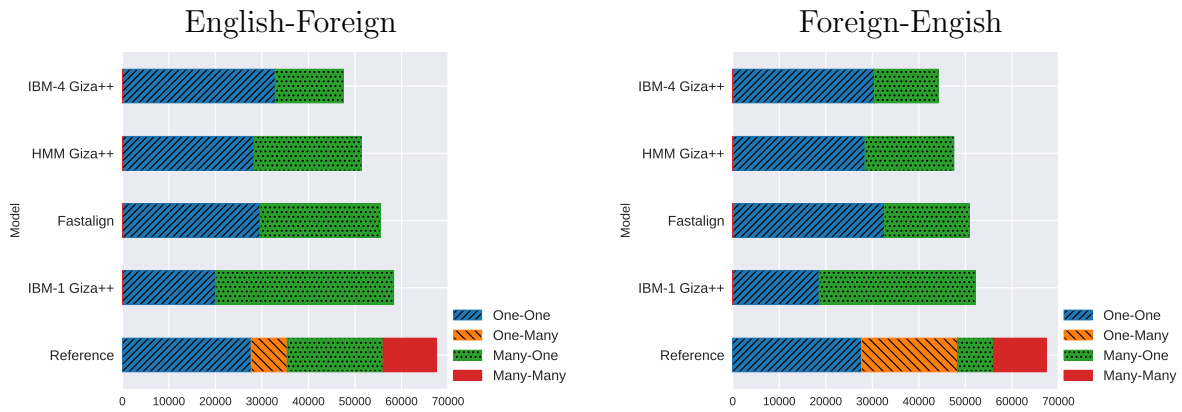


Figure A.16: Results of our baselines: Alignment types for English-Czech

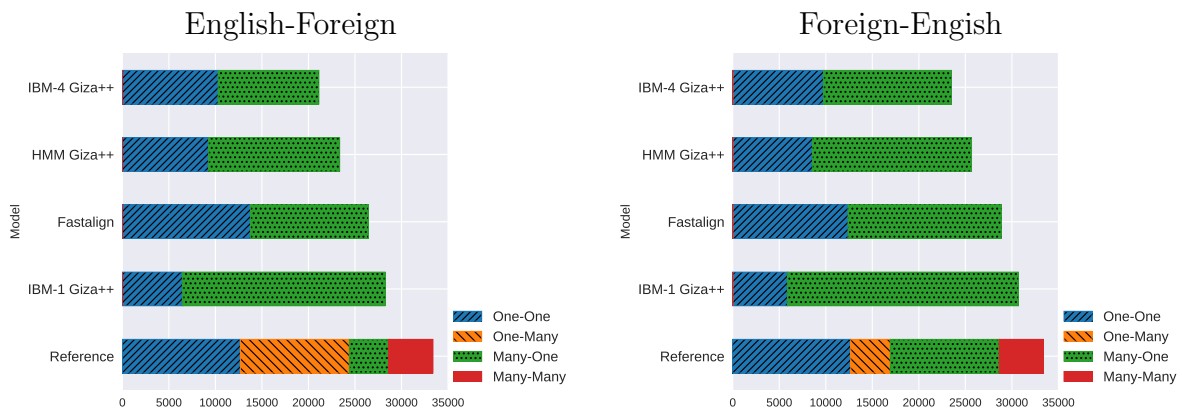


Figure A.17: Results of our baselines: Alignment types for English-Japanese

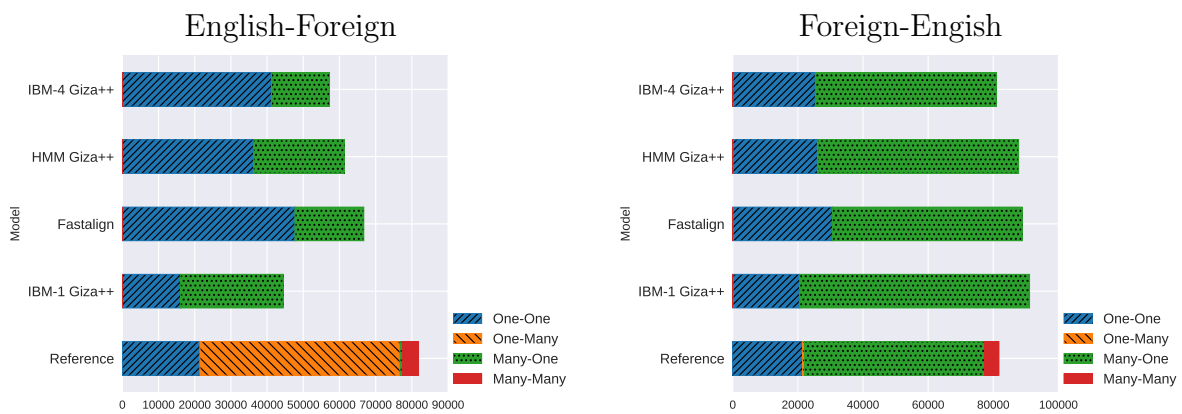


Figure A.18: Results of our baselines: Alignment types for English-Vietnamese

A.4.2 Baselines: Correct/incorrect links as a function of alignment types

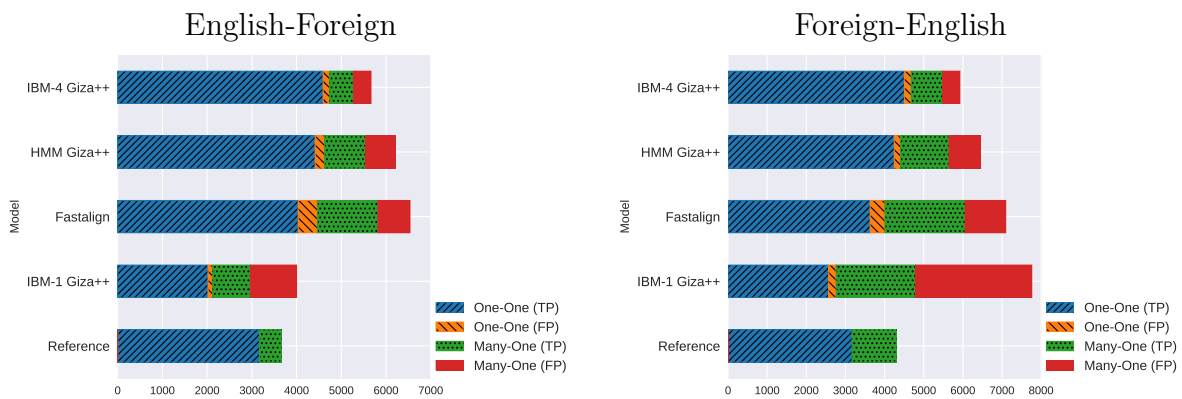


Figure A.19: Results of our baselines: Alignment types for English-French

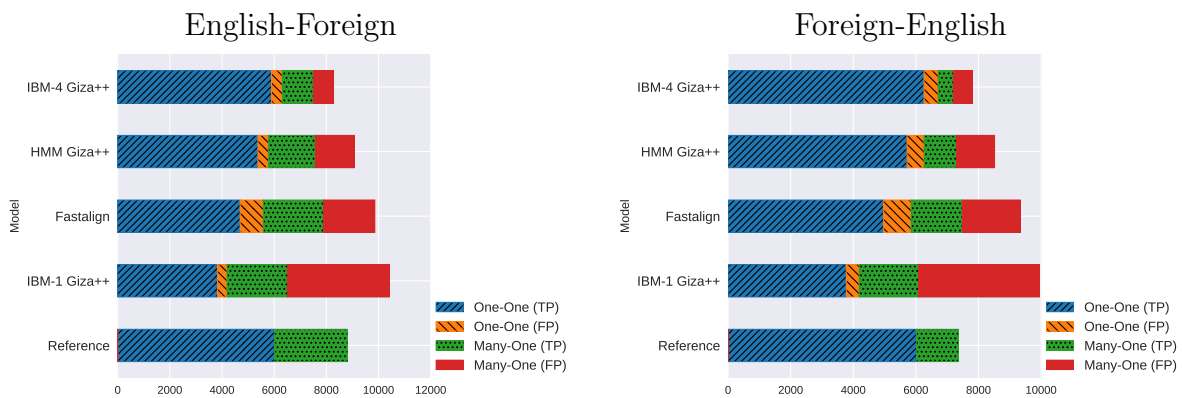


Figure A.20: Results of our baselines: Alignment types for English-German

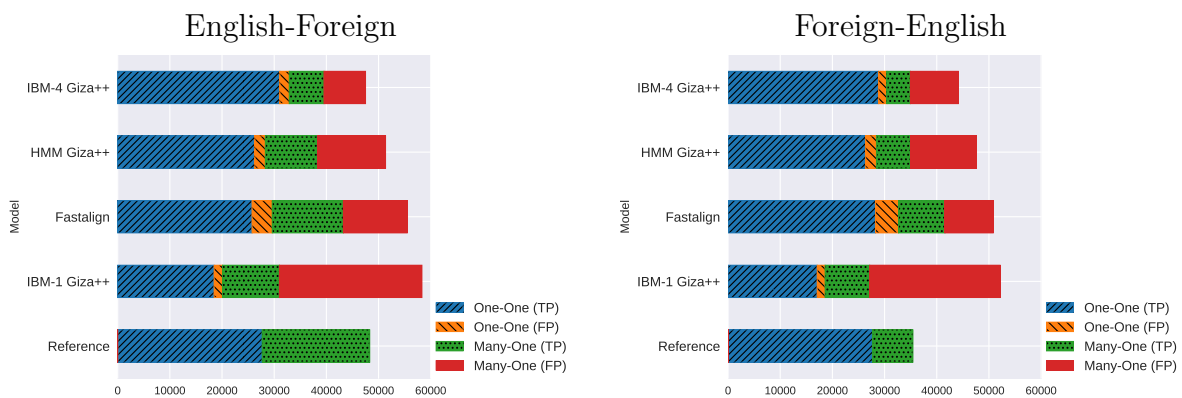


Figure A.21: Results of our baselines: Alignment types for English-Czech

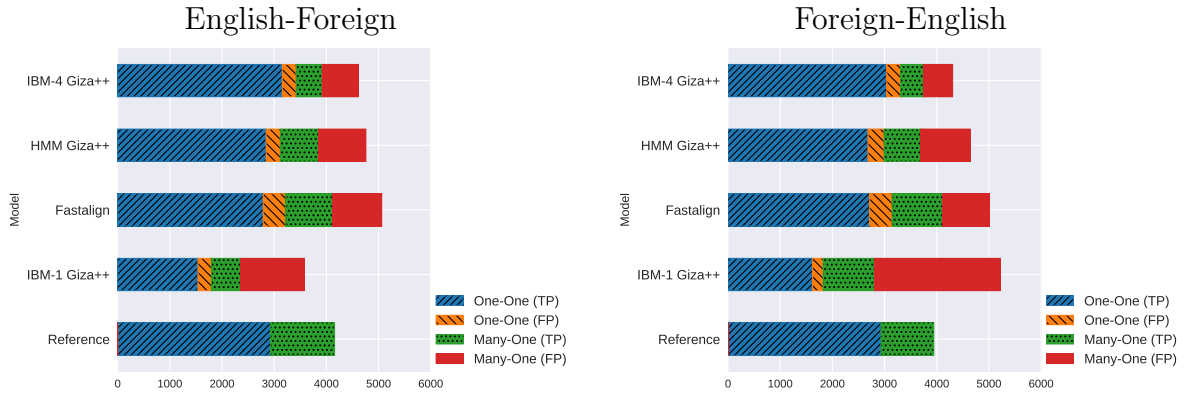


Figure A.22: Results of our baselines: Alignment types for English-Romanian

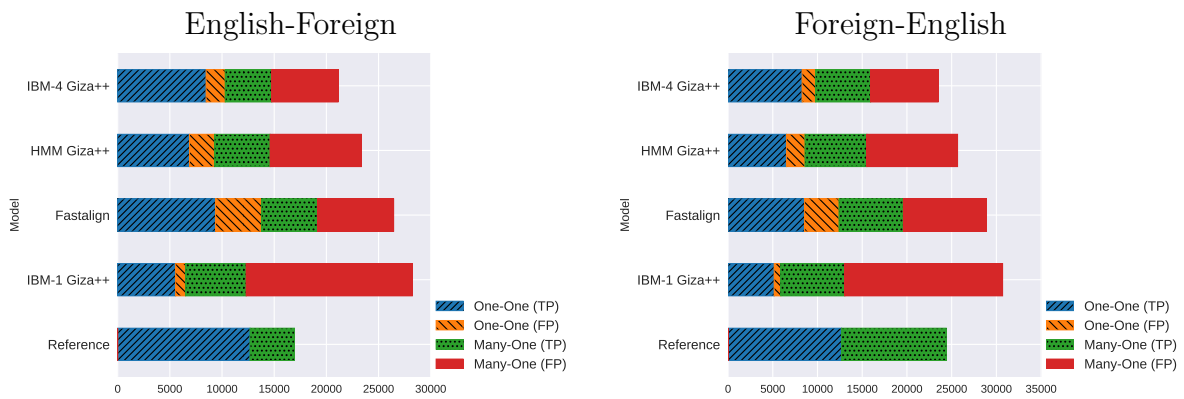


Figure A.23: Results of our baselines: Alignment types for English-Japanese

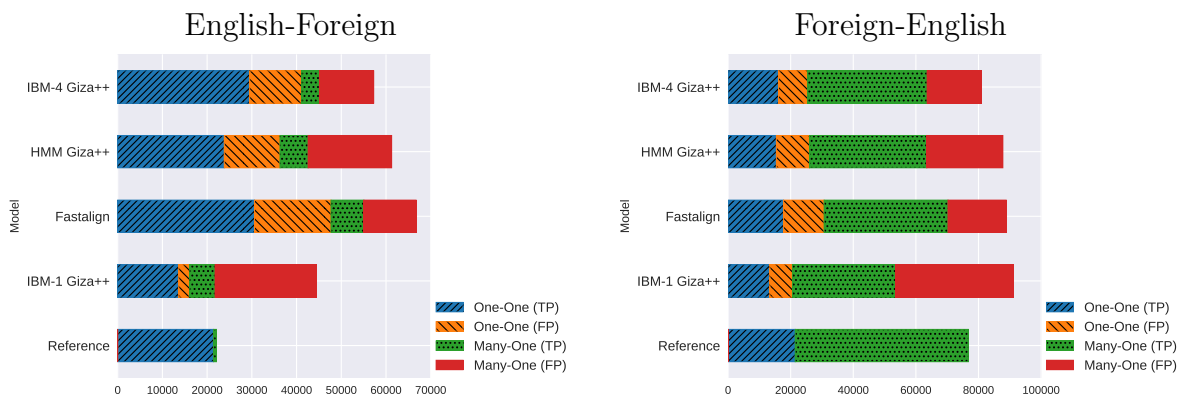


Figure A.24: Results of our baselines: Alignment types for English-Vietnamese

A.5 Jump width statistics

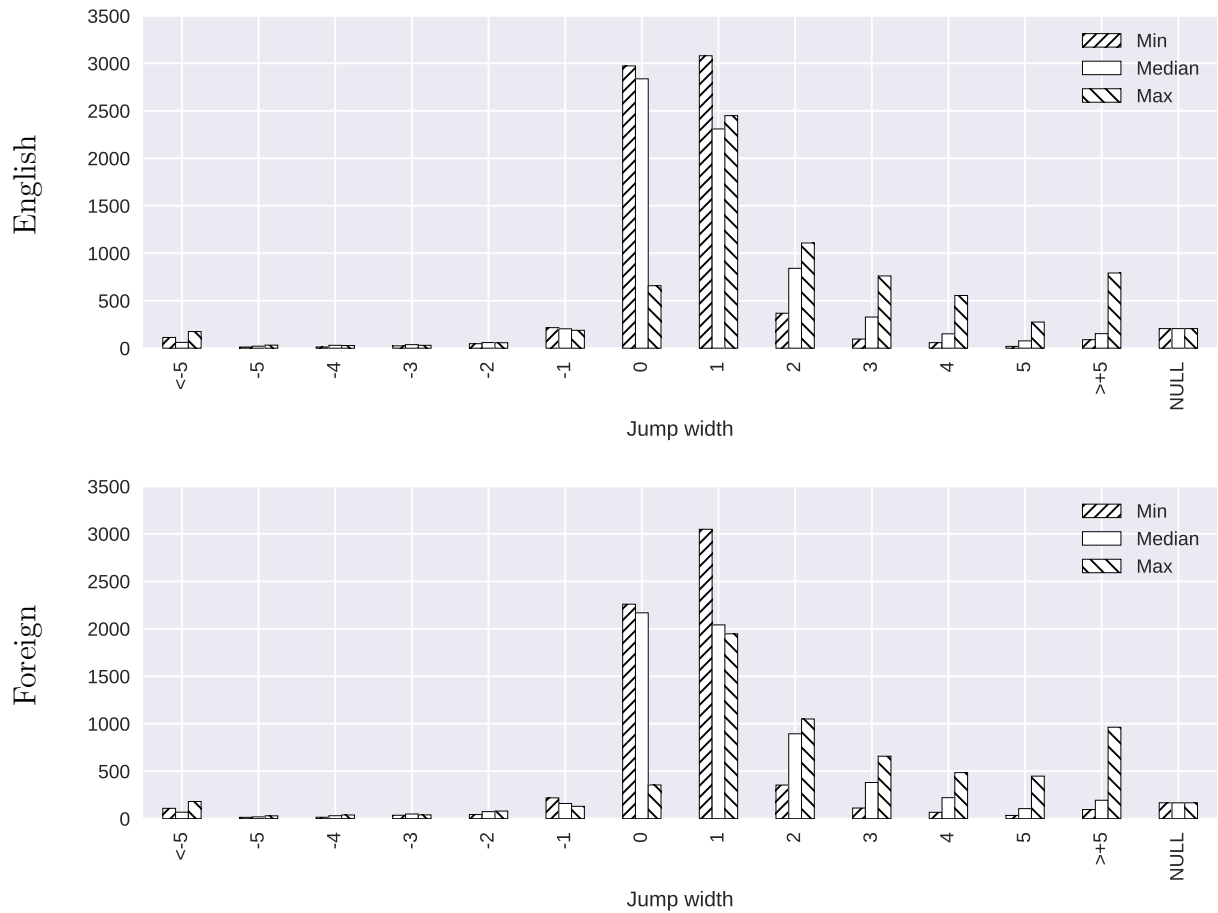


Figure A.25: Jump patterns for English words in the direction French-English and jump patterns for foreign words in the direction English-French

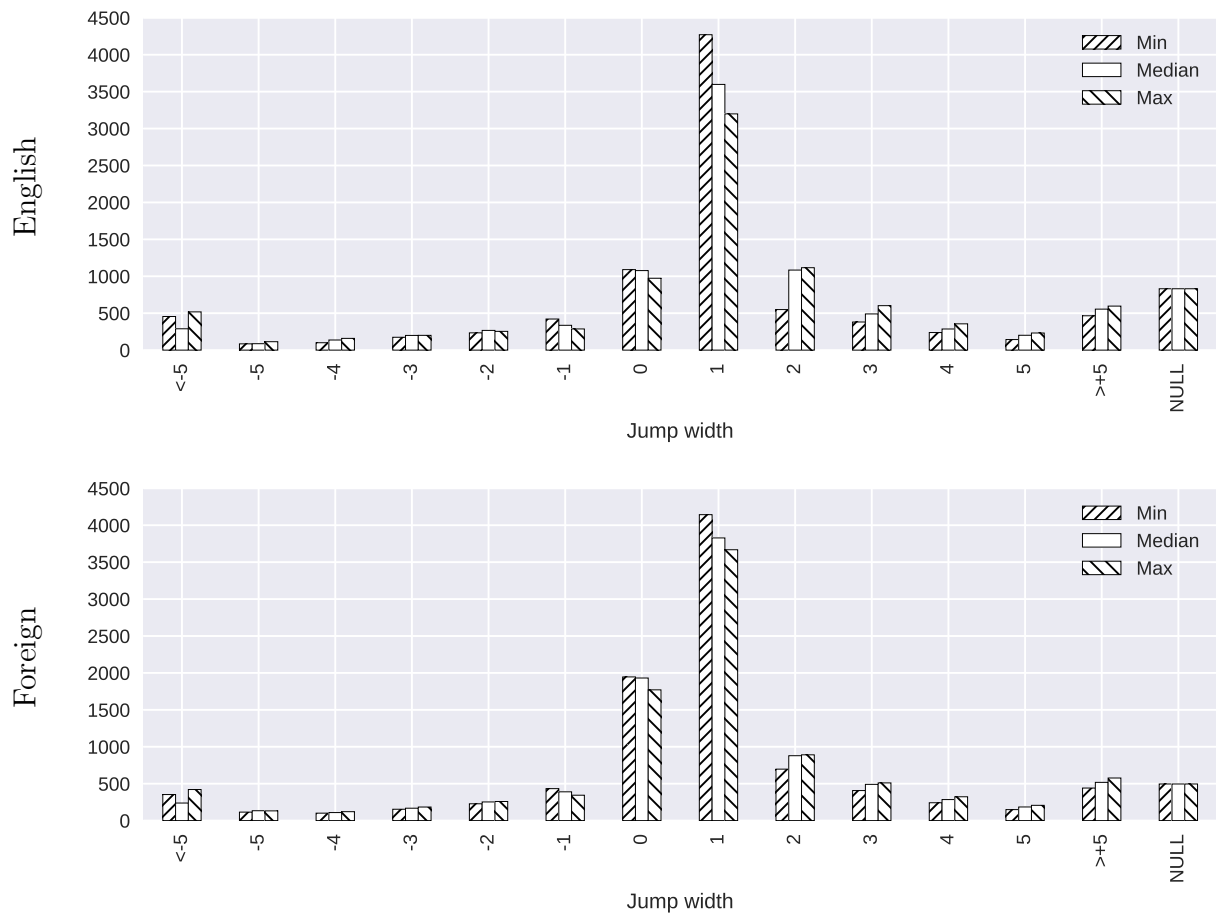


Figure A.26: Jump patterns for English words in the direction German-English and jump patterns for foreign words in the direction English-German

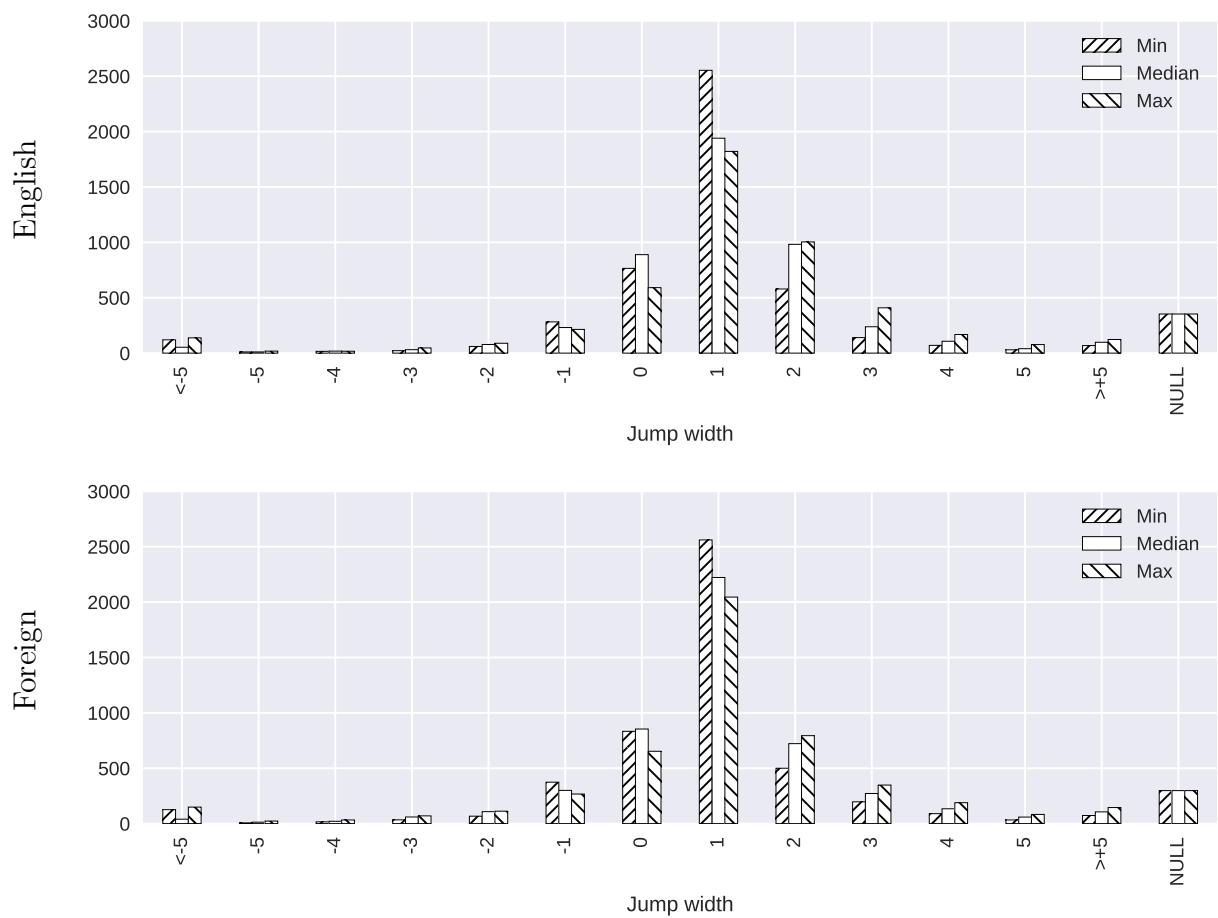


Figure A.27: Jump patterns for English words in the direction Romanian-English and jump patterns for foreign words in the direction English-Romanian

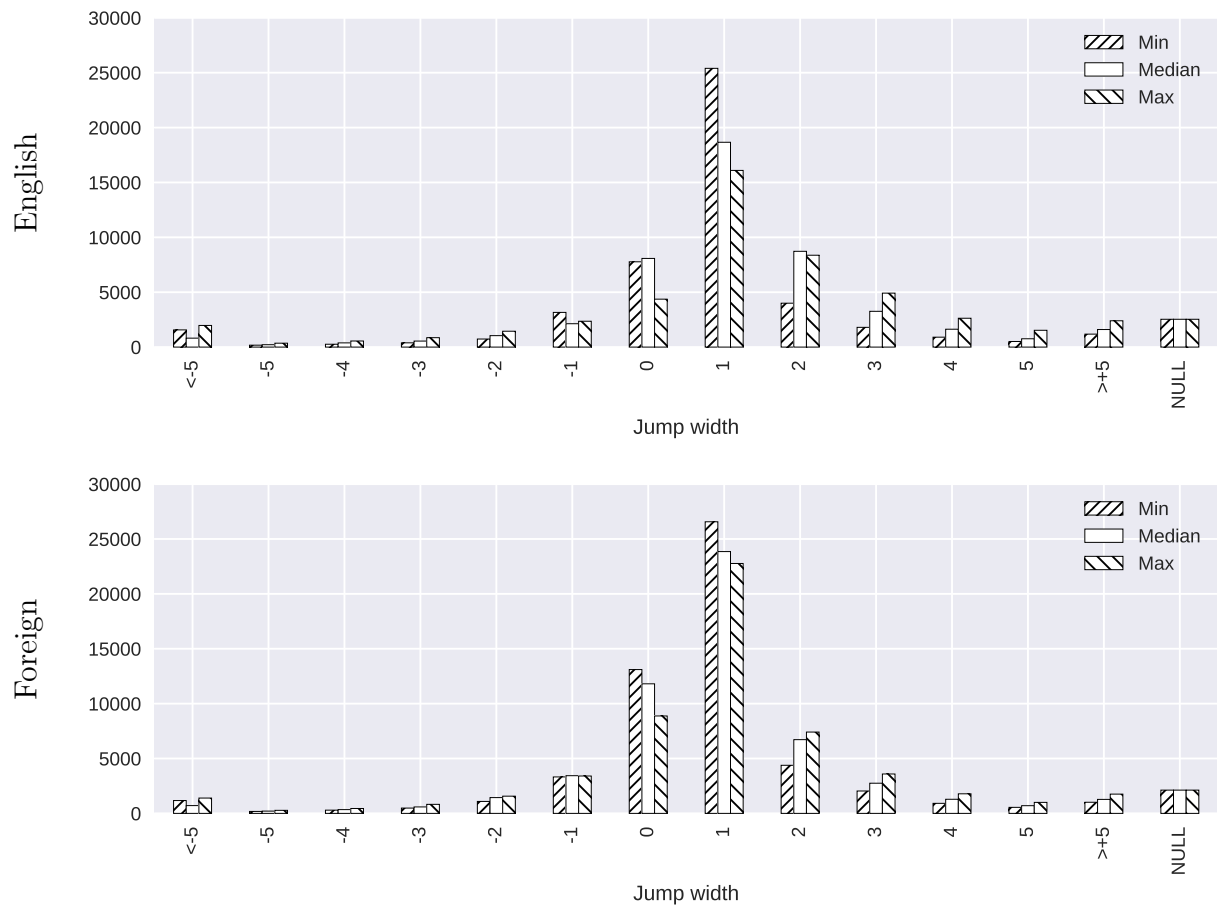


Figure A.28: Jump patterns for English words in the direction Czech-English and jump patterns for foreign words in the direction English-Czech

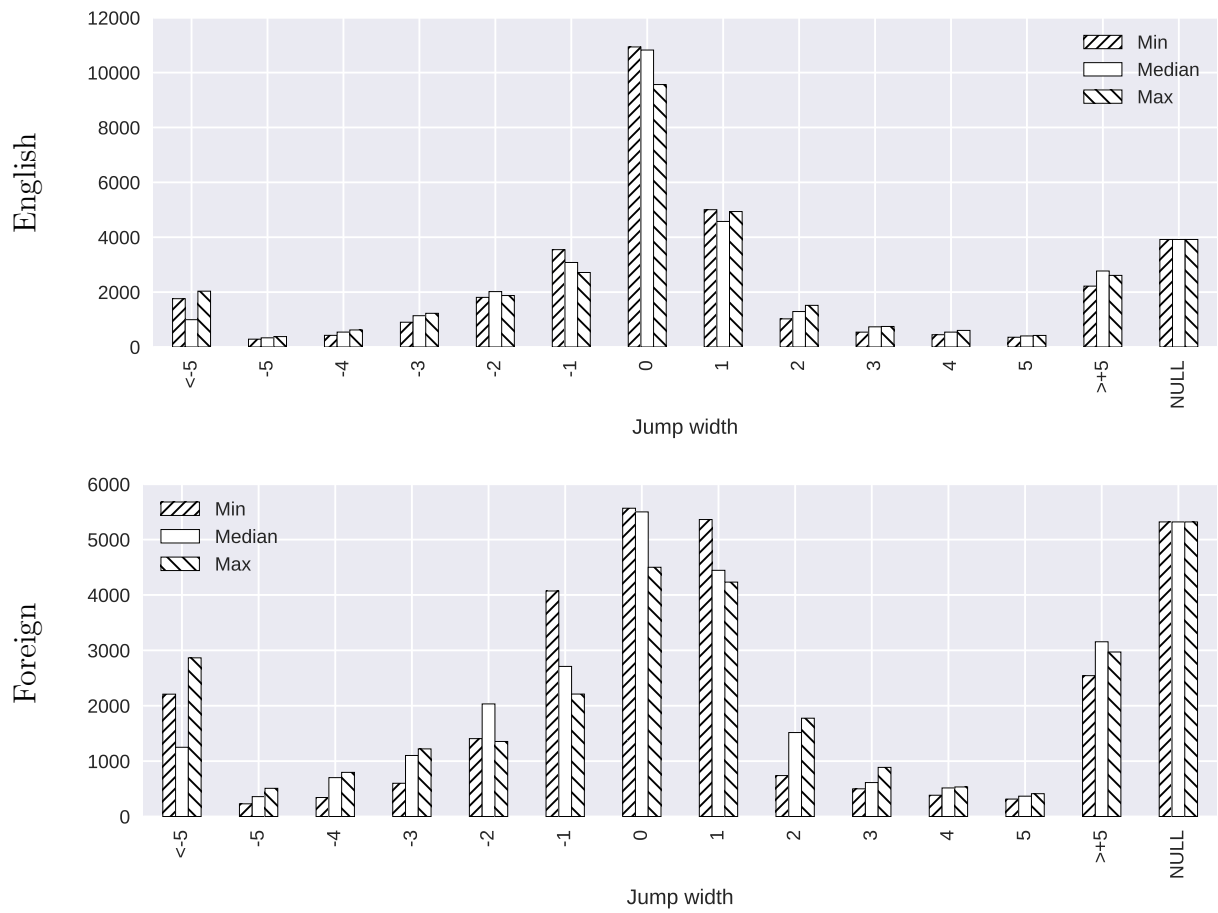


Figure A.29: Jump patterns for English words in the direction Japanese-English and jump patterns for foreign words in the direction English-Japanese

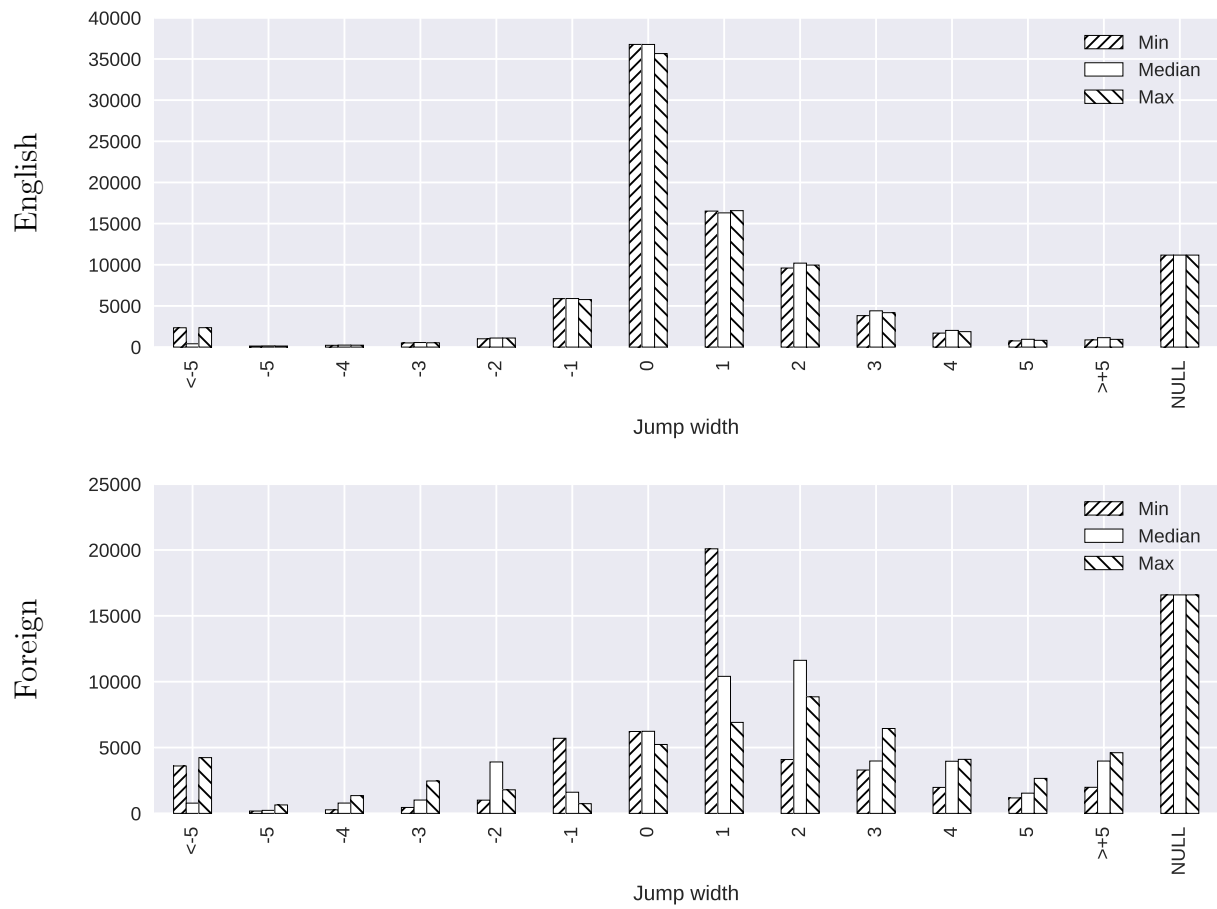


Figure A.30: Jump patterns for English words in the direction Vietnamese-English and jump patterns for foreign words in the direction English-Vietnamese

A.5.1 Baselines: Jump width statistics

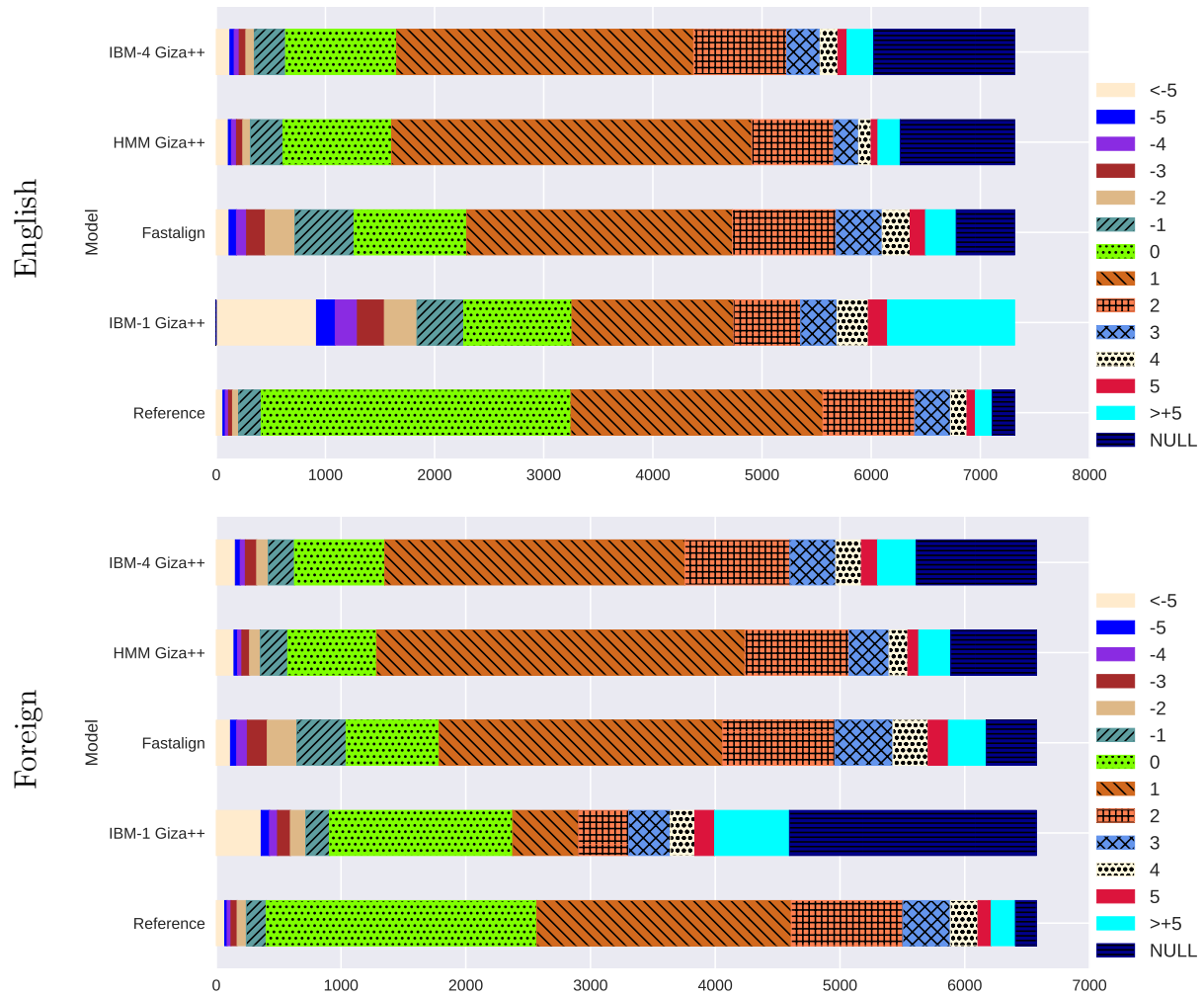


Figure A.31: Baselines: Jump patterns for English words in the direction French-English and jump patterns for foreign words in the direction English-French

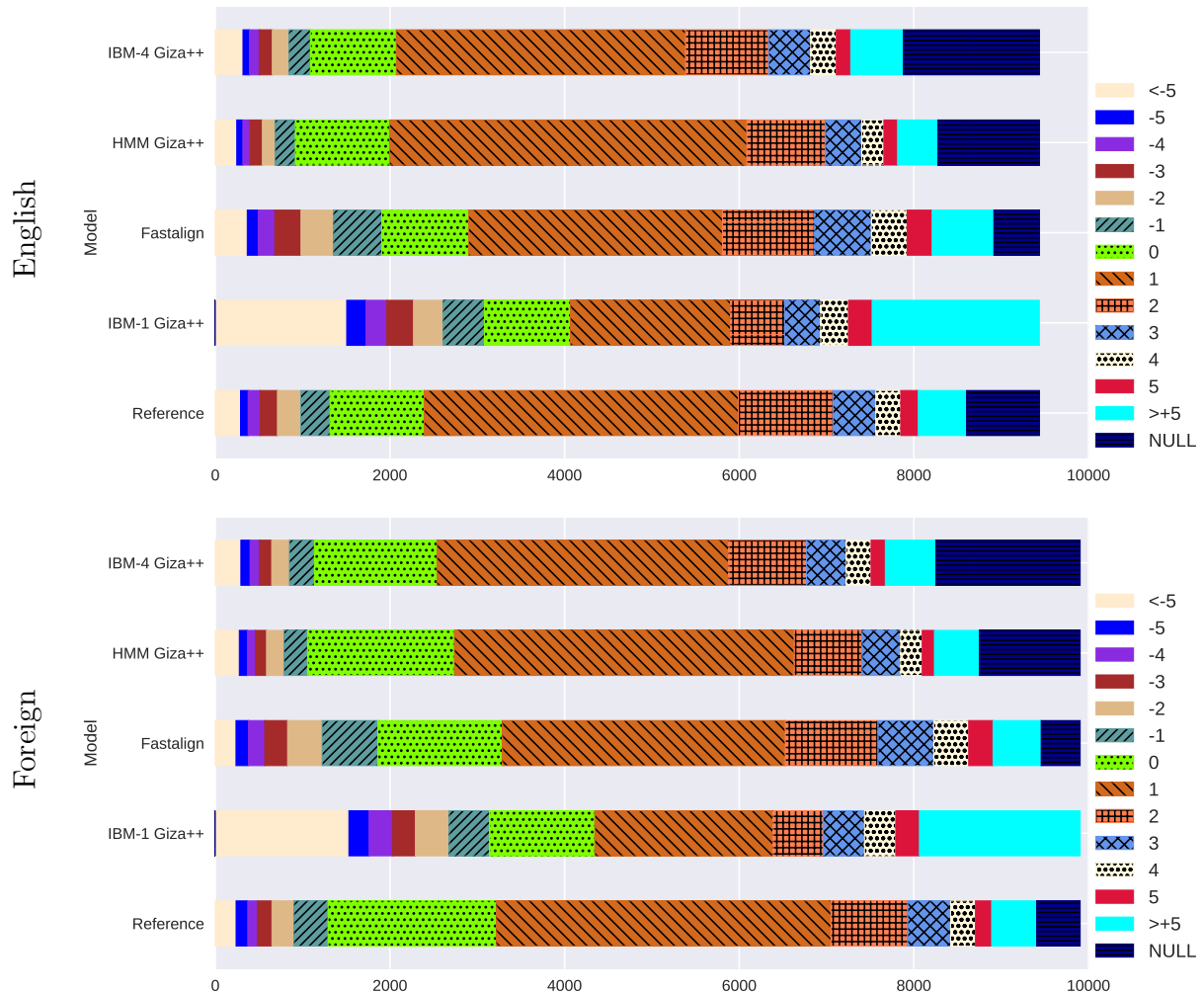


Figure A.32: Baselines: Jump patterns for English words in the direction German-English and jump patterns for foreign words in the direction English-German

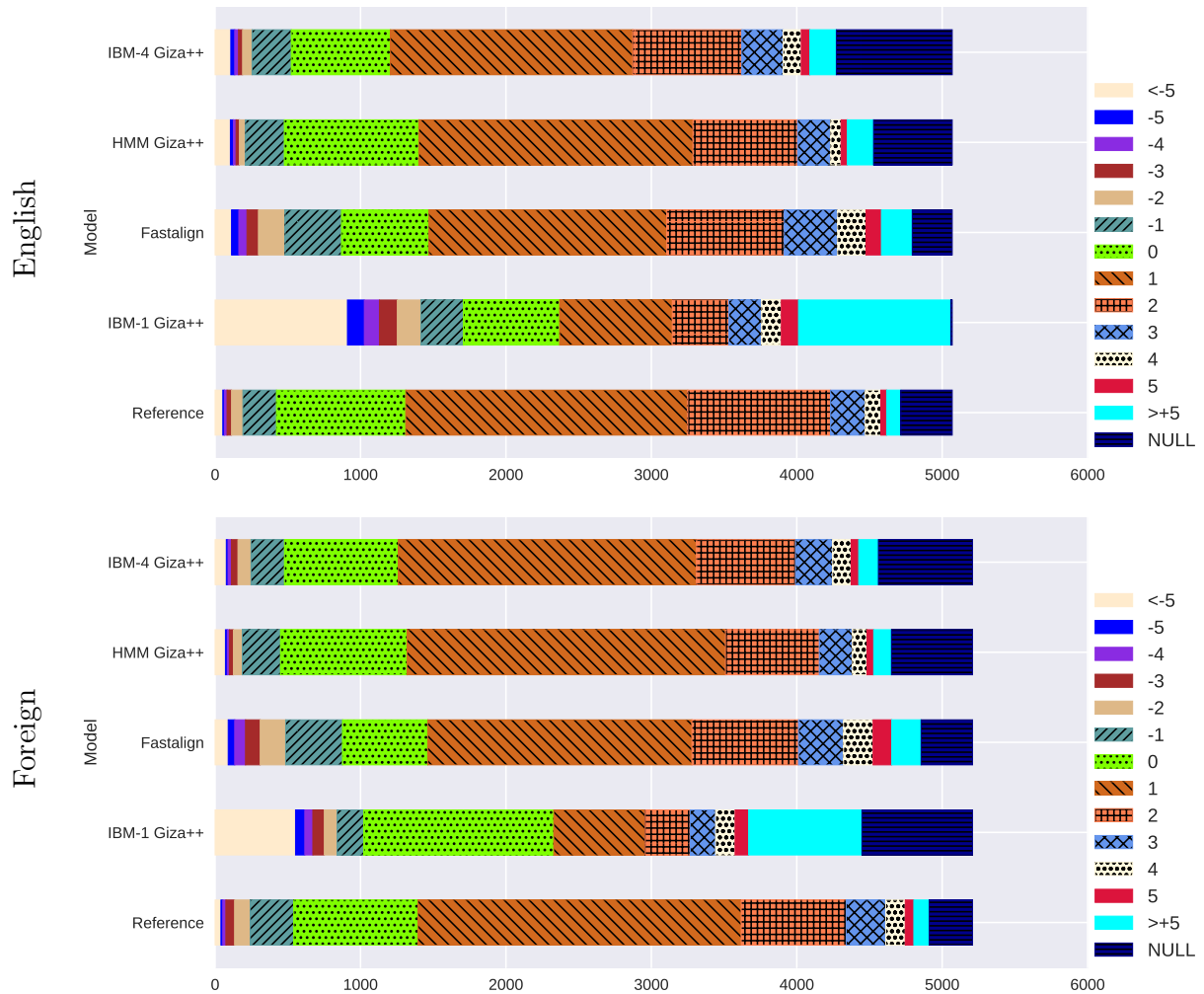


Figure A.33: Baselines: Jump patterns for English words in the direction Romanian-English and jump patterns for foreign words in the direction English-Romanian

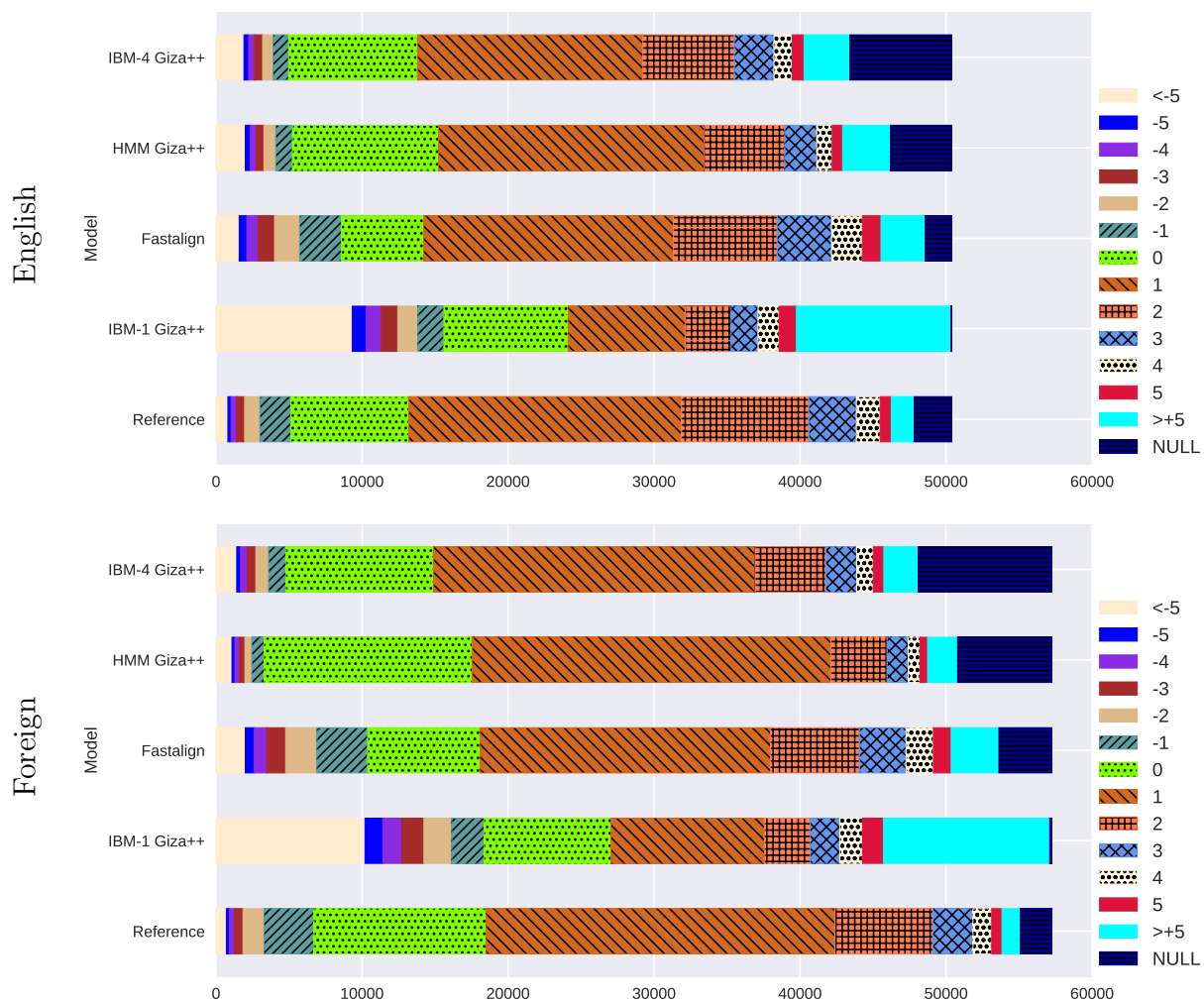


Figure A.34: Baselines: Jump patterns for English words in the direction Czech-English and jump patterns for foreign words in the direction English-Czech

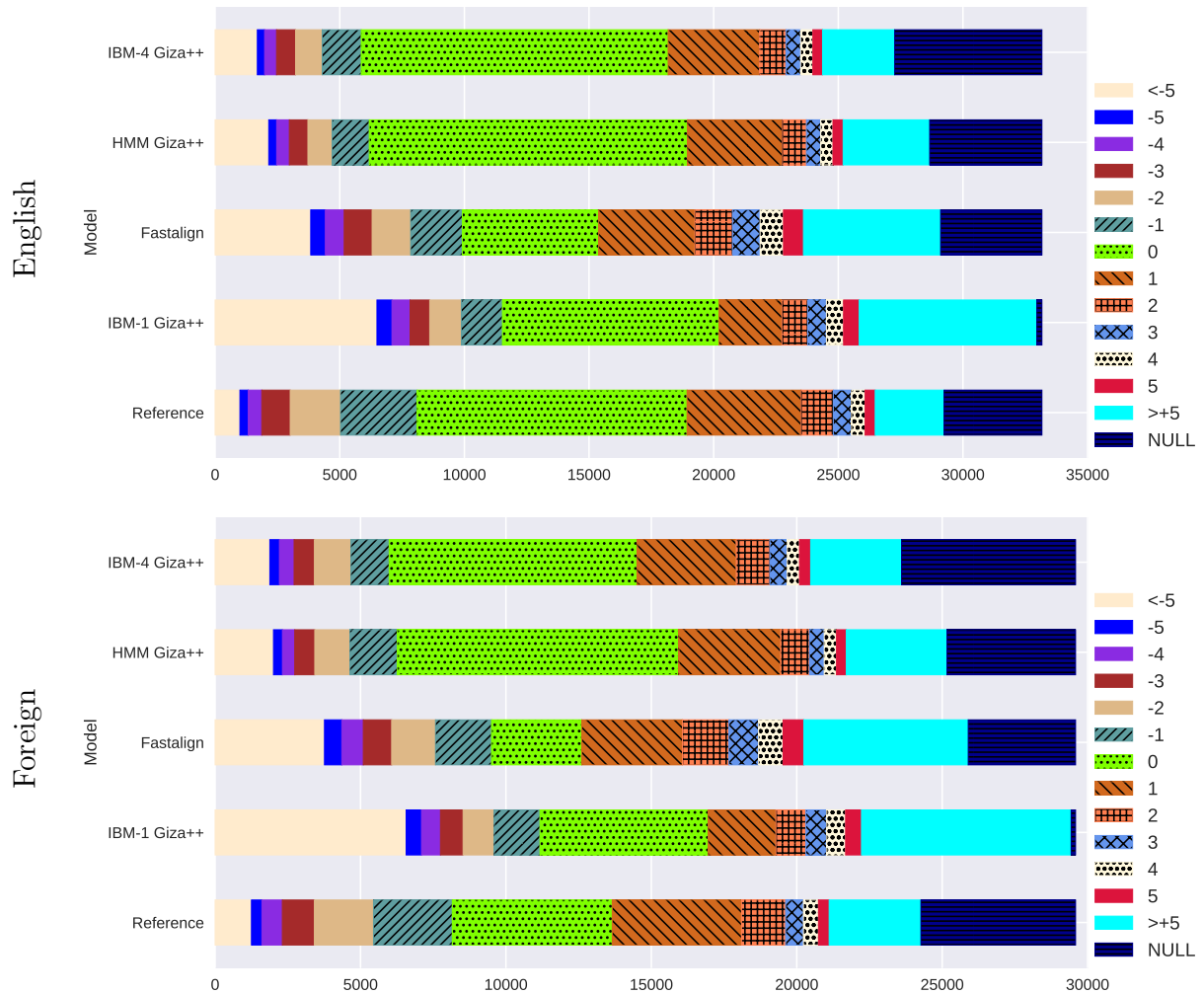


Figure A.35: Baselines: Jump patterns for English words in the direction Japanese-English and jump patterns for foreign words in the direction English-Japanese

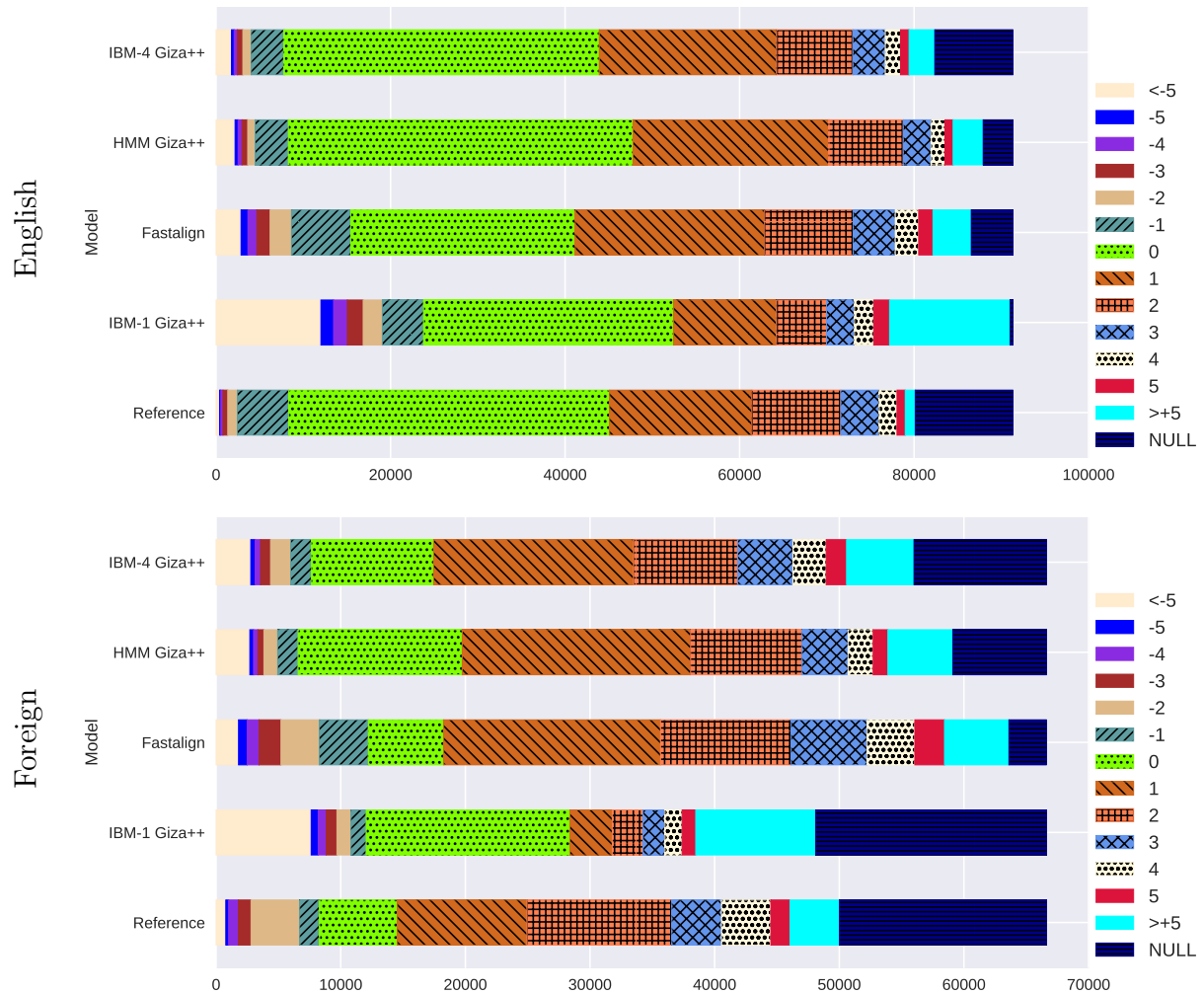


Figure A.36: Baselines: Jump patterns for English words in the direction Vietnamese-English and jump patterns for foreign words in the direction English-Vietnamese

A.5.2 Baselines: Correct and incorrect jumps

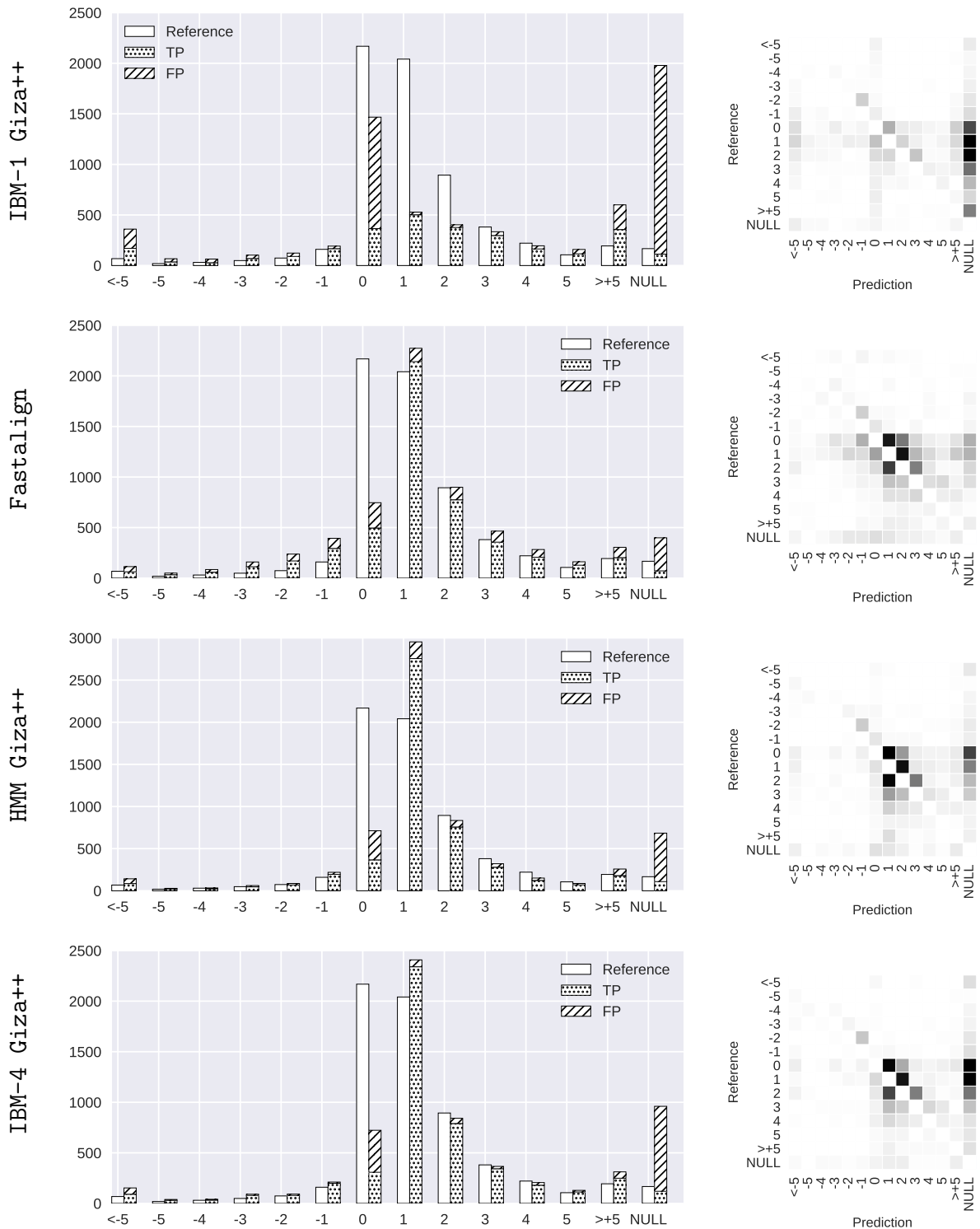


Figure A.37: Baselines for the direction English-French: Correct (TP) and incorrect (FP) jump widths for French words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.

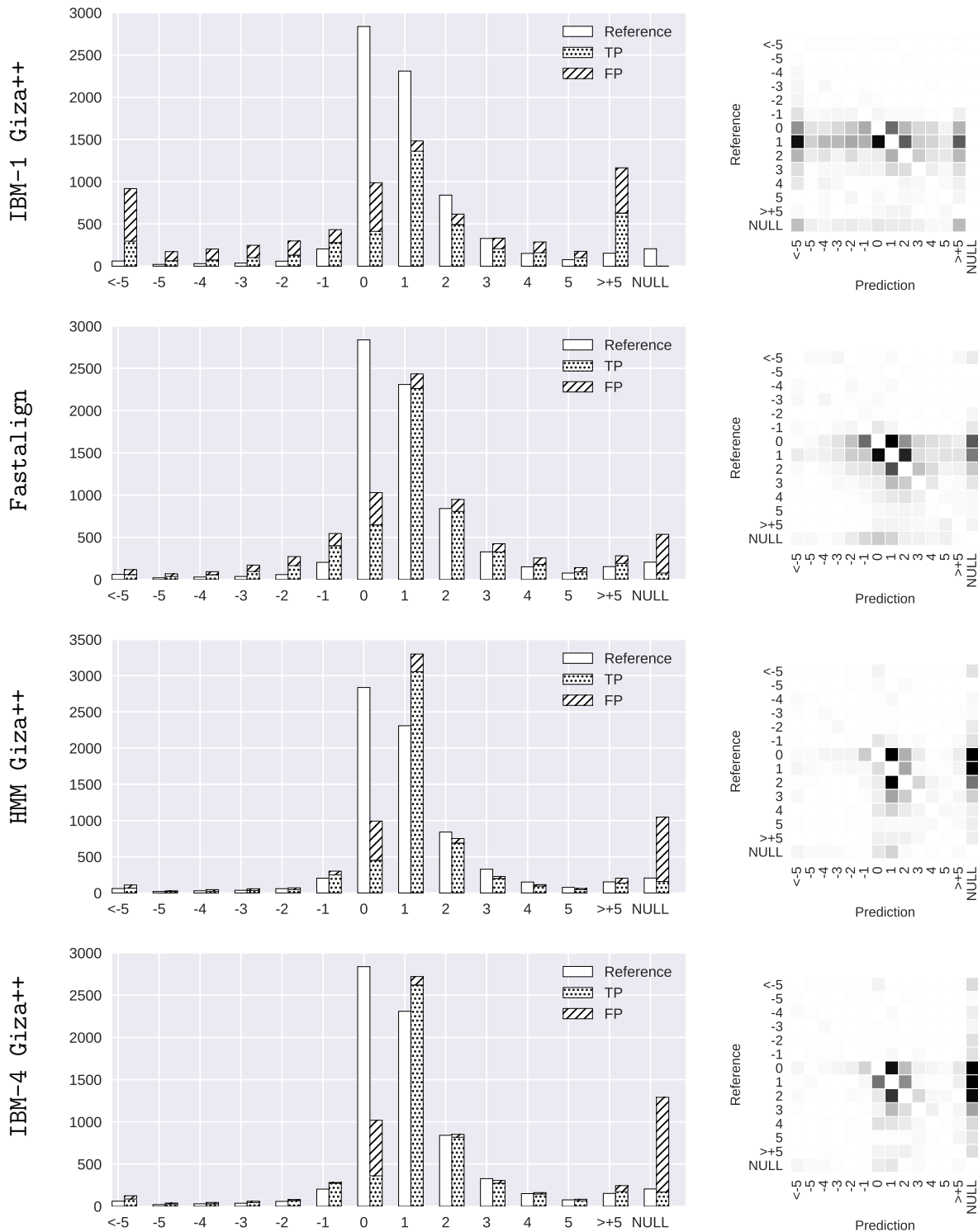


Figure A.38: Baselines for the direction French-English: Correct (TP) and incorrect (FP) jump widths for English words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.

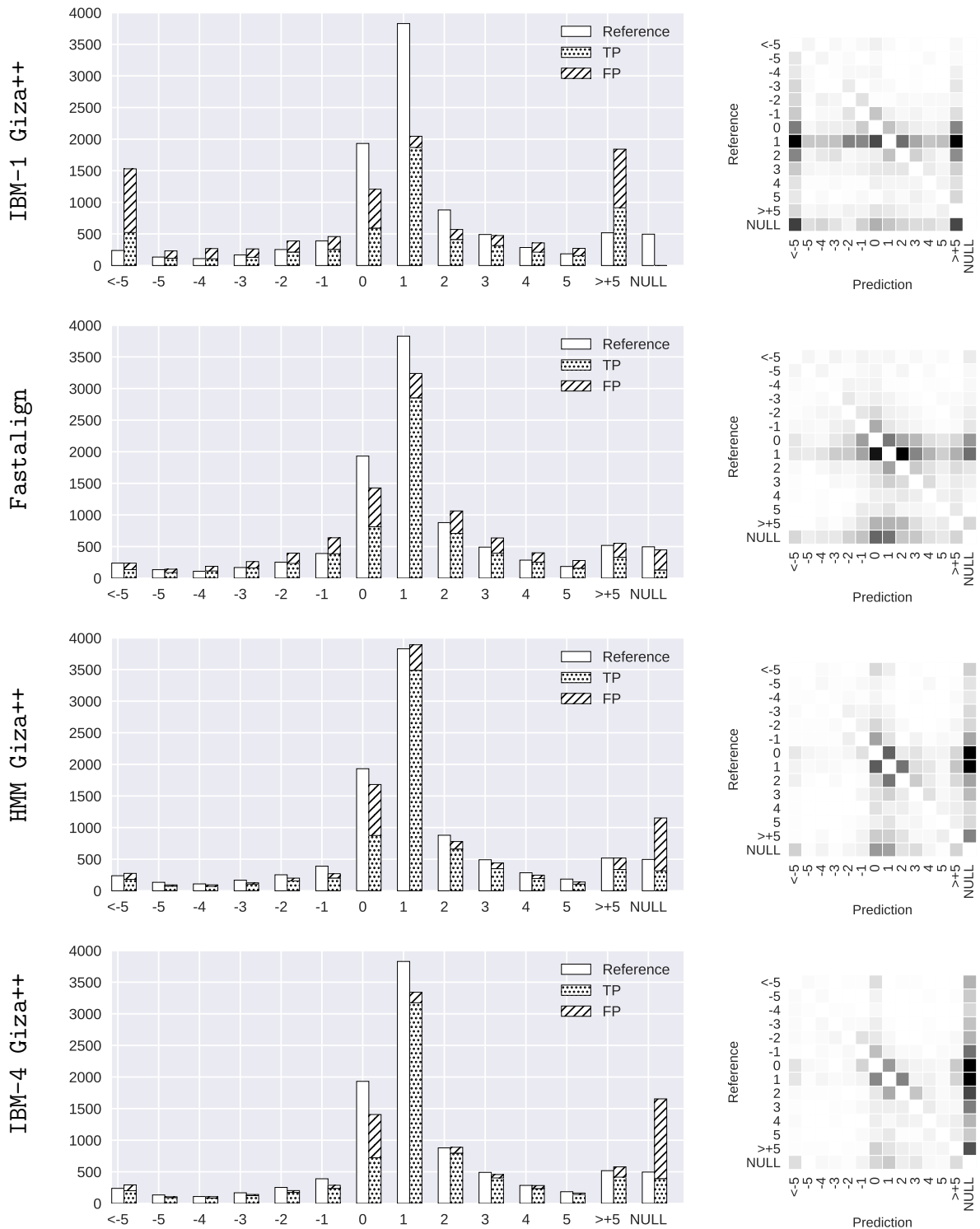


Figure A.39: Baselines for the direction English-German: Correct (TP) and incorrect (FP) jump widths for German words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.

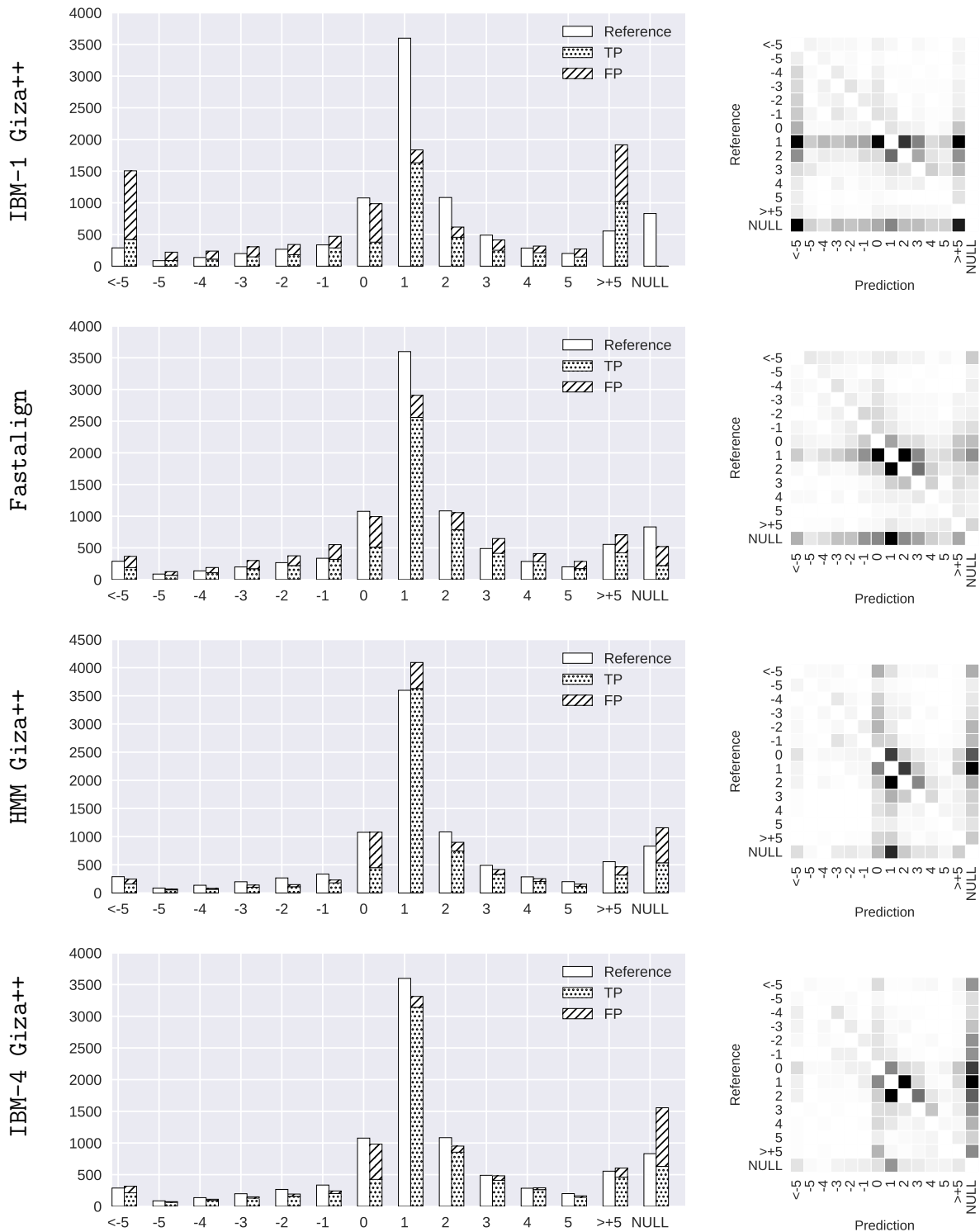


Figure A.40: Baselines for the direction German-English: Correct (TP) and incorrect (FP) jump widths for English words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.

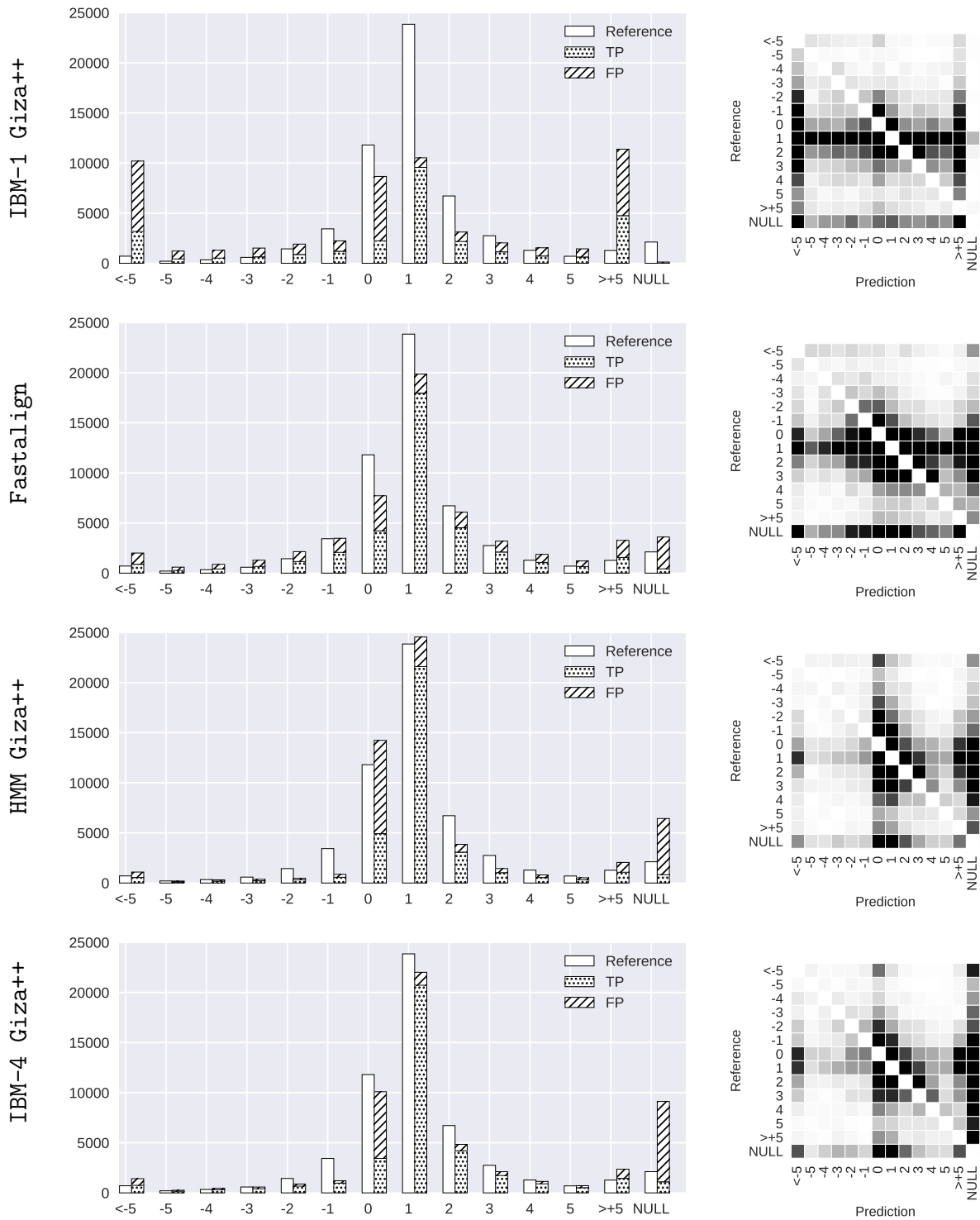


Figure A.41: Baselines for the direction English-Czech: Correct (TP) and incorrect (FP) jump widths for Czech words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.

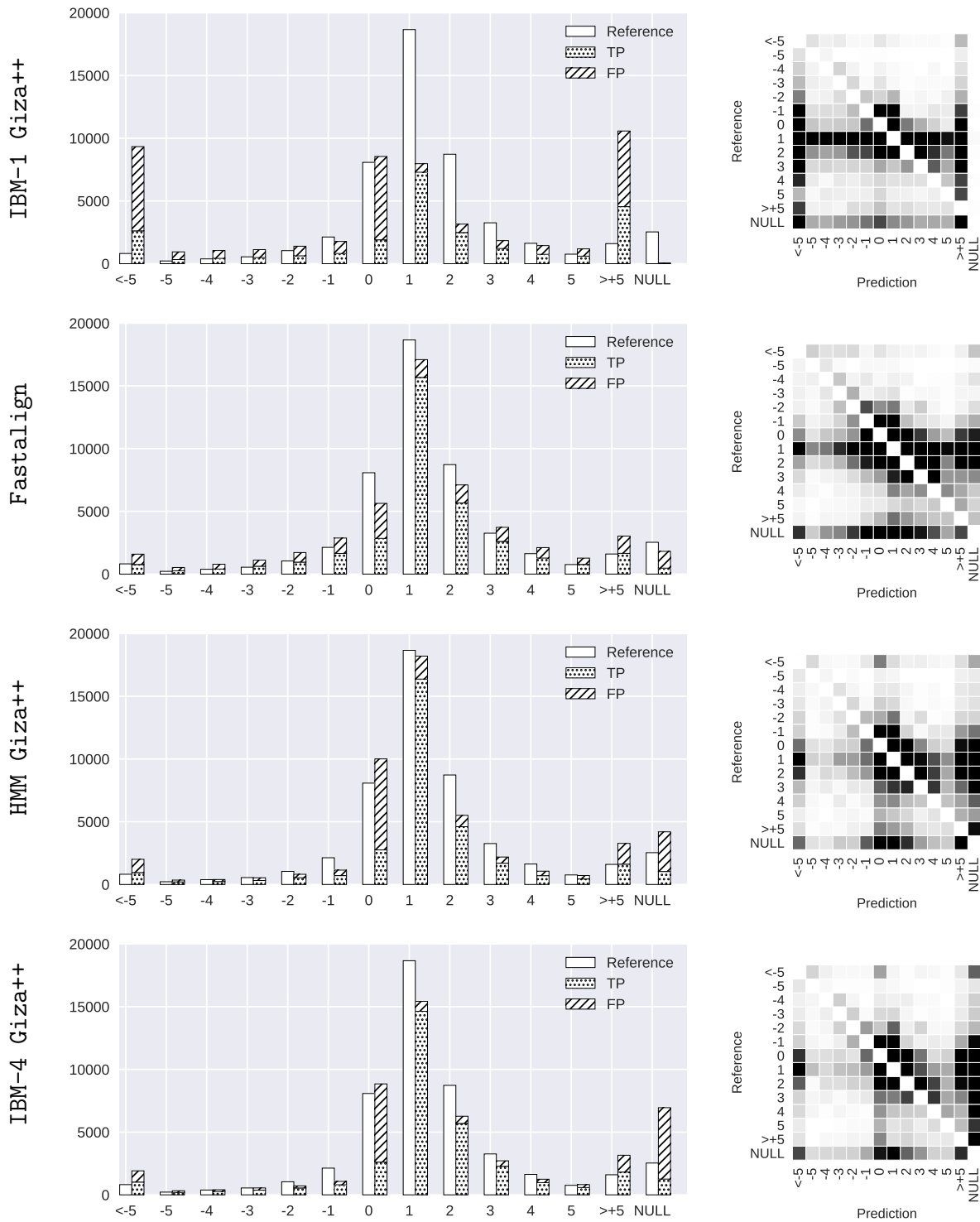


Figure A.42: Baselines for the direction Czech-English: Correct (TP) and incorrect (FP) jump widths for English words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.

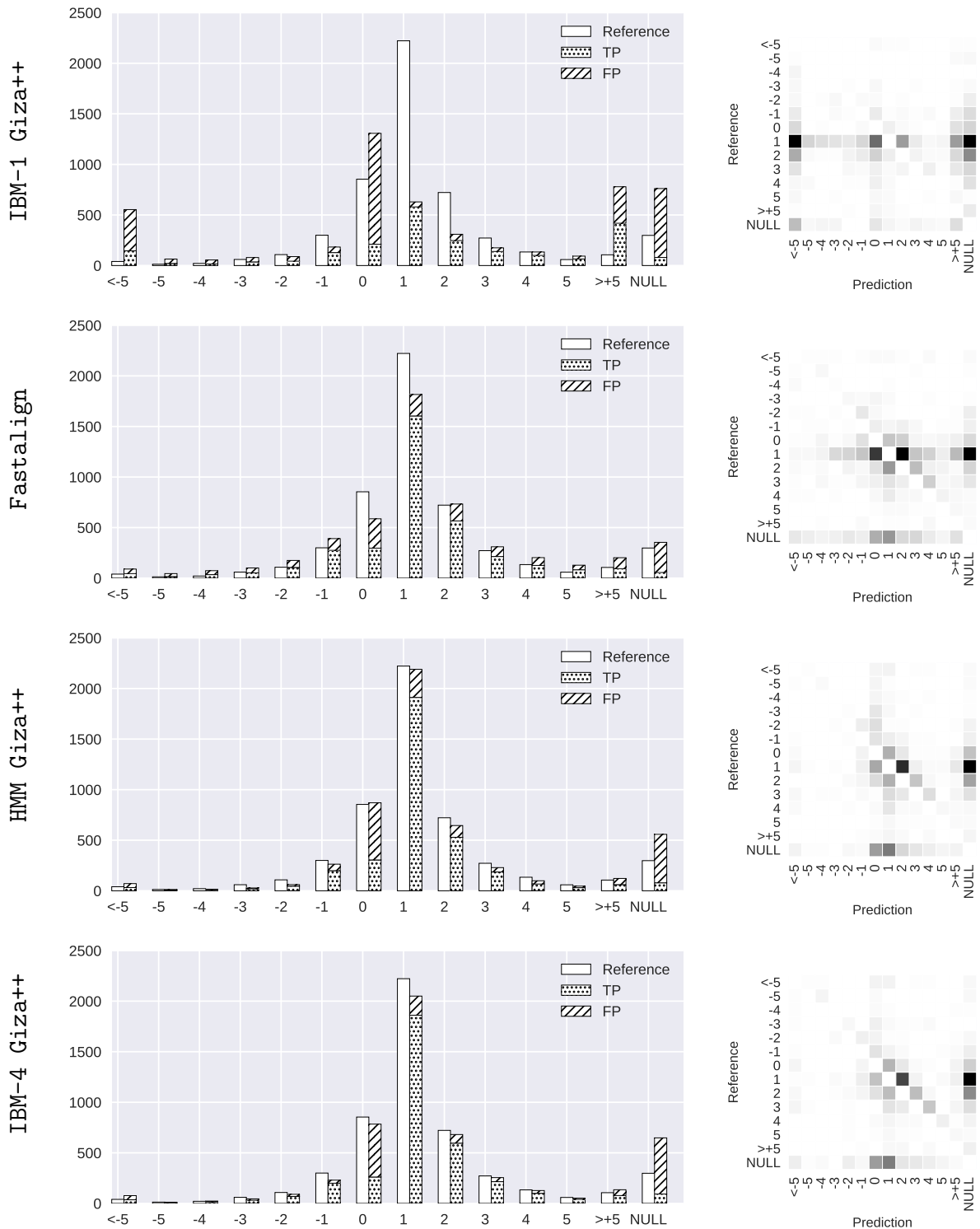


Figure A.43: Baselines for the direction English-Romanian: Correct (TP) and incorrect (FP) jump widths for Romanian words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.

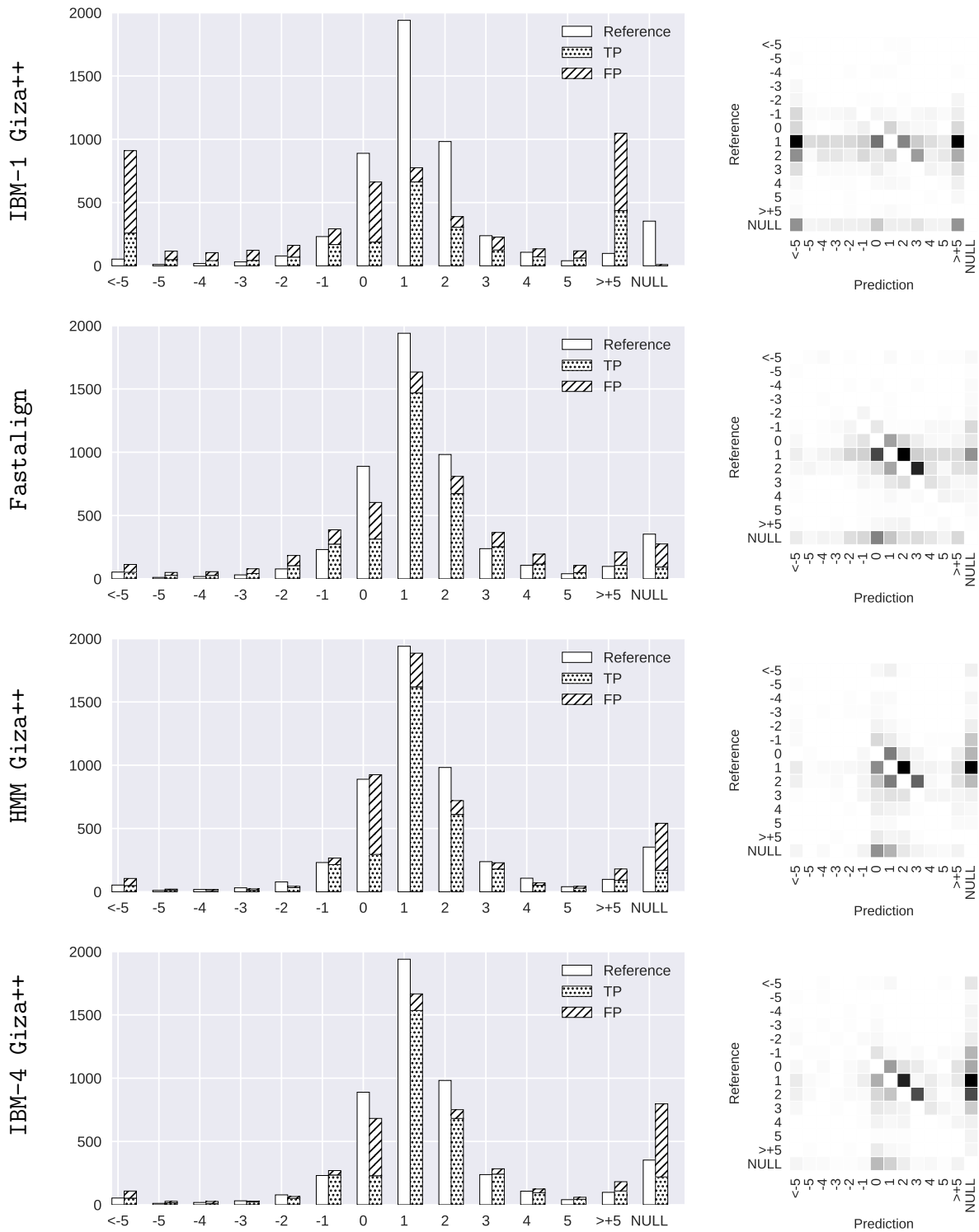


Figure A.44: Baselines for the direction Romanian-English: Correct (TP) and incorrect (FP) jump widths for English words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.

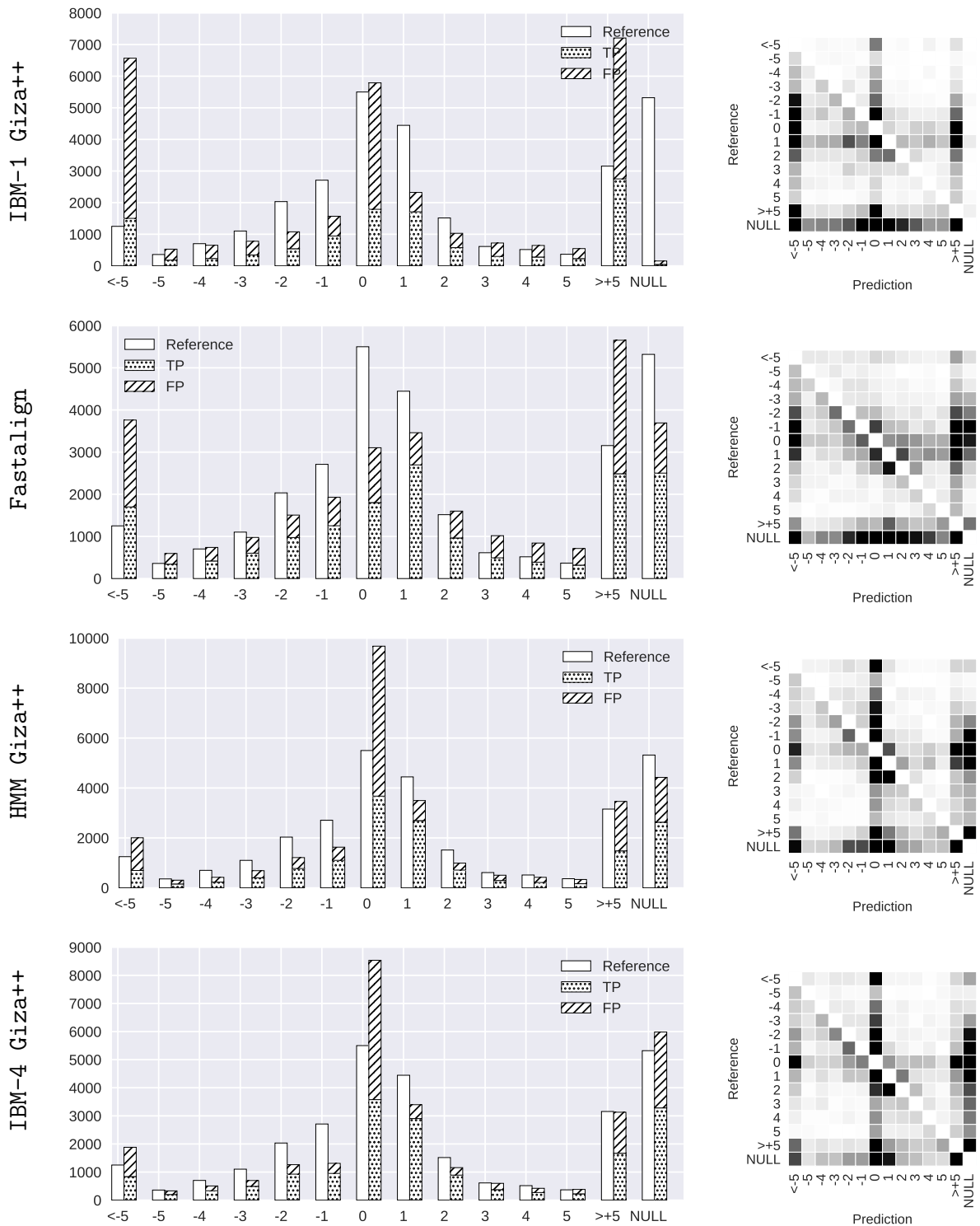


Figure A.45: Baselines for the direction English-Japanese: Correct (TP) and incorrect (FP) jump widths for Japanese words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.

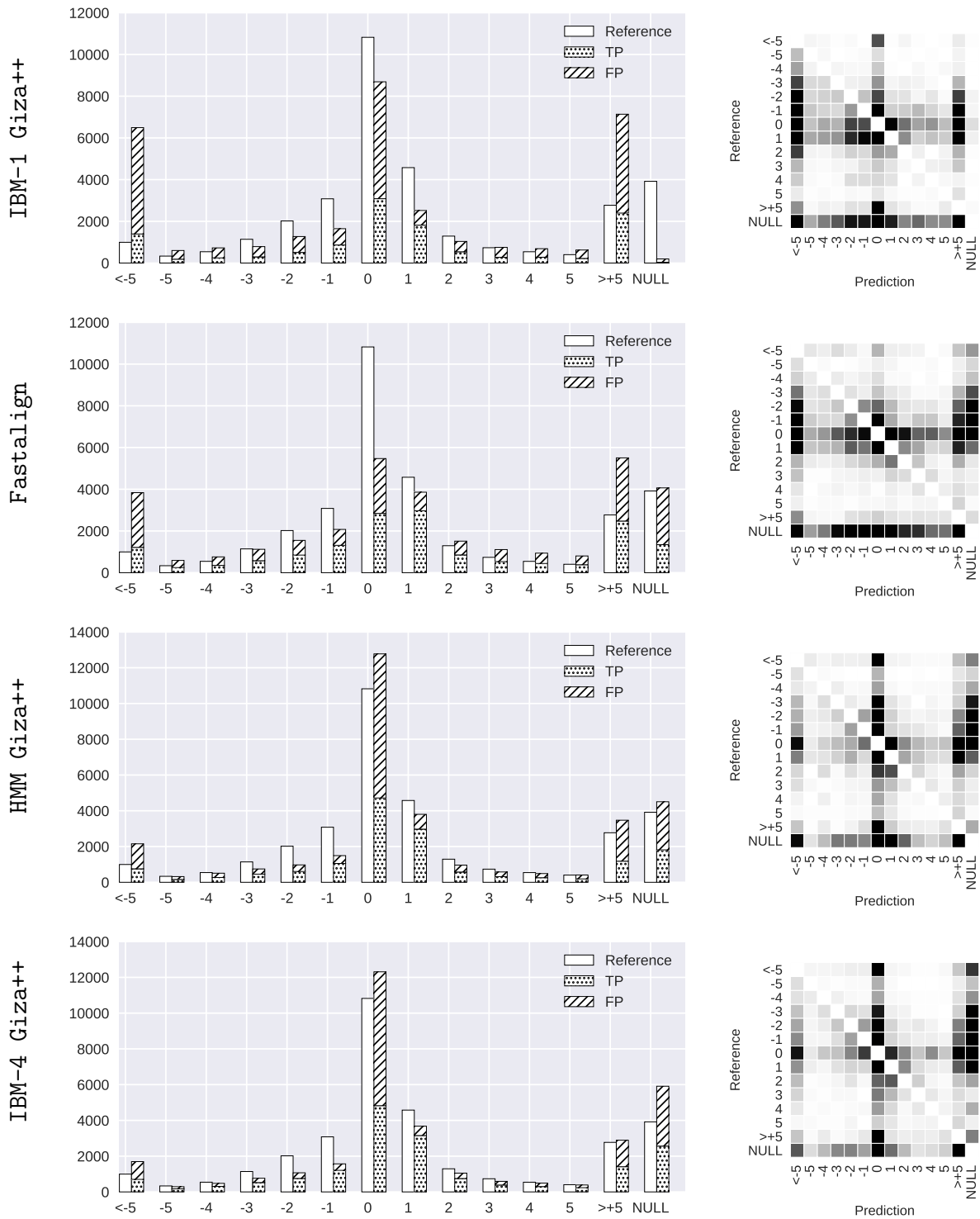


Figure A.46: Baselines for the direction Japanese-English: Correct (TP) and incorrect (FP) jump widths for English words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.

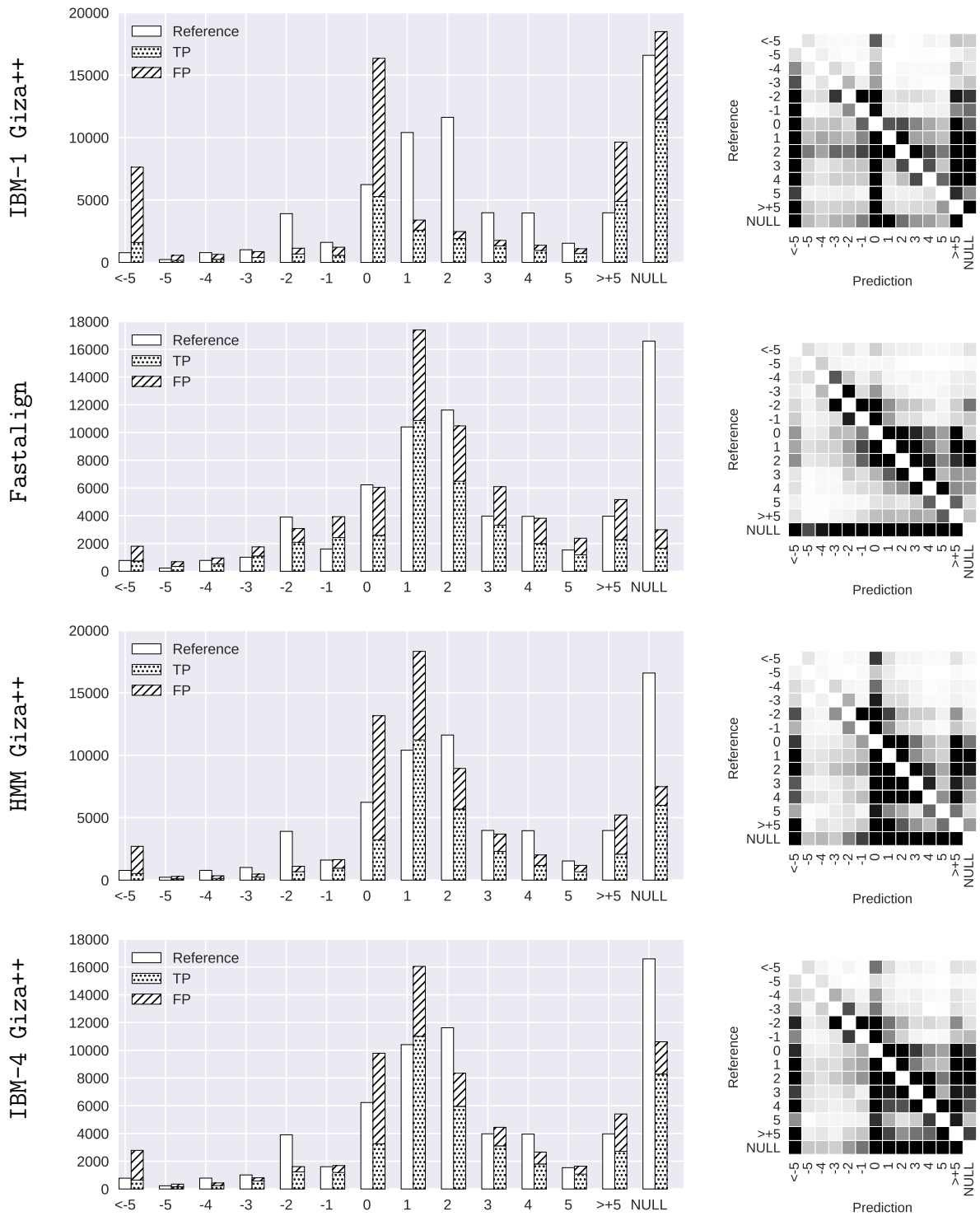


Figure A.47: Baselines for the direction English-Vietnamese: Correct (TP) and incorrect (FP) jump widths for Vietnamese words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.

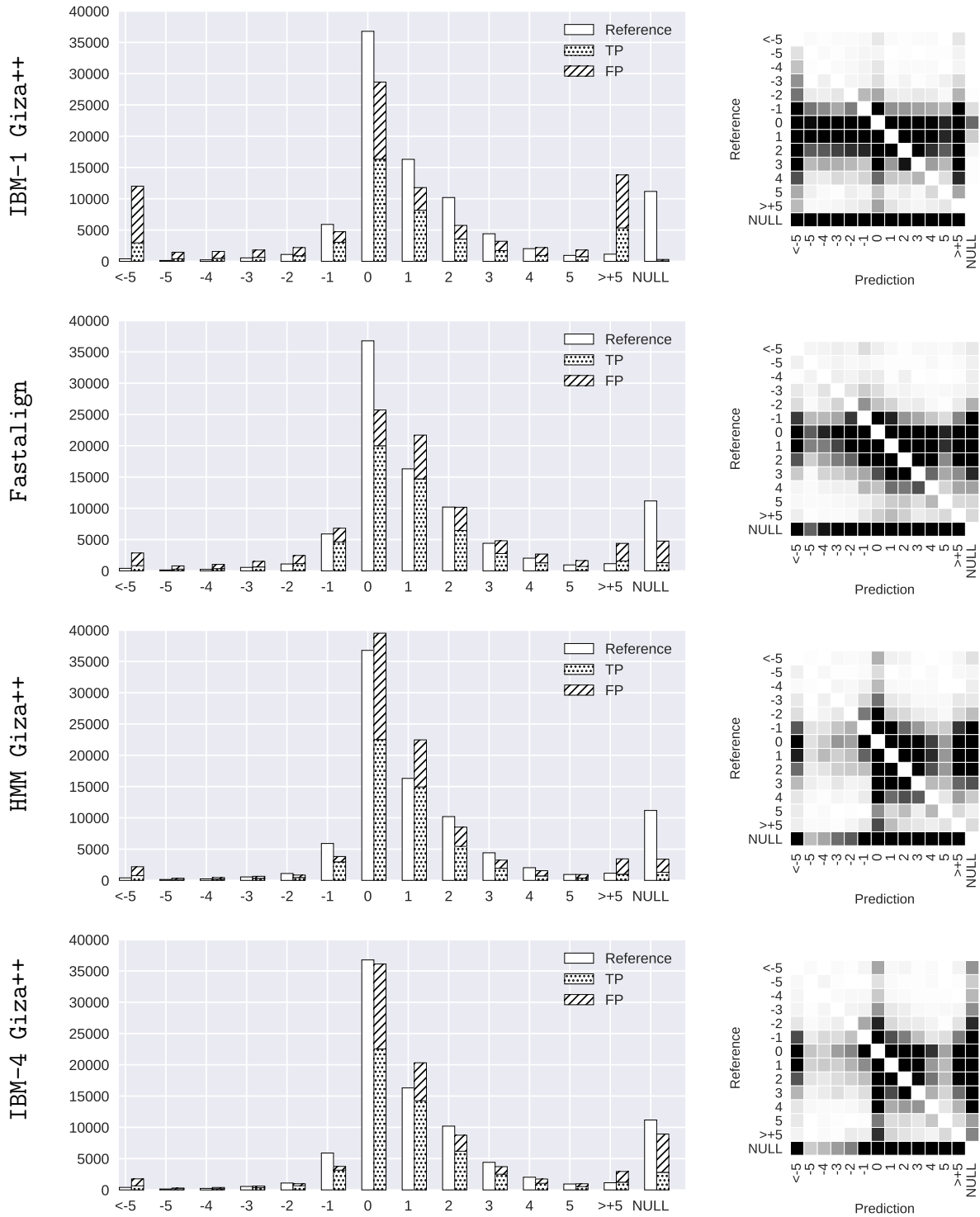


Figure A.48: Baselines for the direction Vietnamese-English: Correct (TP) and incorrect (FP) jump widths for English words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.

A.6 Word length as a function of word occurrence

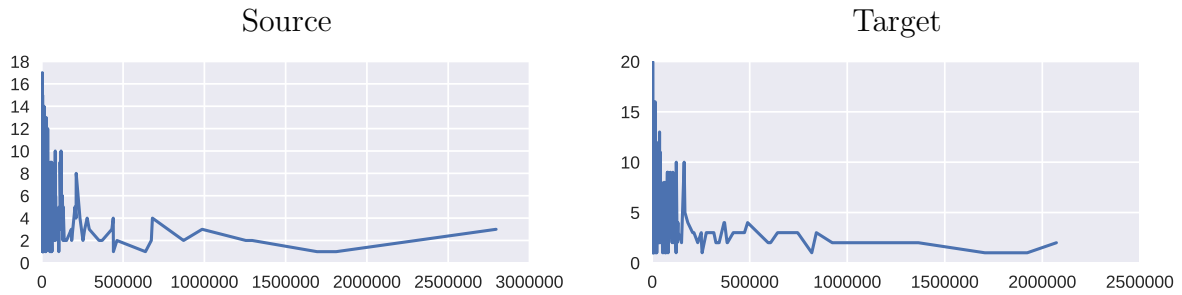


Figure A.49: The direction English-French: Word length as a function of word occurrence.

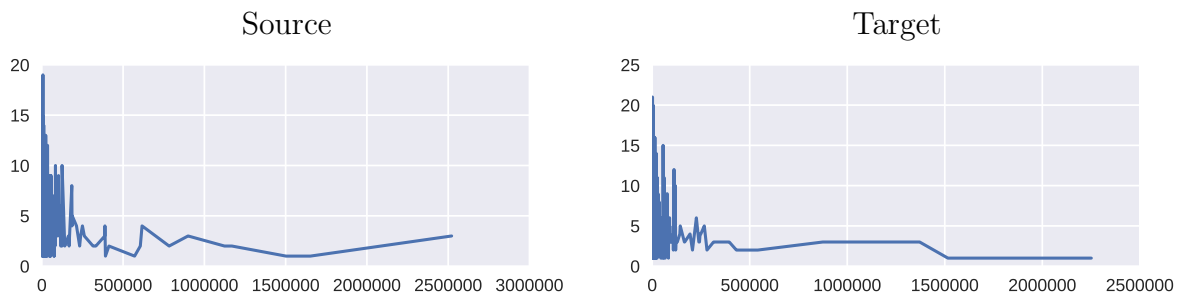


Figure A.50: The direction English-German: Word length as a function of word occurrence.

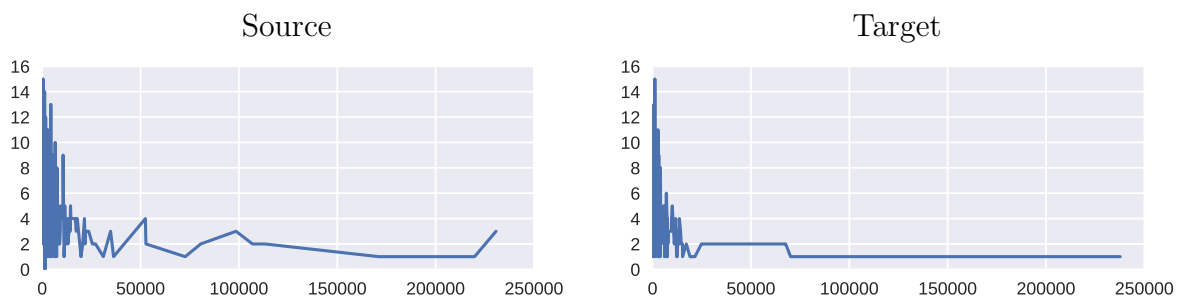


Figure A.51: The direction English-Czech: Word length as a function of word occurrence.

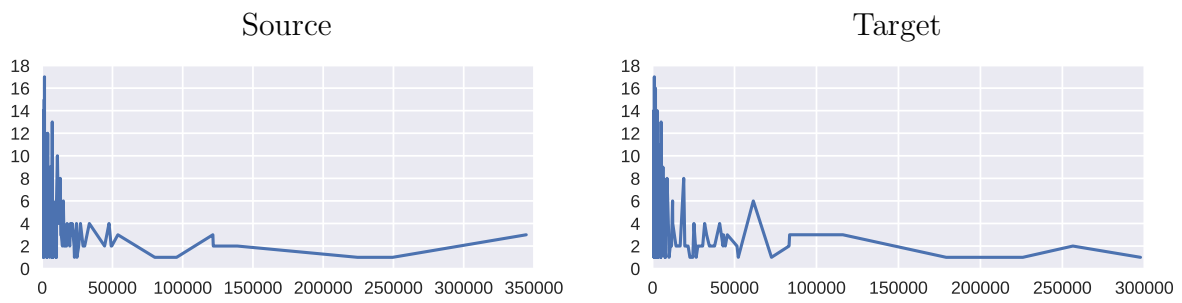


Figure A.52: The direction English-Romanian: Word length as a function of word occurrence.

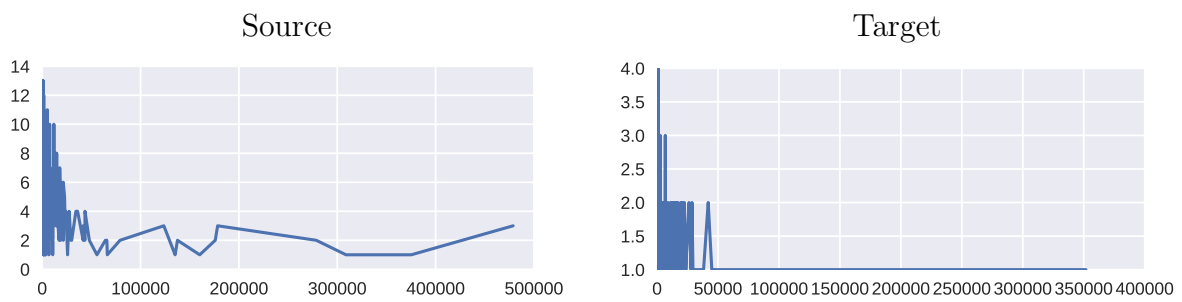


Figure A.53: The direction English-Japanese: Word length as a function of word occurrence.

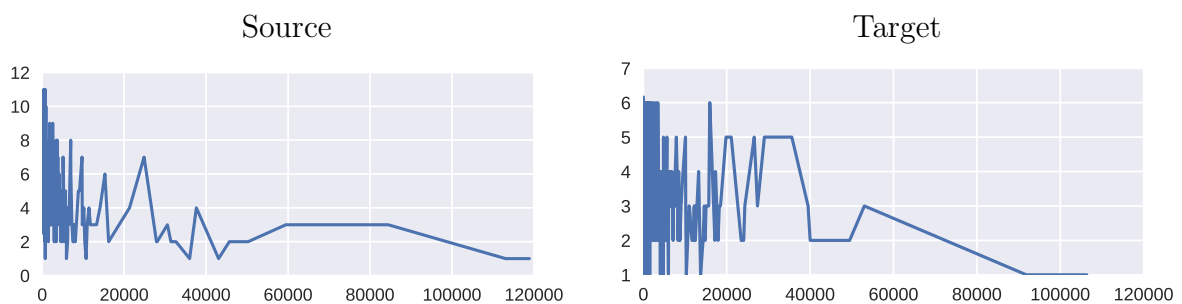


Figure A.54: The direction English-Vietnamese: Word length as a function of word occurrence.

A.7 Baselines: Rare words

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	45	4.09	86.61	31.11	93.33	46.67	64	2.78	86.21	23.44	40.54	29.7
Fastalign	11	1	95.82	72.73	53.33	61.54	20	0.87	95.15	80	43.24	56.14
HMM Giza++	33	3	91.63	42.42	93.33	58.33	72	3.13	87.38	30.56	59.46	40.37
IBM-4 Giza++	21	1.91	94.14	52.38	73.33	61.11	55	2.39	89.9	36.36	54.05	43.48

Table A.7: Baselines for English-French: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction French-English and in the direction English-French

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	26	4.33	78.02	23.08	100	37.5	88	4.63	85.37	35.23	77.5	48.44
Fastalign	8	1.33	95.6	62.5	83.33	71.43	27	1.42	95.34	85.19	57.5	68.66
HMM Giza++	22	3.67	80.22	22.73	83.33	35.71	79	4.16	88.69	43.04	85	57.14
IBM-4 Giza++	15	2.5	87.91	33.33	83.33	47.62	72	3.79	90.24	47.22	85	60.71

Table A.8: Baselines for English-German: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction German-English and in the direction English-German

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	1961	4.25	85.54	15.96	56.09	24.85	3365	2.86	90.68	23.6	46.06	31.2
Fastalign	700	1.52	95.94	51.86	65.05	57.71	1489	1.27	95.84	55.41	47.85	51.35
HMM Giza++	1623	3.52	89.42	24.52	71.33	36.5	2878	2.45	93.61	38.26	63.86	47.85
IBM-4 Giza++	1468	3.18	90.83	28.13	74.01	40.77	2430	2.07	95	46.79	65.95	54.74

Table A.9: Baselines for English-Czech: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Czech-English and in the direction English-Czech

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	82	6.31	78.74	19.51	76.19	31.07	137	2.49	89.77	31.39	48.86	38.22
Fastalign	17	1.31	97.01	82.35	66.67	73.68	57	1.04	95.36	71.93	46.59	56.55
HMM Giza++	53	4.08	87.43	30.19	76.19	43.24	182	3.31	88.96	32.97	68.18	44.44
IBM-4 Giza++	50	3.85	89.52	36	85.71	50.7	167	3.04	89.92	35.33	67.05	46.27

Table A.10: Baselines for English-Romanian: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Romanian-English and in the direction English-Romanian

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	642	3.75	92.43	31.62	65.48	42.65	412	4.12	89.35	16.99	51.47	25.55
Fastalign	231	1.35	96.8	67.1	50	57.3	100	1	96.45	50	36.76	42.37
HMM Giza++	639	3.74	92.58	32.39	66.77	43.62	442	4.42	89.03	17.87	58.09	27.34
IBM-4 Giza++	667	3.9	92.74	34.03	73.23	46.47	425	4.25	89.48	18.59	58.09	28.16

Table A.11: Baselines for English-Japanese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Japanese-English and in the direction English-Japanese

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	3054	3.39	91.52	32.65	56.94	41.5	1147	2.76	90.04	23.63	64.68	34.61
Fastalign	1280	1.42	96.32	70.78	51.74	59.78	488	1.18	97.11	62.5	72.79	67.25
HMM Giza++	2853	3.16	93.37	42.13	68.65	52.22	1090	2.63	91.1	27.25	70.88	39.36
IBM-4 Giza++	2586	2.87	94.71	49.92	73.73	59.53	956	2.3	92.75	32.95	75.18	45.82

Table A.12: Baselines for English-Vietnamese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Vietnamese-English and in the direction English-Vietnamese

A.8 Baselines: Unknown words

Baselines	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
Concatenation												
IBM-1 Giza++	680	4.33	79.06	22.65	52.38	31.62	298	4.66	78.04	21.14	62.38	31.58
Fastalign	269	1.71	91.86	56.51	51.7	54	82	1.28	93.81	64.63	52.48	57.92
HMM Giza++	432	2.75	88.05	40.05	58.84	47.66	226	3.53	83.67	27.43	61.39	37.92
IBM-4 Giza++	388	2.47	88.81	42.01	55.44	47.8	194	3.03	86.4	32.47	62.38	42.71
Replacing unknown words with the token UNK												
IBM-1 Giza++	443	2.82	83.87	25.28	38.1	30.39	141	2.2	87.77	31.91	44.55	37.19
Fastalign	251	1.6	91.1	52.19	44.56	48.07	77	1.2	93.4	62.34	47.52	53.93
HMM Giza++	267	1.7	91.8	56.18	51.02	53.48	105	1.64	91.79	49.52	51.49	50.49
IBM-4 Giza++	214	1.36	92.46	62.62	45.58	52.76	110	1.72	91.07	45.45	49.5	47.39

Table A.13: Baselines for English-French: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction French-English and in the direction English-French.

Baselines	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
	Concatenation											
IBM-1 Giza++	112	7.47	73.74	13.39	68.18	22.39	393	6.78	79.13	25.45	77.52	38.31
Fastalign	15	1	94.19	46.67	31.82	37.84	48	0.83	92.81	68.75	25.58	37.29
HMM Giza++	88	5.87	81.82	21.59	86.36	34.55	323	5.57	84.83	33.75	84.5	48.23
IBM-4 Giza++	76	5.07	84.85	25	86.36	38.78	270	4.66	88.66	41.48	86.82	56.14
	Replacing unknown words with the token UNK											
IBM-1 Giza++	72	4.8	83.33	19.44	63.64	29.79	95	1.64	91.19	46.32	34.11	39.29
Fastalign	20	1.33	93.94	45	40.91	42.86	37	0.64	91.96	56.76	16.28	25.3
HMM Giza++	55	3.67	88.64	29.09	72.73	41.56	127	2.19	93.26	59.84	58.91	59.38
IBM-4 Giza++	57	3.8	88.64	29.82	77.27	43.04	118	2.03	92.29	54.24	49.61	51.82

Table A.14: Baselines for English-German: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction German-English and in the direction English-German.

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
	Concatenation											
IBM-1 Giza++	6931	4.33	85.4	16.87	55.53	25.87	8487	3.33	89.86	20.9	48.91	29.29
Fastalign	2118	1.32	96.29	59.54	59.9	59.72	3056	1.2	96.22	57.04	48.06	52.16
HMM Giza++	5702	3.57	89.75	27.24	73.78	39.78	7488	2.94	92.59	32.41	66.91	43.67
IBM-4 Giza++	5132	3.21	91.11	30.79	75.06	43.66	6058	2.38	94.39	40.82	68.18	51.07
	Replacing unknown words with the token UNK											
IBM-1 Giza++	2124	1.33	93.18	25.94	26.18	26.06	2077	0.82	94.5	25.37	14.53	18.48
Fastalign	2076	1.3	95.2	47.69	47.03	47.36	2820	1.11	95.39	45.25	35.18	39.58
HMM Giza++	1854	1.16	95.37	49.51	43.61	46.38	1869	0.73	95.88	53.93	27.79	36.68
IBM-4 Giza++	1977	1.24	95.1	46.38	43.56	44.93	1839	0.72	95.71	50.19	25.45	33.77

Table A.15: Baselines for English-Czech: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Czech-English and in the direction English-Czech.

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
	Concatenation											
IBM-1 Giza++	238	6.61	78.68	13.45	65.31	22.3	263	4.24	86.77	21.29	58.33	31.2
Fastalign	36	1	96.46	66.67	48.98	56.47	61	0.98	95.45	59.02	37.5	45.86
HMM Giza++	189	5.25	83.94	18.52	71.43	29.41	242	3.9	89.93	30.99	78.12	44.38
IBM-4 Giza++	159	4.42	87.76	25.16	81.63	38.46	214	3.45	91.22	34.11	76.04	47.1
	Replacing unknown words with the token UNK											
IBM-1 Giza++	57	1.58	93.5	33.33	38.78	35.85	26	0.42	94.64	42.31	11.46	18.03
Fastalign	37	1.03	95.98	59.46	44.9	51.16	44	0.71	95.39	61.36	28.12	38.57
HMM Giza++	62	1.72	95.51	51.61	65.31	57.66	56	0.9	95.61	62.5	36.46	46.05
IBM-4 Giza++	51	1.42	95.41	50.98	53.06	52	46	0.74	94.86	50	23.96	32.39

Table A.16: Baselines for English-Romanian: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Romanian-English and in the direction English-Romanian.

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
	Concatenation											
IBM-1 Giza++	2649	4.73	89.29	25.78	57.44	35.59	1192	4.97	88.47	15.44	58.04	24.39
Fastalign	684	1.22	95.6	62.72	36.08	45.81	167	0.7	96.97	55.09	29.02	38.02
HMM Giza++	2610	4.66	89.95	28.35	62.24	38.96	1121	4.67	89.49	17.75	62.78	27.68
IBM-4 Giza++	2502	4.47	90.46	29.74	62.57	40.31	1081	4.5	90.06	19.15	65.3	29.61
	Replacing unknown words with the token UNK											
IBM-1 Giza++	242	0.43	94.57	36.78	7.49	12.44	202	0.84	96.11	33.17	21.14	25.82
Fastalign	501	0.89	95.05	54.69	23.04	32.43	275	1.15	96.06	36.73	31.86	34.12
HMM Giza++	369	0.66	94.94	52.85	16.4	25.03	211	0.88	96.34	39.34	26.18	31.44
IBM-4 Giza++	316	0.56	94.85	50	13.29	21	187	0.78	96.52	42.78	25.24	31.75

Table A.17: Baselines for English-Japanese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Japanese-English and in the direction English-Japanese.

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
	Concatenation											
IBM-1 Giza++	14589	3	92.22	23.11	56.59	32.82	7181	2.55	90.93	17.1	64.56	27.04
Fastalign	5808	1.2	96.62	49.62	48.36	48.98	3141	1.11	96.89	44.09	72.82	54.93
HMM Giza++	13469	2.77	93.79	31.23	70.58	43.3	6888	2.44	91.85	20.6	74.61	32.29
IBM-4 Giza++	11283	2.32	95.07	37.69	71.35	49.32	5977	2.12	93.41	25.63	80.55	38.89
	Replacing unknown words with the token UNK											
IBM-1 Giza++	4525	0.93	95.29	23.56	17.89	20.34	3824	1.36	94.42	21.6	43.43	28.85
Fastalign	5500	1.13	96.01	39.91	36.84	38.31	3430	1.22	96.1	36.15	65.19	46.51
HMM Giza++	3814	0.79	96.26	41.27	26.41	32.21	3464	1.23	96.06	35.94	65.46	46.4
IBM-4 Giza++	3730	0.77	96.24	40.43	25.31	31.13	3444	1.22	96.07	35.92	65.04	46.28

Table A.18: Baselines for English-Vietnamese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Vietnamese-English and in the direction English-Vietnamese.

A.9 Baselines: Part-of-Speech

A.9.1 Baselines: Alignment links for content and function words

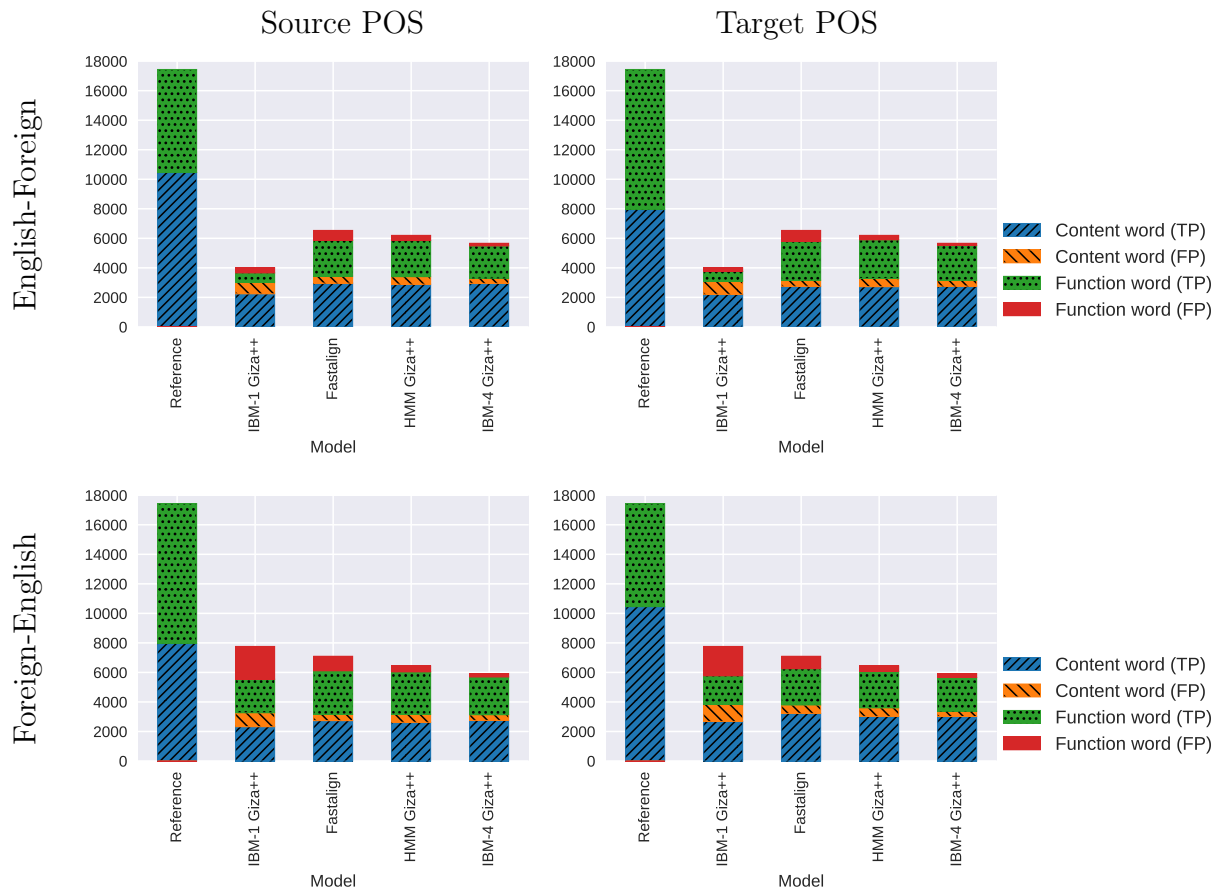


Figure A.55: Baselines for English-French: The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).

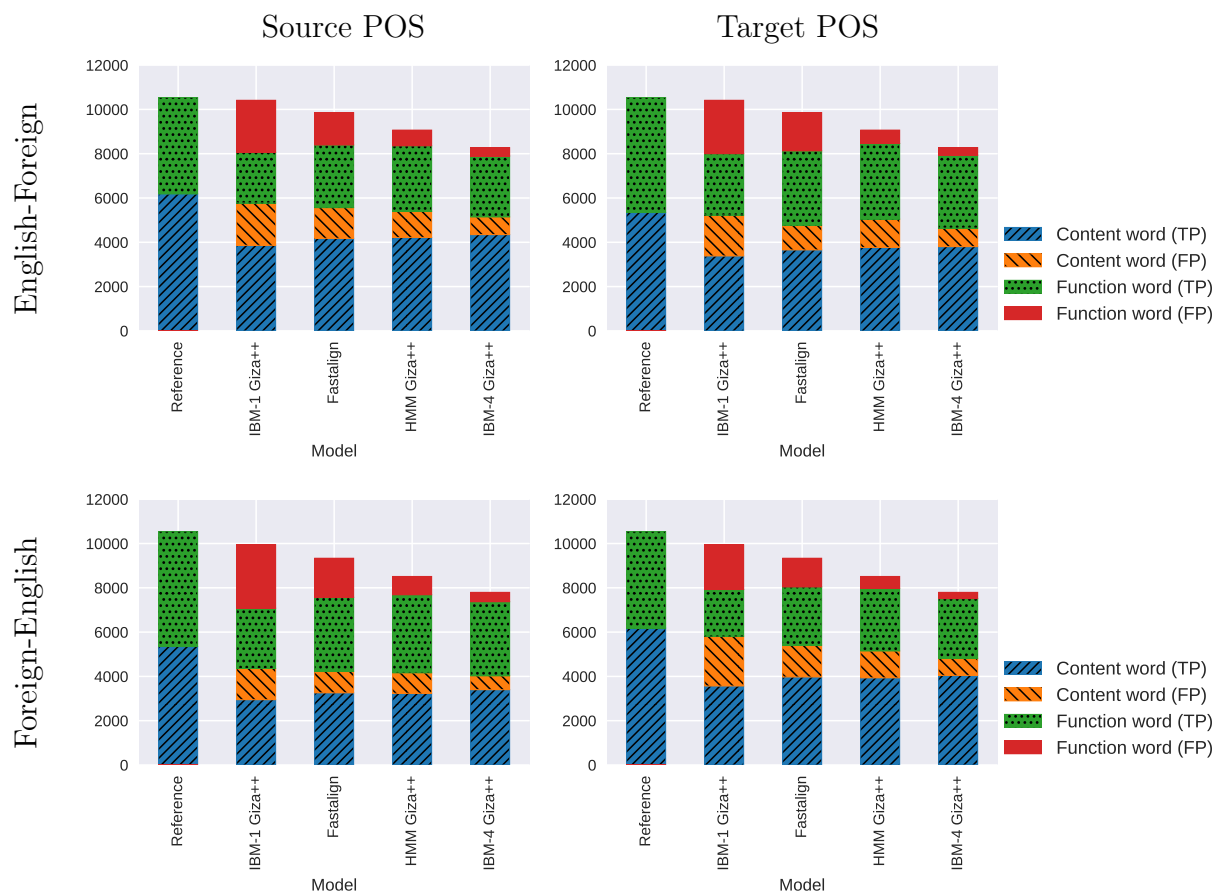


Figure A.56: Baselines for English-German: Baselines for English-French: The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).

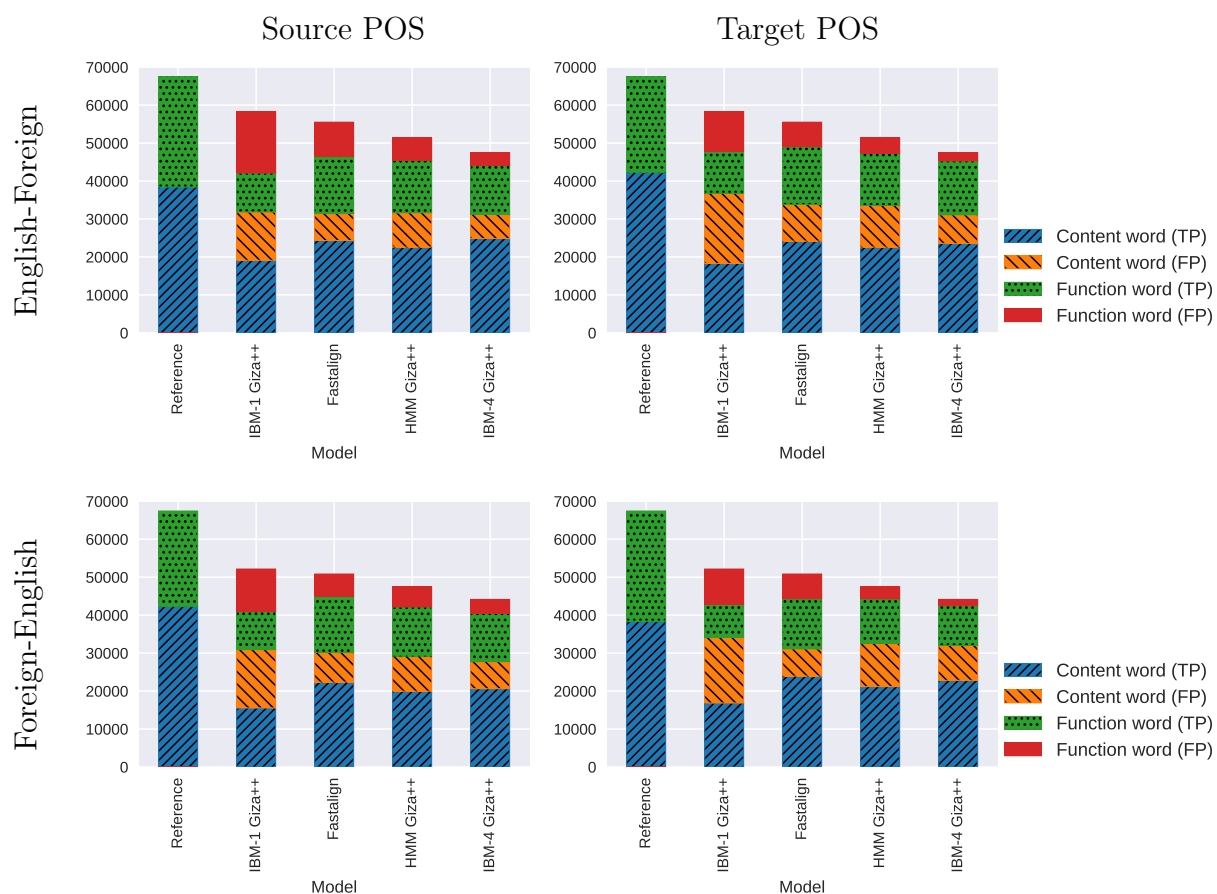


Figure A.57: Baselines for English-Czech: Baselines for English-French: The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).

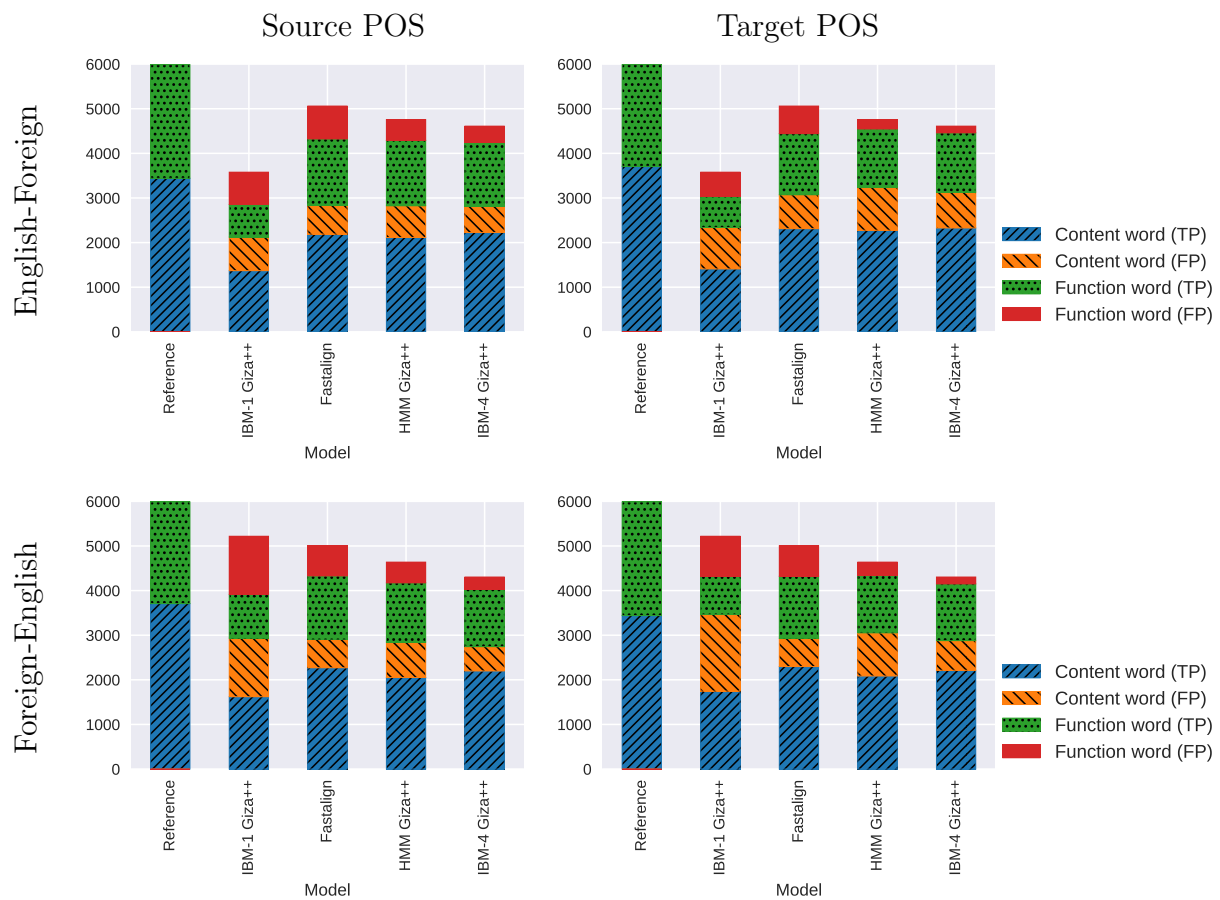


Figure A.58: Baselines for English-Romanian: Baselines for English-French: The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).

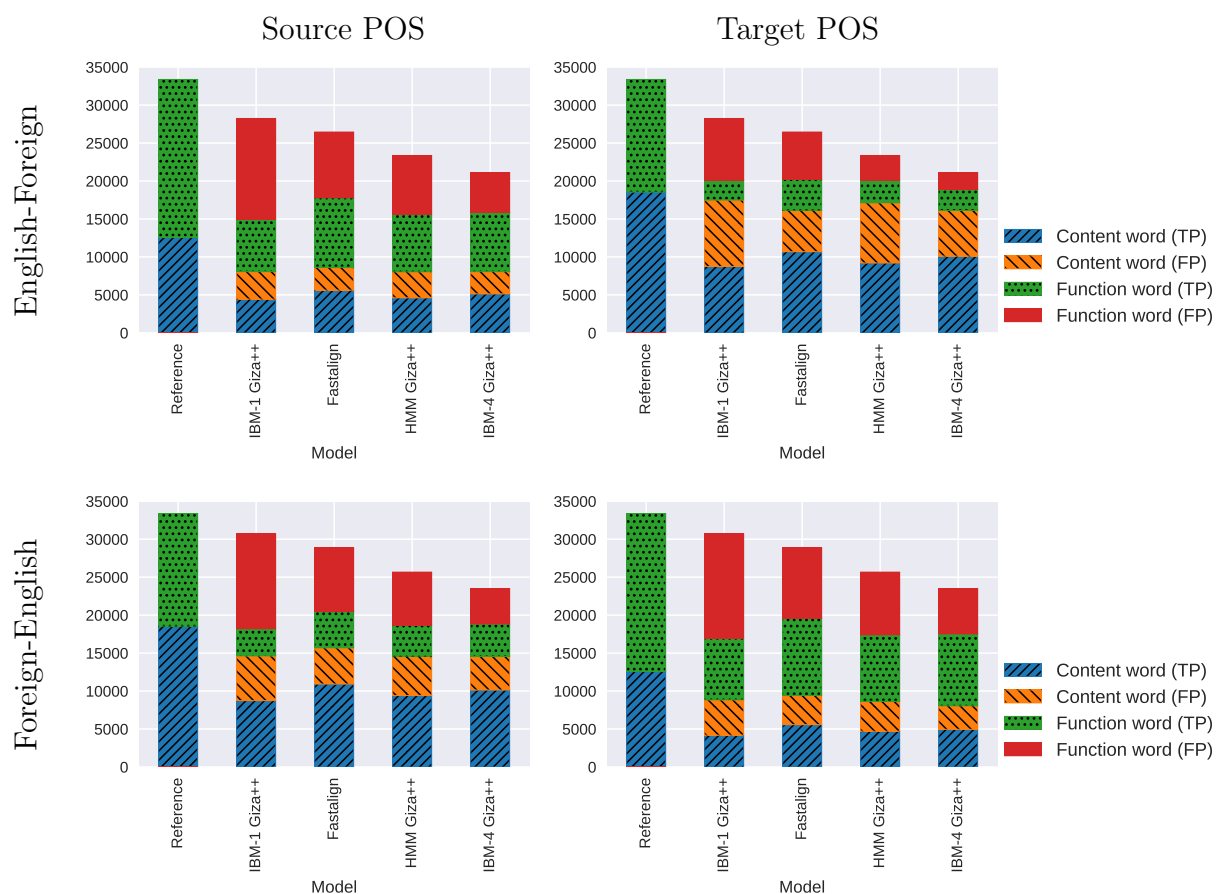


Figure A.59: Baselines for English-Japanese: Baselines for English-French: The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).

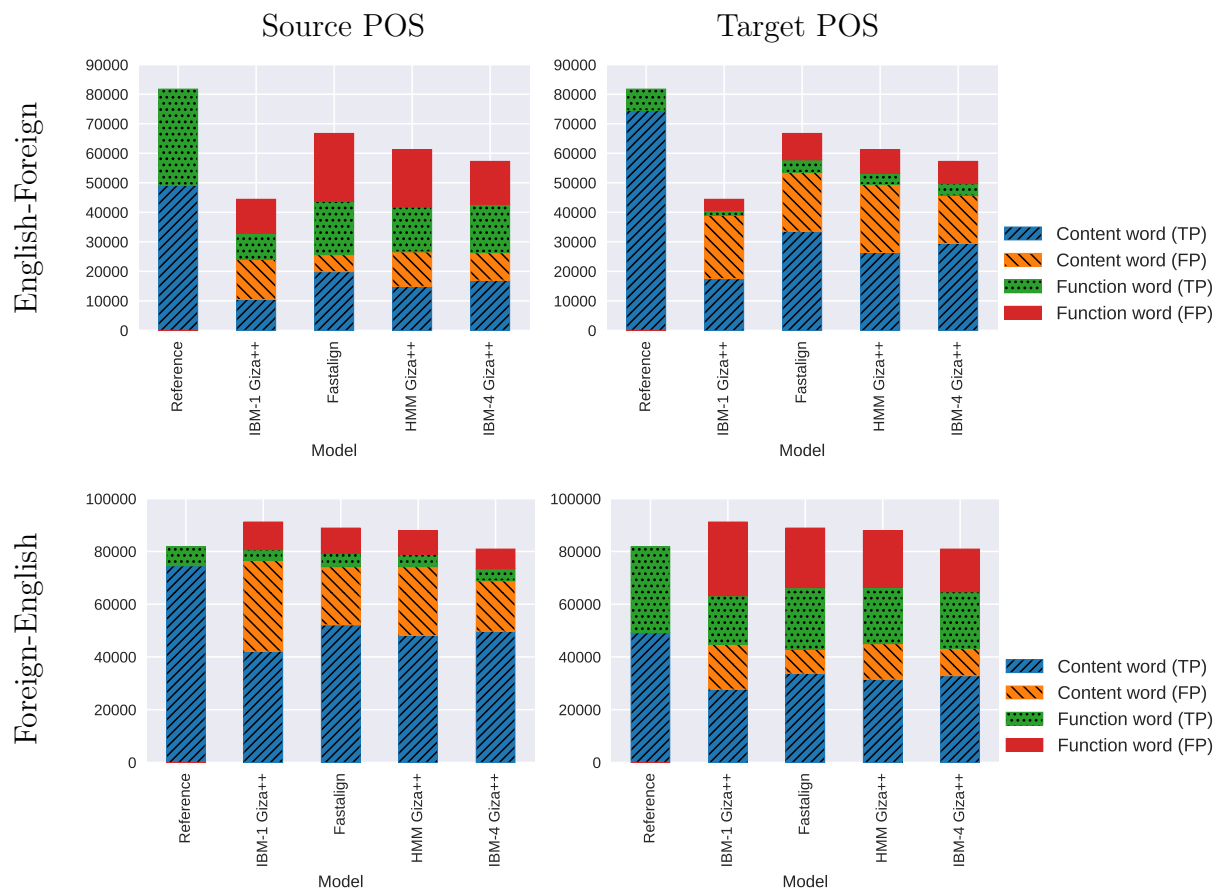


Figure A.60: Baselines for English-Vietnamese: Baselines for English-French: The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).

A.9.2 Baselines: Unaligned words for content and function words

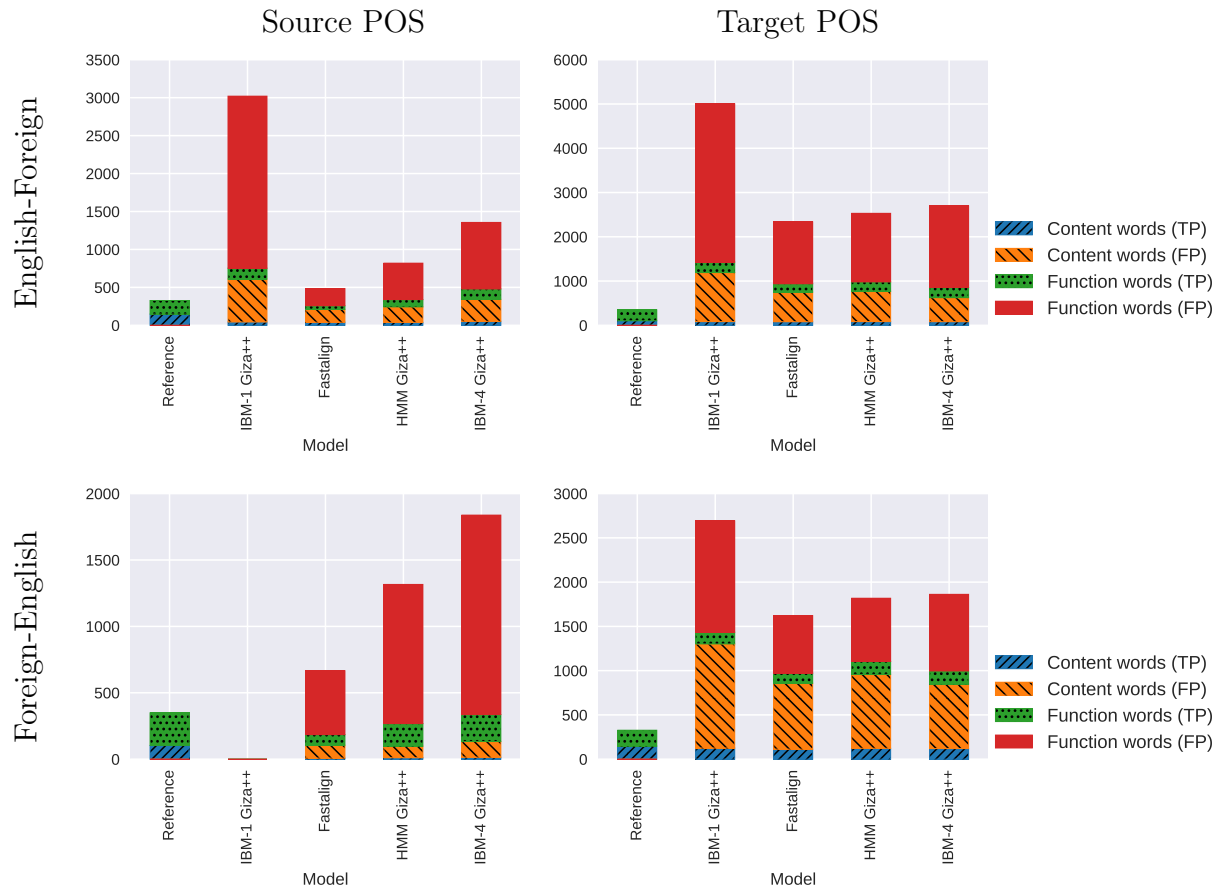


Figure A.61: Baselines for English-French: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).

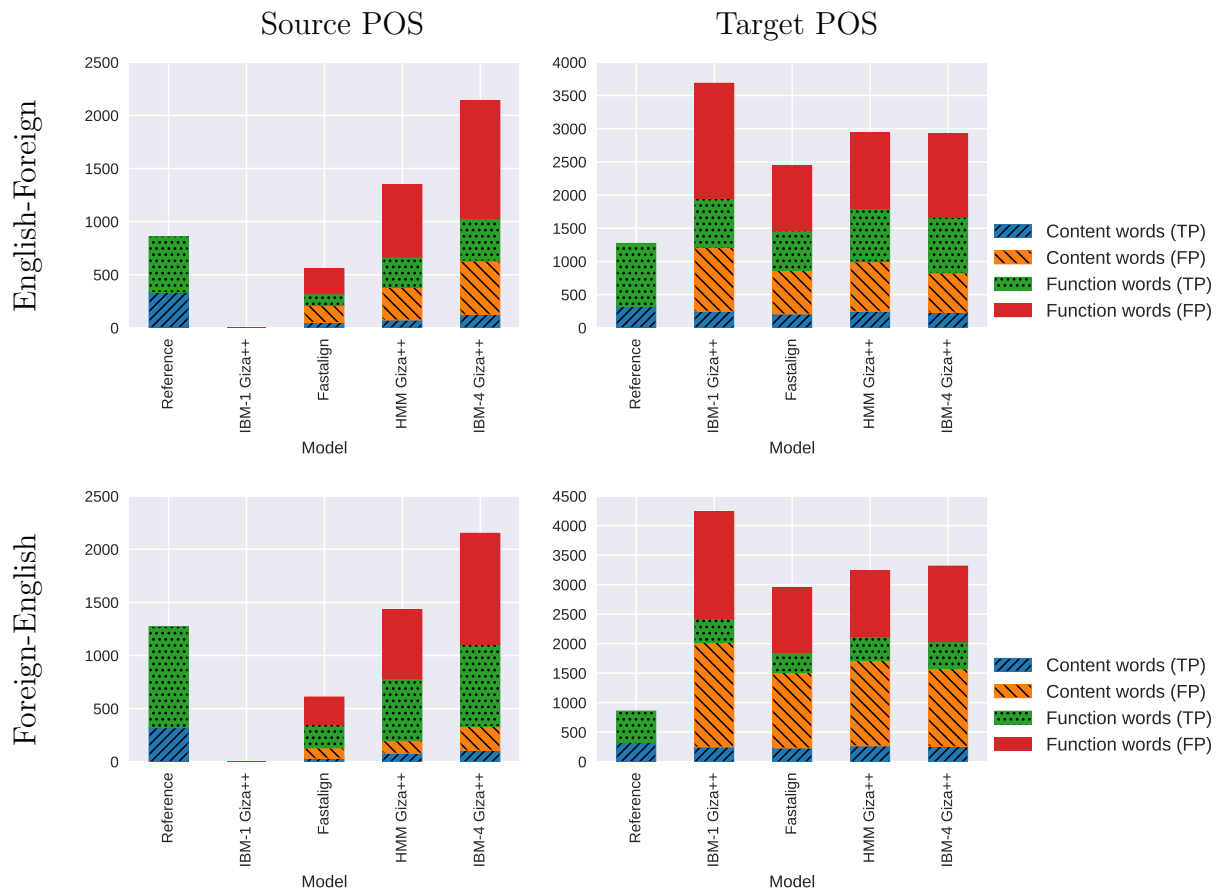


Figure A.62: Baselines for English-German: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).

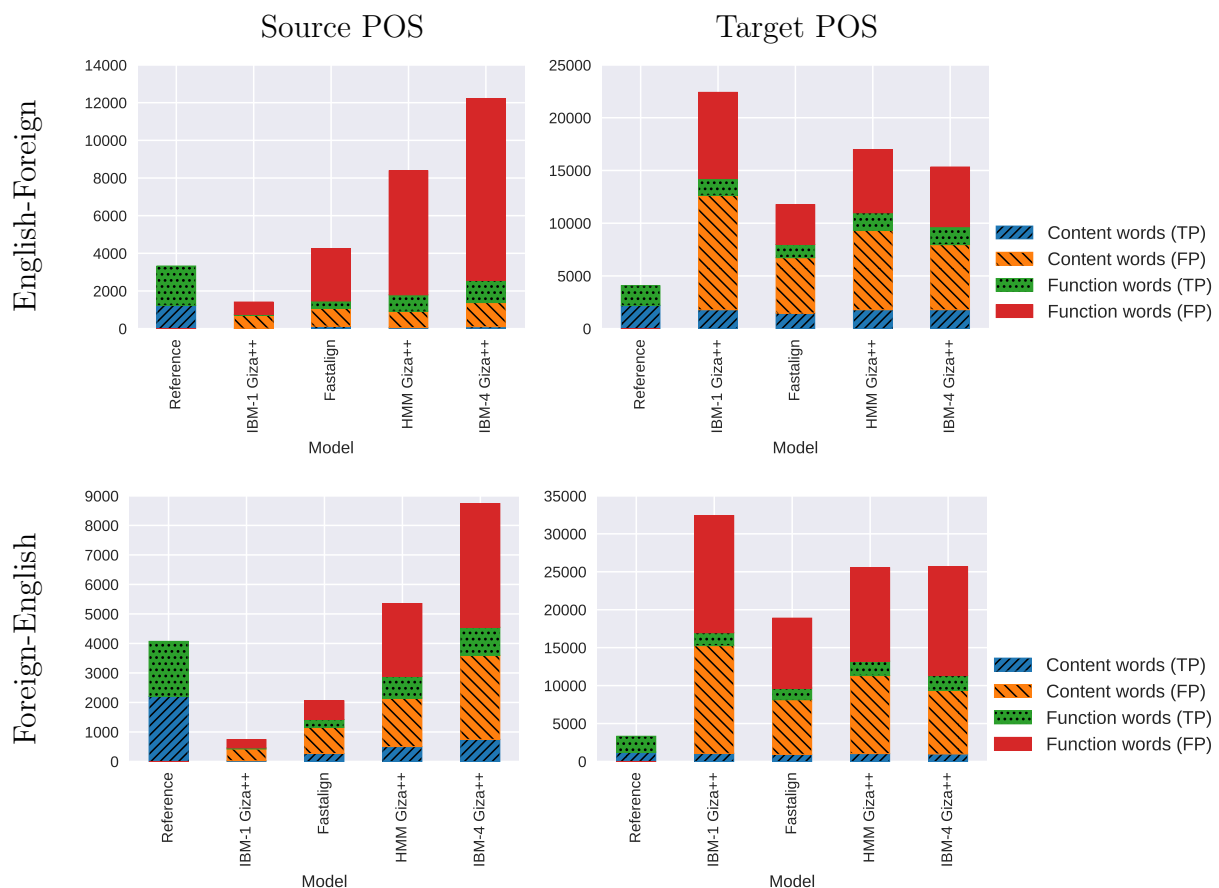


Figure A.63: Baselines for English-Czech: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).

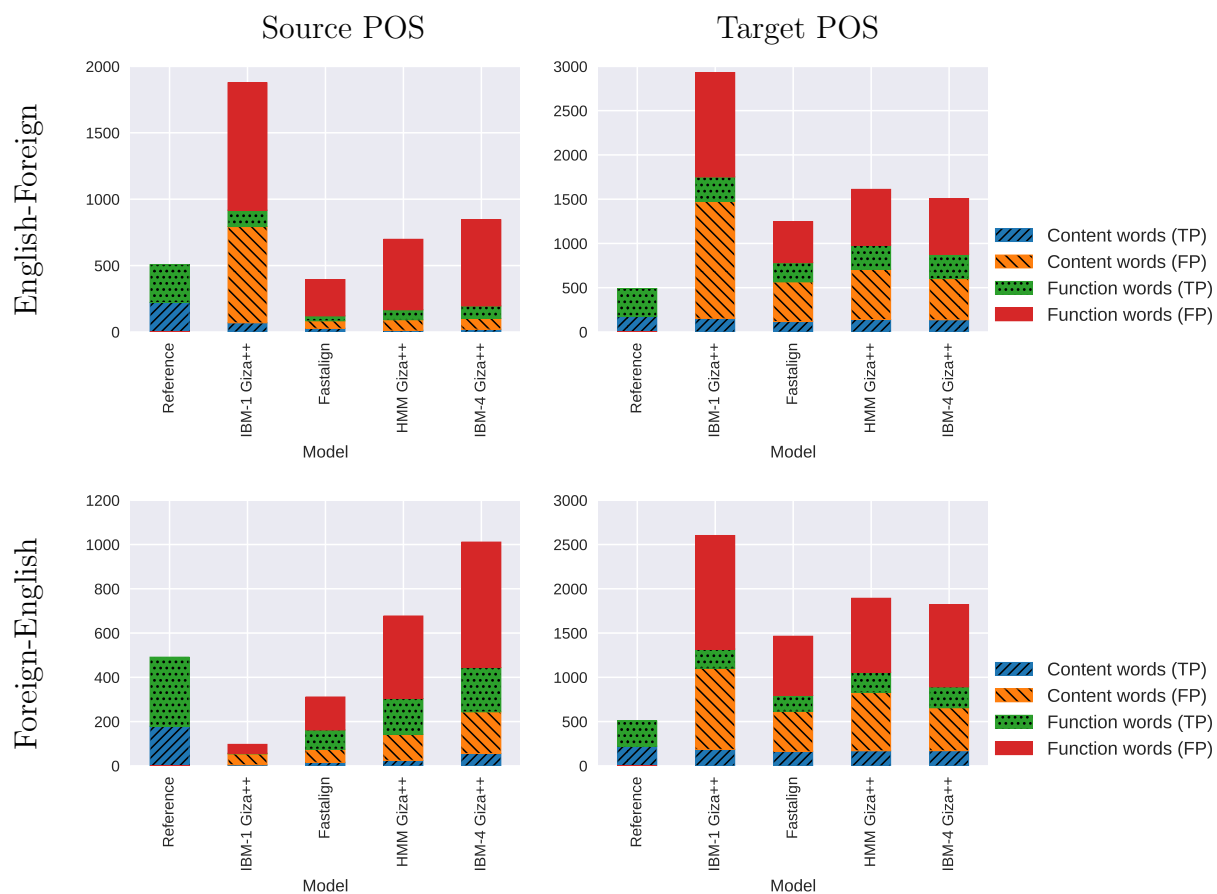


Figure A.64: Baselines for English-Romanian: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).

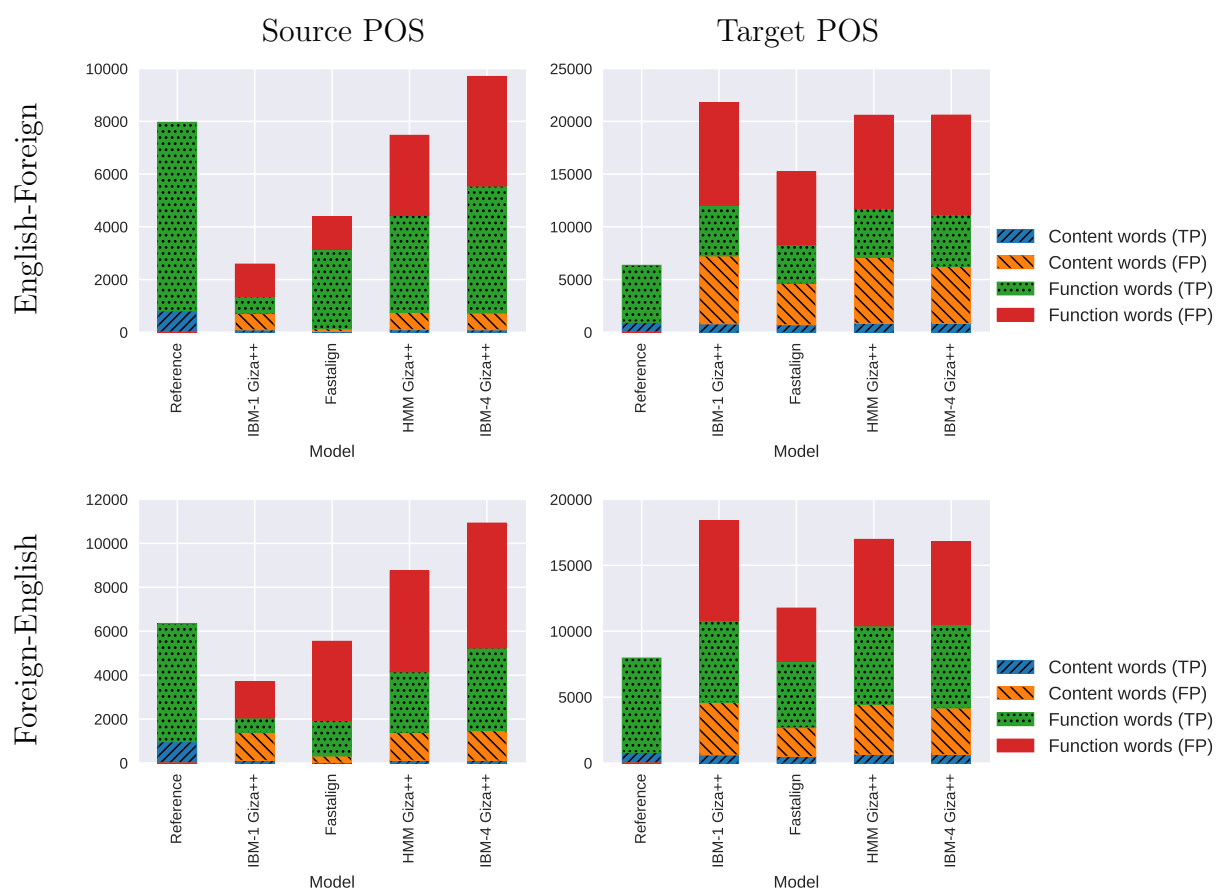


Figure A.65: Baselines for English-Japanese: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).

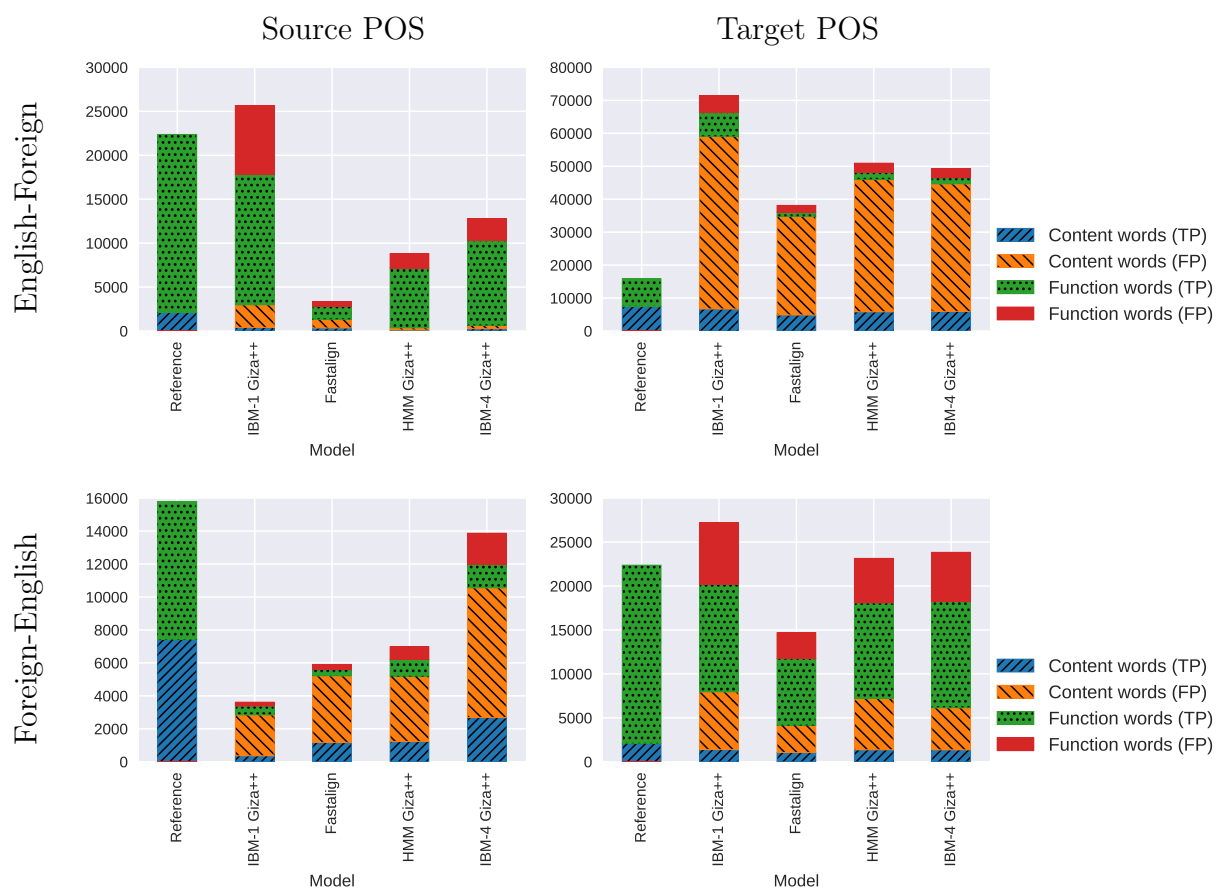


Figure A.66: Baselines for English-Vietnamese: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).

A.10 Baselines: Symmetrization

A.10.1 Baselines: Intersection

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
IBM-1 Giza++	2441	0.31	0.61	36.9	23.04	93.81	13.13	89.3	0.43	0.55
Fastalign	4879	0.69	0.75	11.09	40.48	92.58	25.9	90.71	0.7	0.73
HMM Giza++	4683	0.73	0.76	7.59	41.16	97.2	26.1	90.9	0.65	0.82
IBM-4 Giza++	4588	0.77	0.81	7.76	40.88	98.13	25.82	90.89	0.65	0.64

Table A.19: Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-French

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
IBM-1 Giza++	5531	0.56	0.53	37.11	59.5	86.4	45.37	97.29	0.68	1
Fastalign	6746	0.72	0.68	28.41	68.31	87.49	56.03	97.72	0.84	1
HMM Giza++	6265	0.74	0.69	25.46	70.91	95.07	56.55	97.97	0.84	0.83
IBM-4 Giza++	6413	0.82	0.78	23.82	72.55	95.85	58.36	98.06	0.89	0.83

Table A.20: Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-German

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
IBM-1 Giza++	23298	0.45	0.4	40.22	45.12	87.85	30.36	97	0.39	0.55
Fastalign	36091	0.71	0.65	20.68	63.06	90.43	48.41	97.7	0.63	0.77
HMM Giza++	28415	0.6	0.55	25.65	57.2	96.46	40.65	97.53	0.51	0.68
IBM-4 Giza++	30648	0.69	0.65	21.43	60.84	97.33	44.24	97.69	0.58	0.73

Table A.21: Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Czech

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
IBM-1 Giza++	1811	0.35	0.51	59.32	40.7	87.63	26.5	97.12	0.49	0.69
Fastalign	3659	0.73	0.72	32.85	67.17	88.55	54.11	98.03	0.67	0.92
HMM Giza++	3119	0.67	0.66	35.7	64.32	93.91	48.91	97.98	0.6	0.92
IBM-4 Giza++	3315	0.77	0.72	32.19	67.83	95.17	52.69	98.14	0.69	0.85

Table A.22: Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Romanian

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
IBM-1 Giza++	9540	0.31	0.34	64.7	35.3	79.4	22.7	98.01	0.58	0.57
Fastalign	15622	0.54	0.59	51.23	48.77	76.49	35.8	98.21	0.53	0.67
HMM Giza++	9372	0.37	0.4	61.12	38.88	88.67	24.9	98.13	0.53	0.57
IBM-4 Giza++	10137	0.43	0.48	56.29	43.71	93.82	28.5	98.25	0.56	0.63

Table A.23: Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Japanese

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
IBM-1 Giza++	20603	0.23	0.46	67.18	32.82	81.52	20.55	97.26	0.62	0.39
Fastalign	48975	0.55	0.73	46.21	53.79	71.78	43.01	97.59	0.78	0.71
HMM Giza++	35598	0.41	0.58	56.09	43.92	72.38	31.52	97.38	0.67	0.47
IBM-4 Giza++	38925	0.48	0.68	51.18	48.82	75.67	36.04	97.54	0.72	0.55

Table A.24: Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Vietnamese

A.10.2 Baselines: Grow-diag-final (GDF)

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
BM-1 Giza++	40.1	26.7	33.9	36.49	25.19	33.83	82.75	21.26
Fastalign	15.19	44.98	16.23	46.32	14.25	49.05	81.84	35.02
HMM Giza++	11.99	45.18	11.97	45.98	8.54	47.5	90.78	32.17
IBM-4 Giza++	10	44.43	9.64	45.43	7.03	46.32	93.55	30.78

Table A.25: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-French

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
BM-1 Giza++	39.03	58.76	42.66	55.39	32.31	65.06	72.19	59.2
Fastalign	28.98	68.75	31.28	66.47	28.21	69.84	70.12	69.57
HMM Giza++	23.92	73.3	26.33	71.04	22.87	74.48	81.74	68.39
IBM-4 Giza++	21.46	75.48	23.31	73.63	19.59	77.29	89.26	68.16

Table A.26: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-German

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
BM-1 Giza++	45.09	46.75	48.47	42.88	35.47	52.67	71.16	41.81
Fastalign	25.75	64.09	25.3	62.86	23.3	66.93	72.95	61.82
HMM Giza++	27.86	61.22	30.38	57.28	25.25	62.96	75.67	53.91
IBM-4 Giza++	20.92	65.7	26.5	59.81	19.13	66.67	84.22	55.17

Table A.27: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Czech

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
BM-1 Giza++	56.02	43.99	53.52	46.49	51.13	48.88	73.82	36.54
Fastalign	33.36	66.65	32.91	67.1	30.42	69.6	72.65	66.8
HMM Giza++	33.36	66.65	36.38	63.64	32.4	67.62	76.63	60.5
IBM-4 Giza++	31.04	68.98	32.3	67.72	29.08	70.94	82.98	61.96

Table A.28: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Romanian

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
BM-1 Giza++	63.12	36.88	61.55	38.45	55.31	44.69	64	34.33
Fastalign	50.67	49.33	49.37	50.63	44.79	55.21	64.99	47.99
HMM Giza++	57.01	42.99	54.41	45.59	52.18	47.82	62.37	38.78
IBM-4 Giza++	52.52	47.48	49.23	50.77	46.68	53.32	74.58	41.49

Table A.29: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Japanese

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
BM-1 Giza++	69.43	30.58	46.45	53.56	47.91	52.09	70.78	41.21
Fastalign	48.89	51.11	32.82	67.19	32.9	67.11	64.41	70.05
HMM Giza++	57.86	42.14	37.57	62.43	39.79	60.21	61.51	58.97
IBM-4 Giza++	51.91	48.1	33.19	66.82	33.29	66.71	70.25	63.51

Table A.30: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Vietnamese

A.11 Word-based sentence length statistics

A.11.1 Sentence length statistics

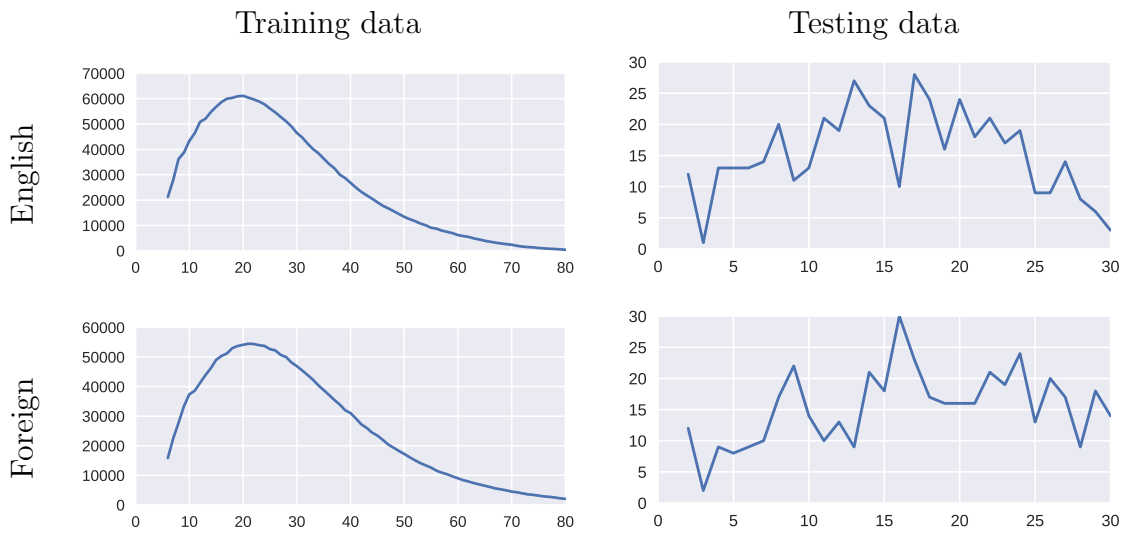


Figure A.67: Sentence lengths in English-French



Figure A.68: Sentence lengths in English-German

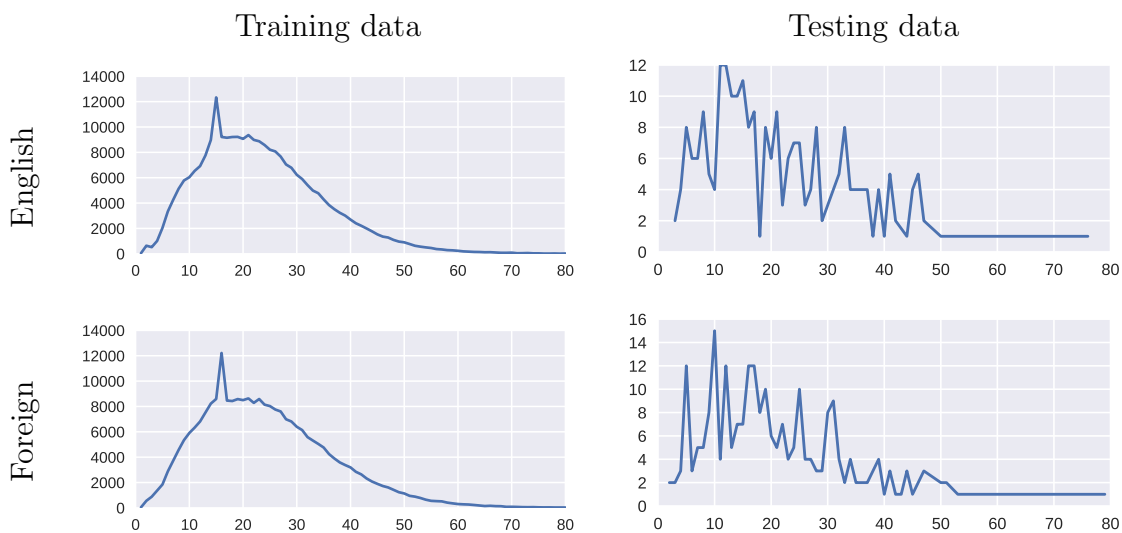


Figure A.69: Sentence lengths in English-Romanian

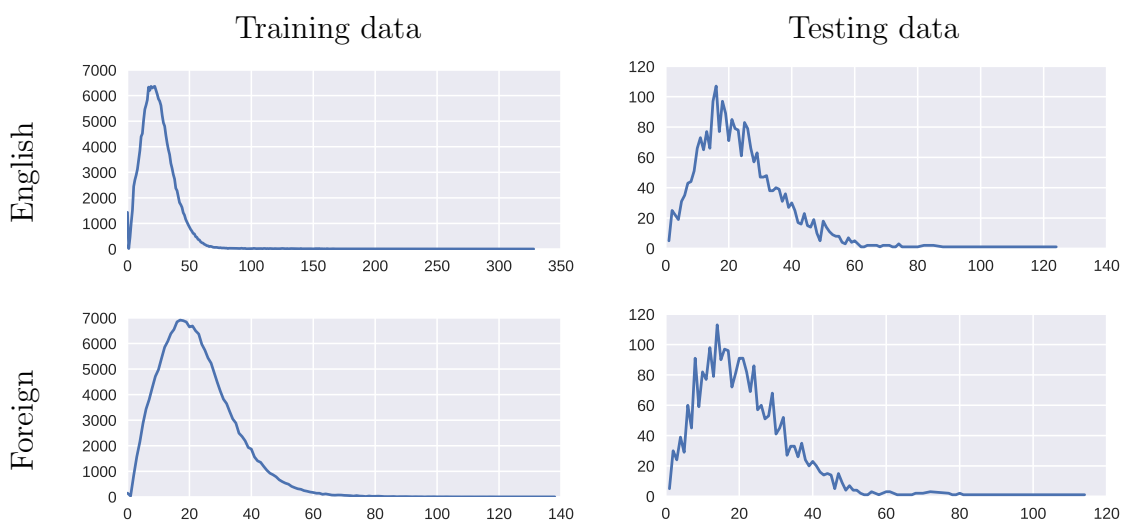


Figure A.70: Sentence lengths in English-Czech

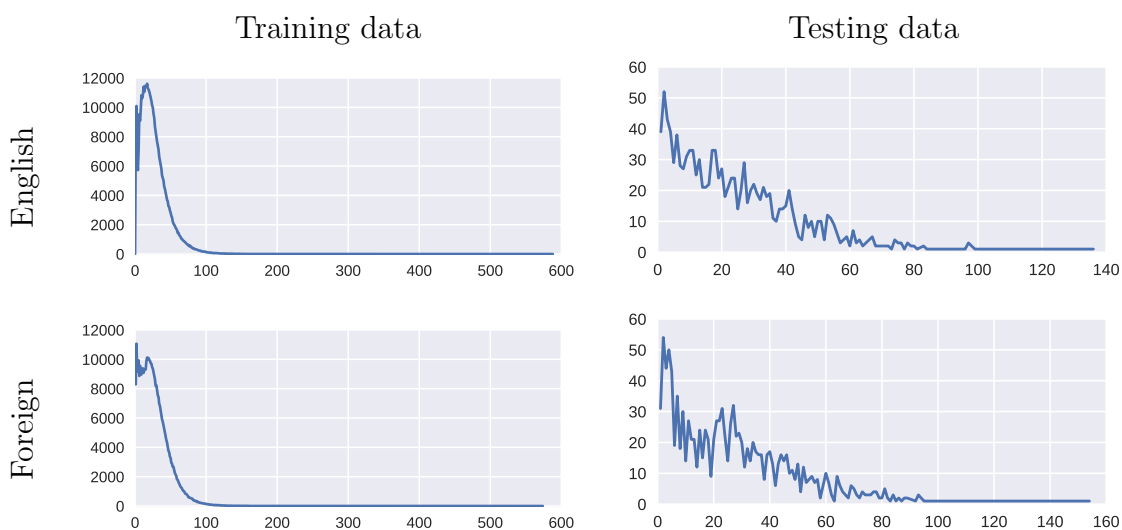


Figure A.71: Sentence lengths in English-Japanese

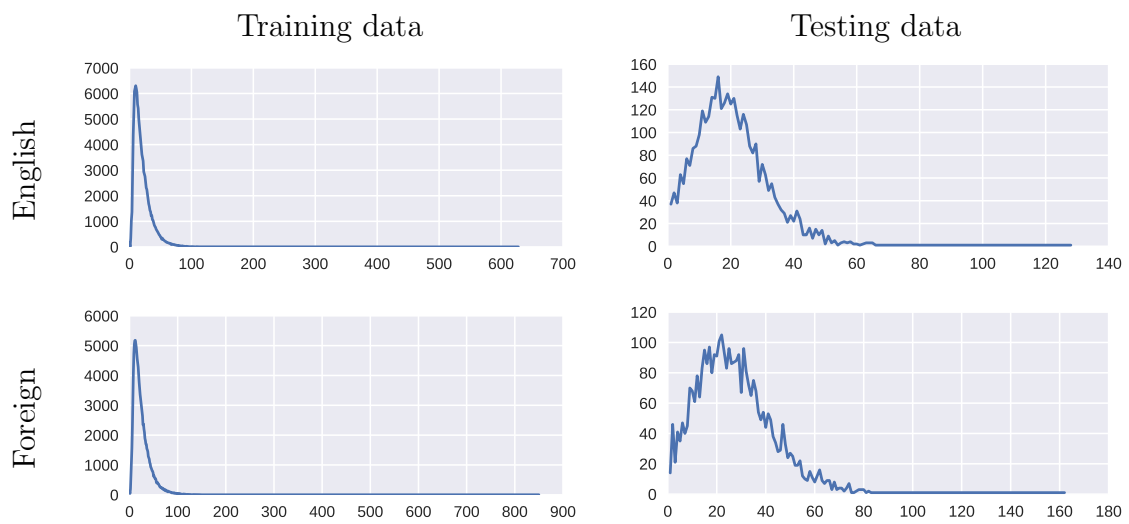


Figure A.72: Sentence lengths in English-Vietnamese

A.11.2 Length difference statistics between English and foreign sentence

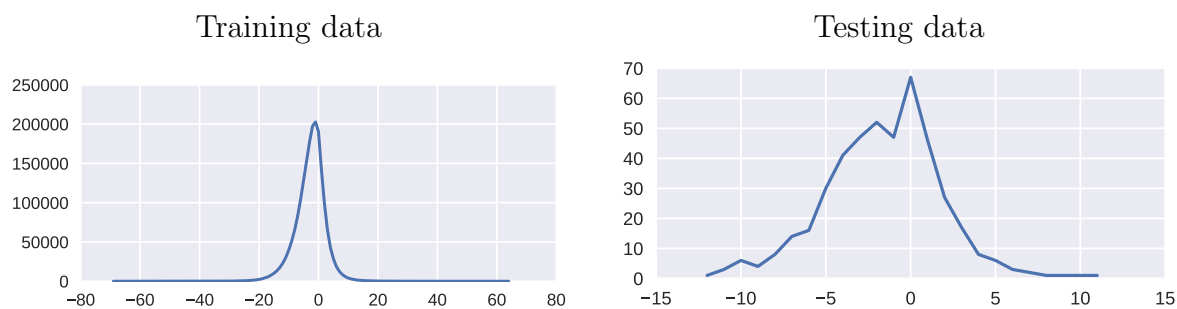


Figure A.73: Length differences in English-French

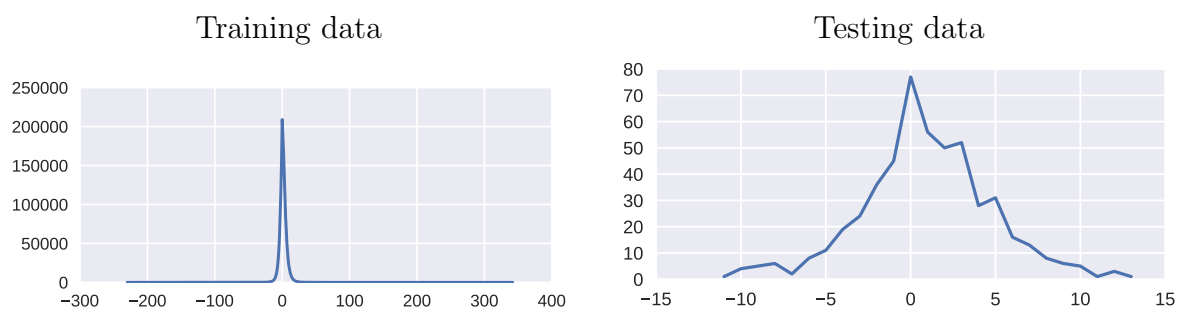


Figure A.74: Length differences in English-German

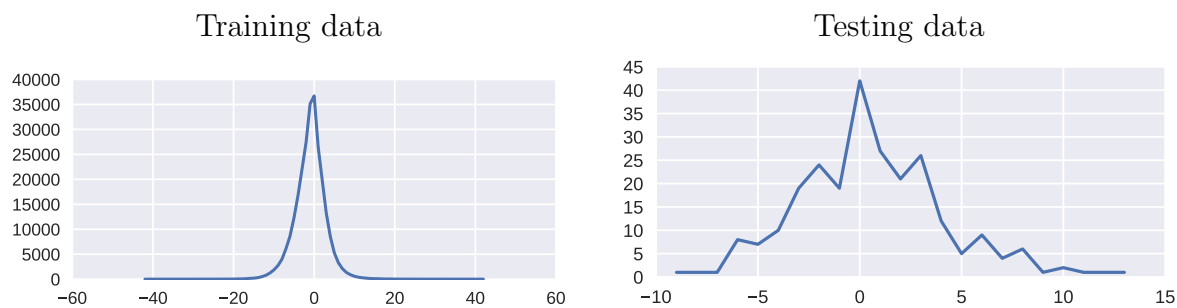


Figure A.75: Length differences in English-Romanian

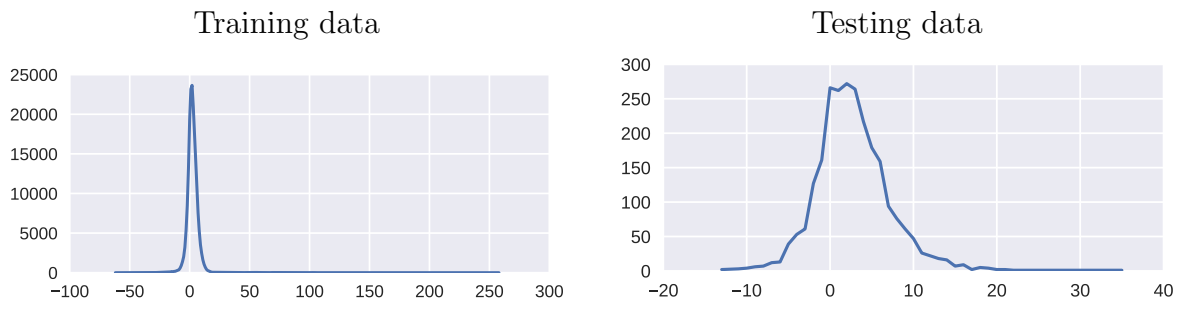


Figure A.76: Length differences in English-Czech

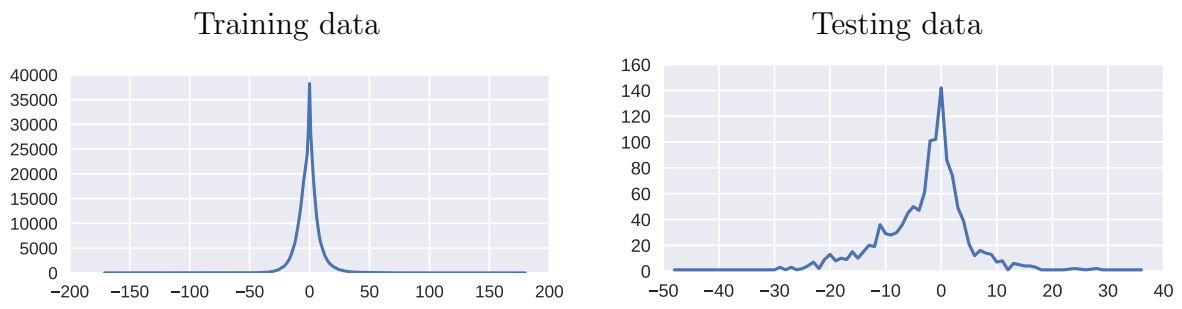


Figure A.77: Length differences in English-Japanese

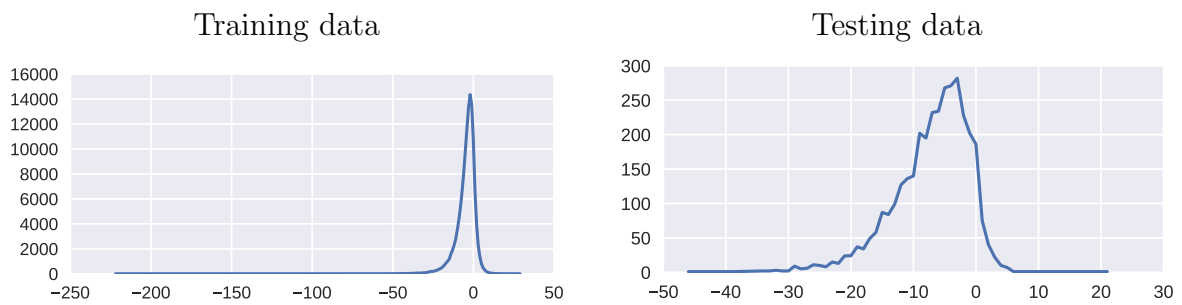


Figure A.78: Length differences in English-Vietnamese

A.11.3 Sentence length and unknown/rare words

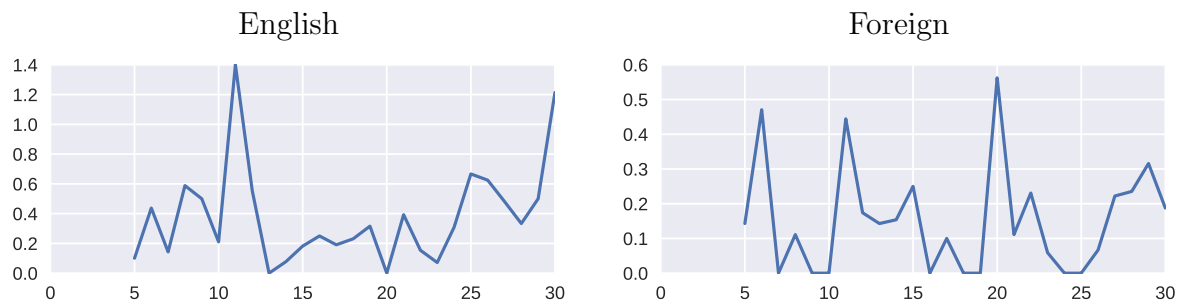


Figure A.79: Number of unknown/rare words as a function of sentence length for English-French

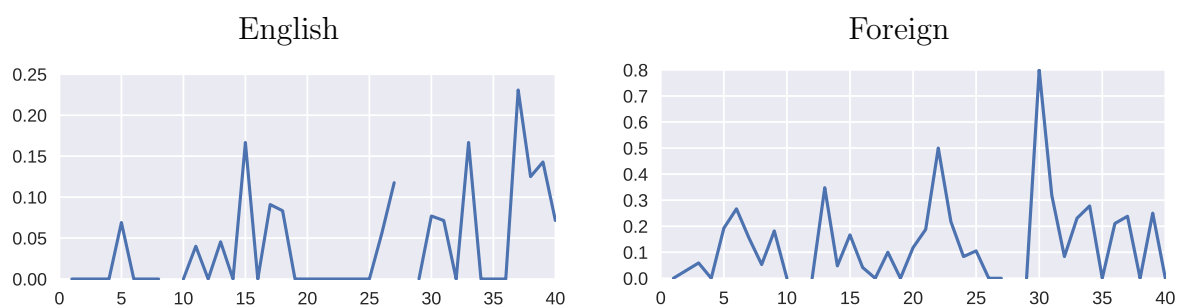


Figure A.80: Number of unknown/rare words as a function of sentence length for English-German

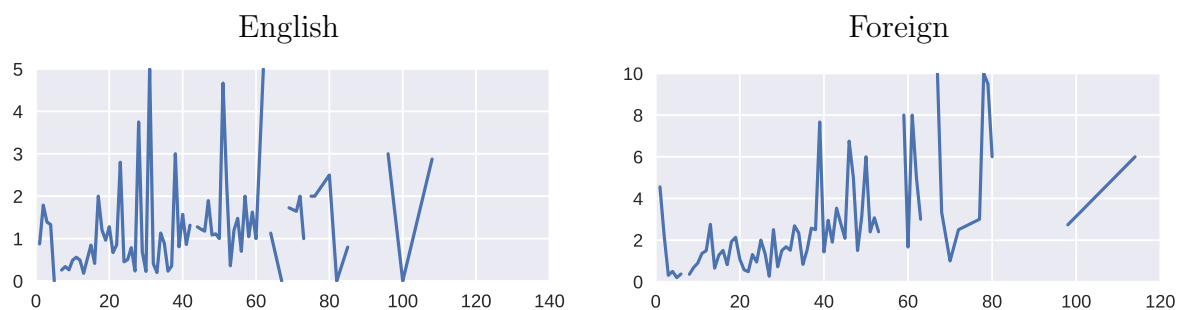


Figure A.81: Number of unknown/rare words as a function of sentence length for English-Czech

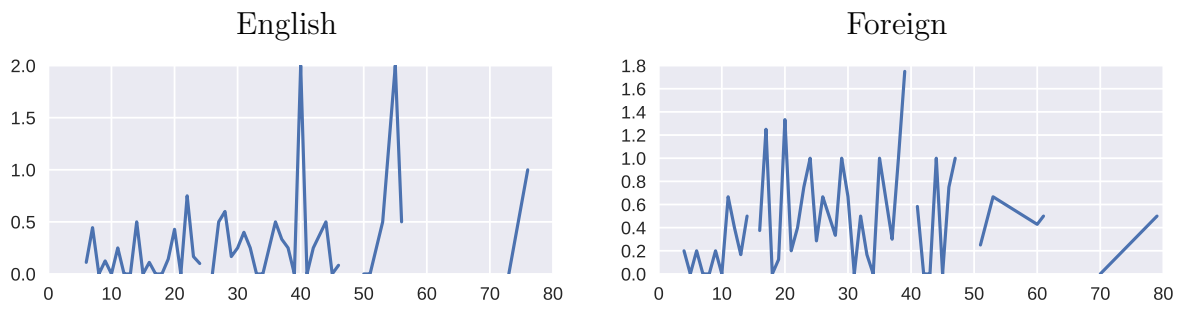


Figure A.82: Number of unknown/rare words as a function of sentence length for English-Romanian

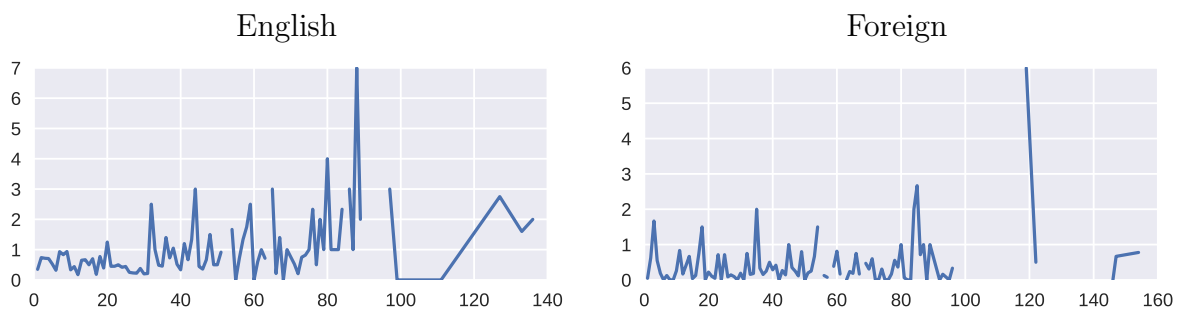


Figure A.83: Number of unknown/rare words as a function of sentence length for English-Japanese

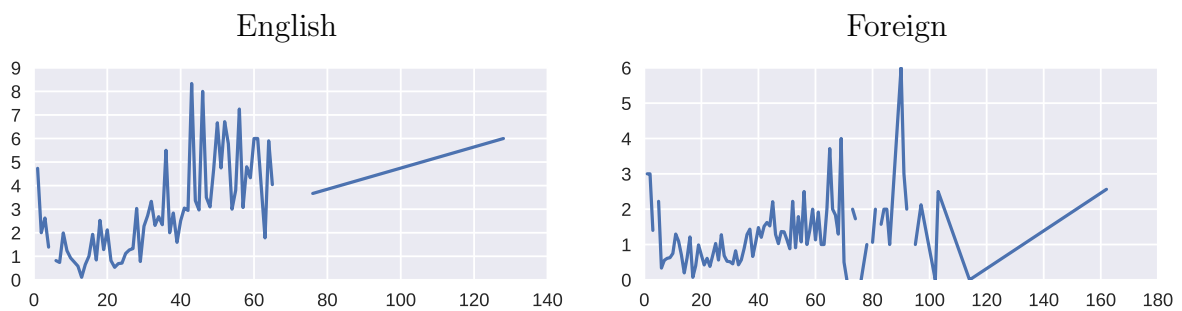


Figure A.84: Number of unknown/rare words as a function of sentence length for English-Vietnamese

A.11.4 Sentence length and word repetition

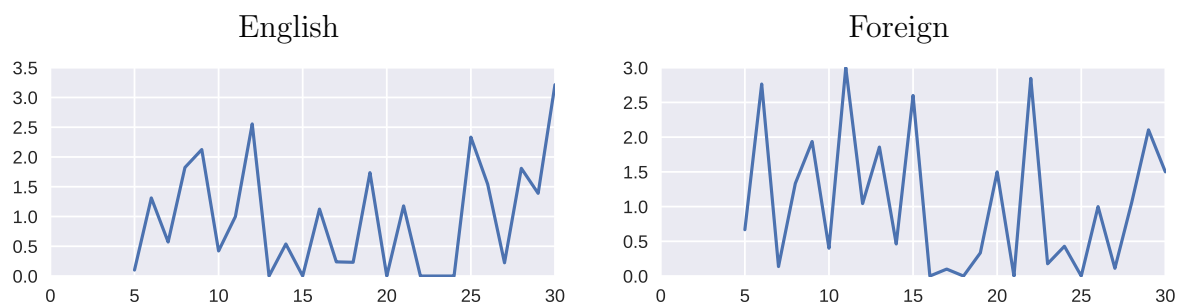


Figure A.85: Number of words that repeat at least twice as a function of sentence length for English-French

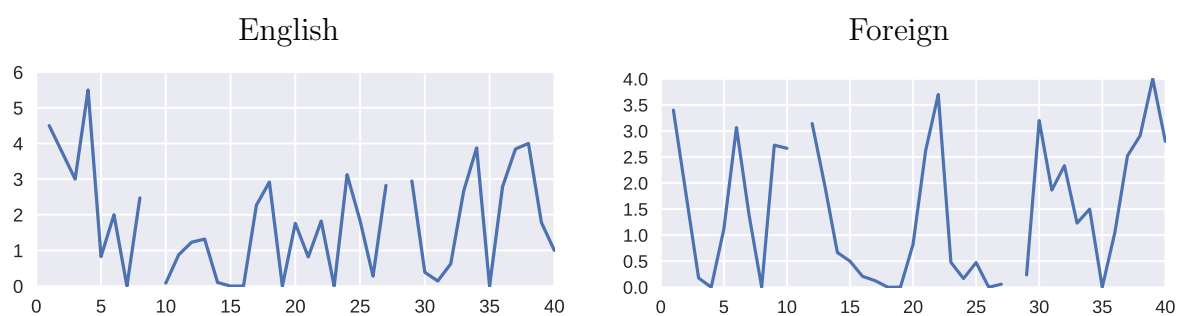


Figure A.86: Number of words that repeat at least twice as a function of sentence length for English-German

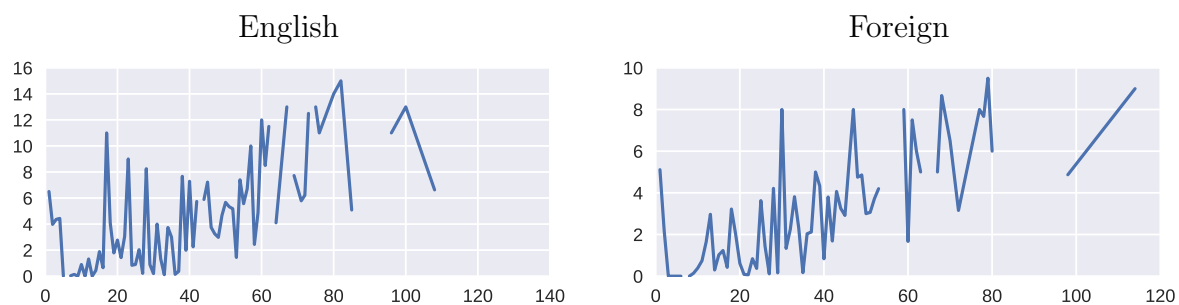


Figure A.87: Number of words that repeat at least twice as a function of sentence length for English-Czech

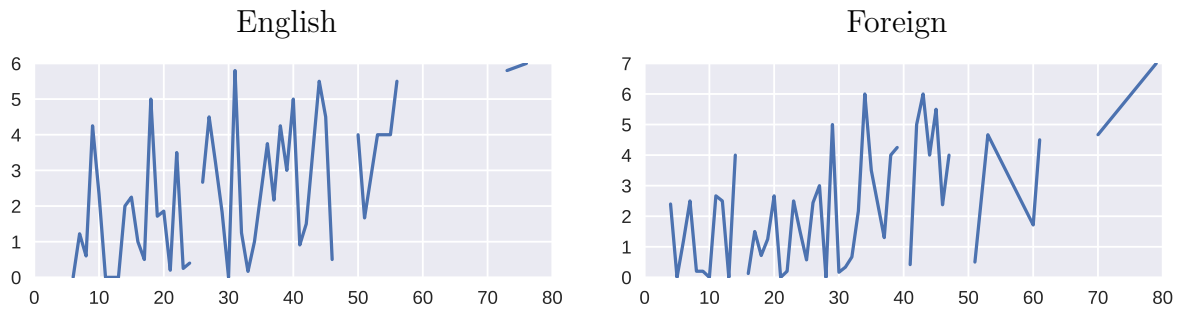


Figure A.88: Number of words that repeat at least twice as a function of sentence length for English-Romanian

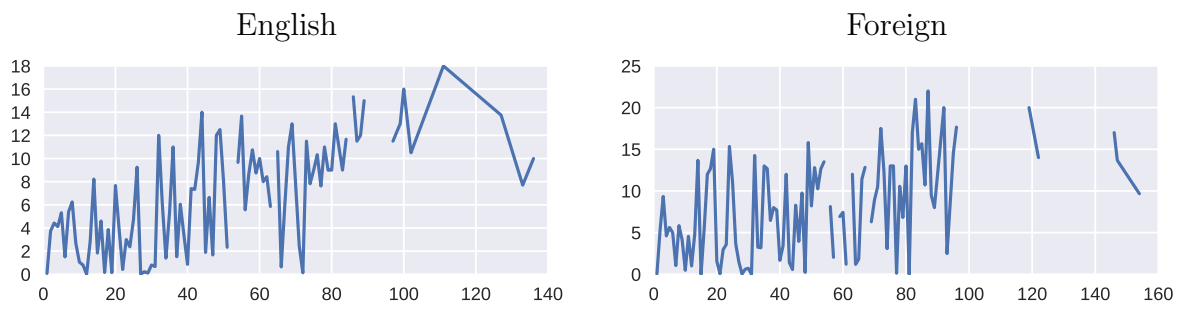


Figure A.89: Number of words that repeat at least twice as a function of sentence length for English-Japanese

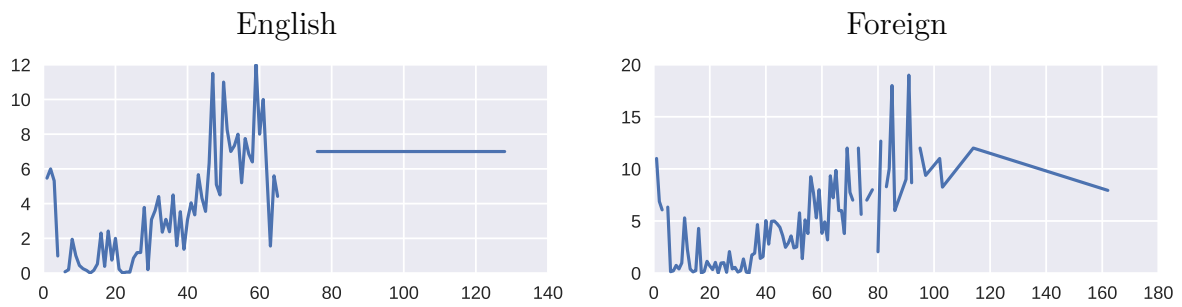


Figure A.90: Number of words that repeat at least twice as a function of sentence length for English-Vietnamese

A.11.5 Baselines: AER as a function of sentence length

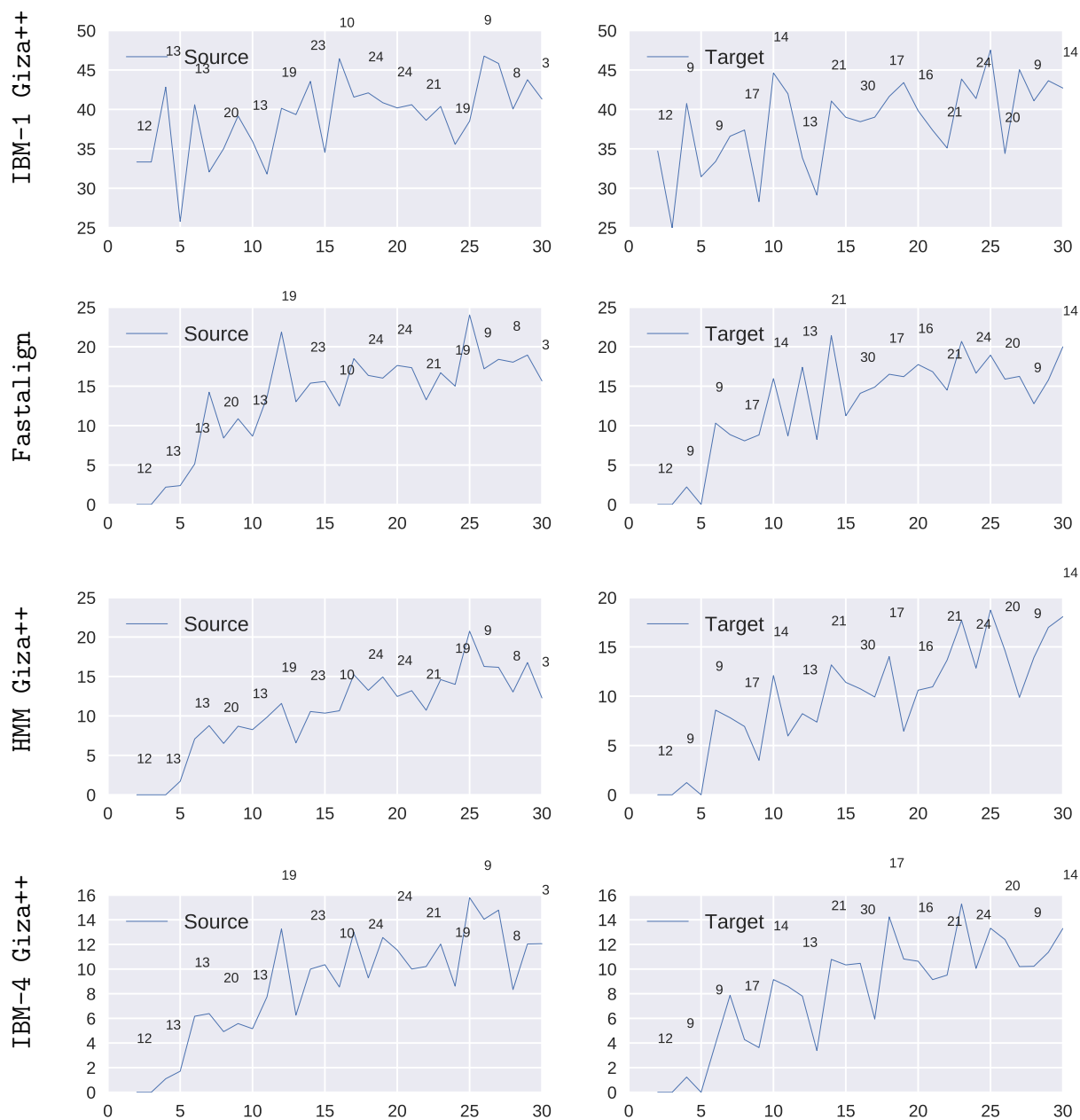


Figure A.91: Baselines for the direction English-French: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.

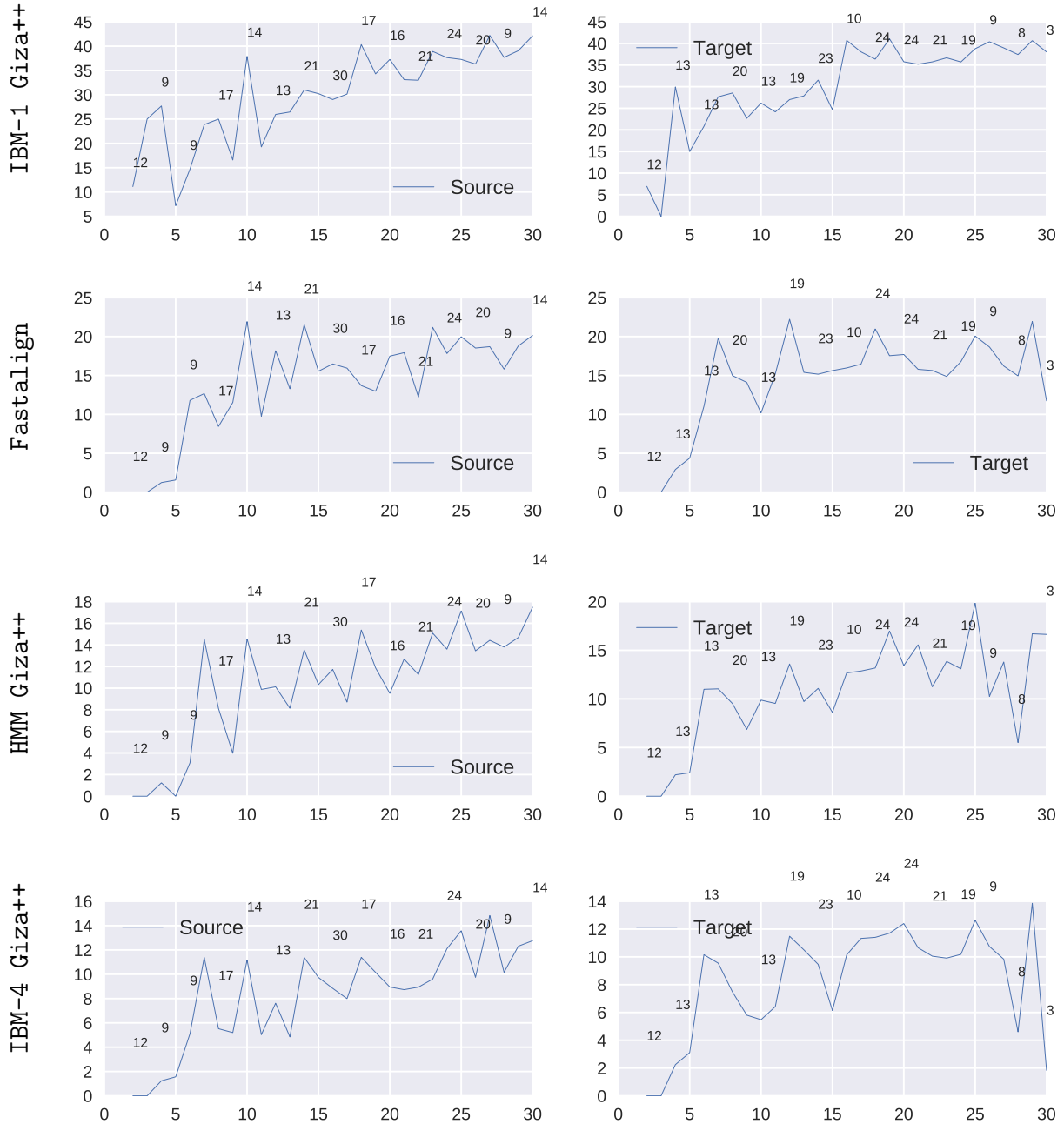


Figure A.92: Baselines for the direction French-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.

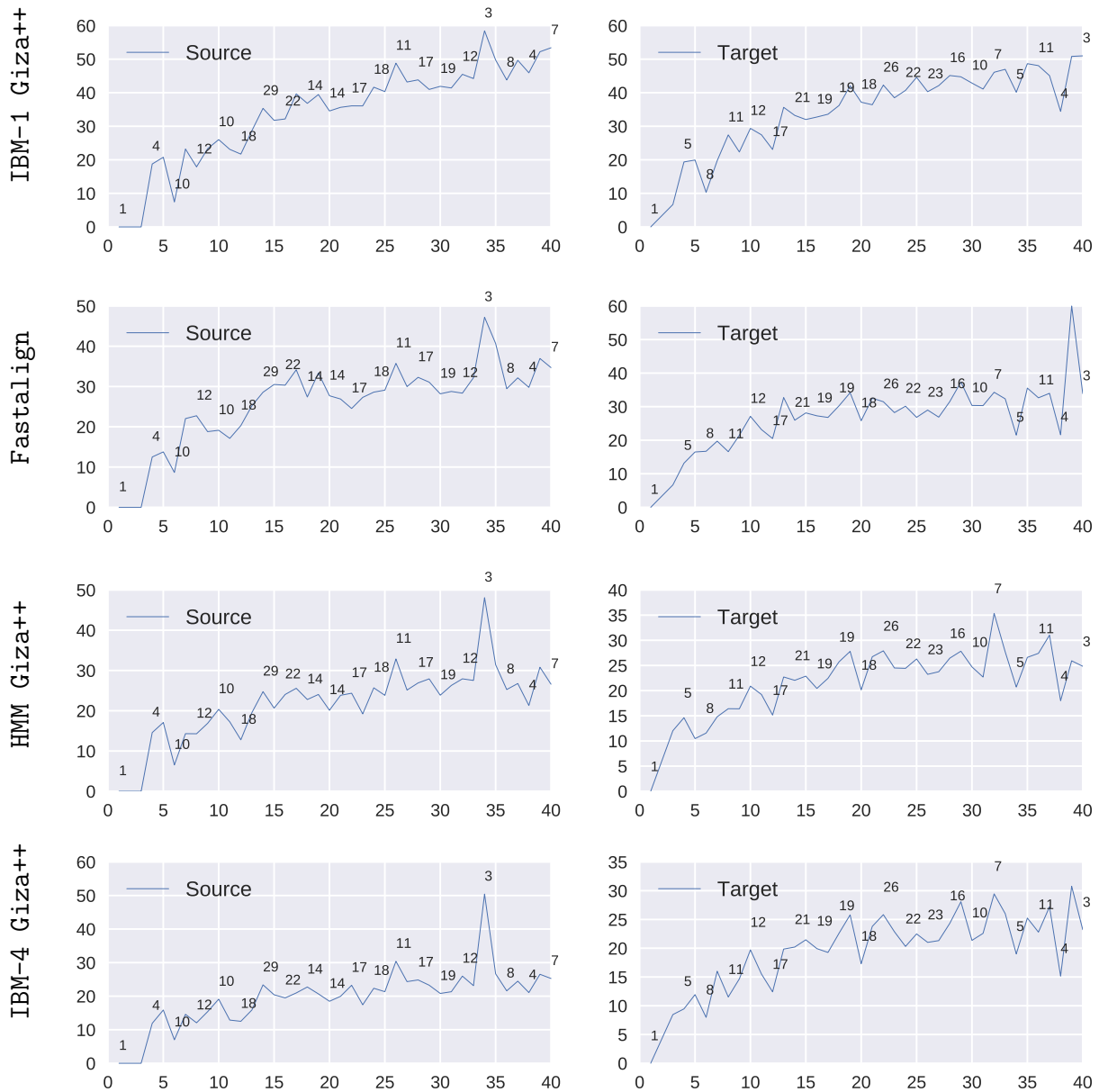


Figure A.93: Baselines for the direction English-German: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.

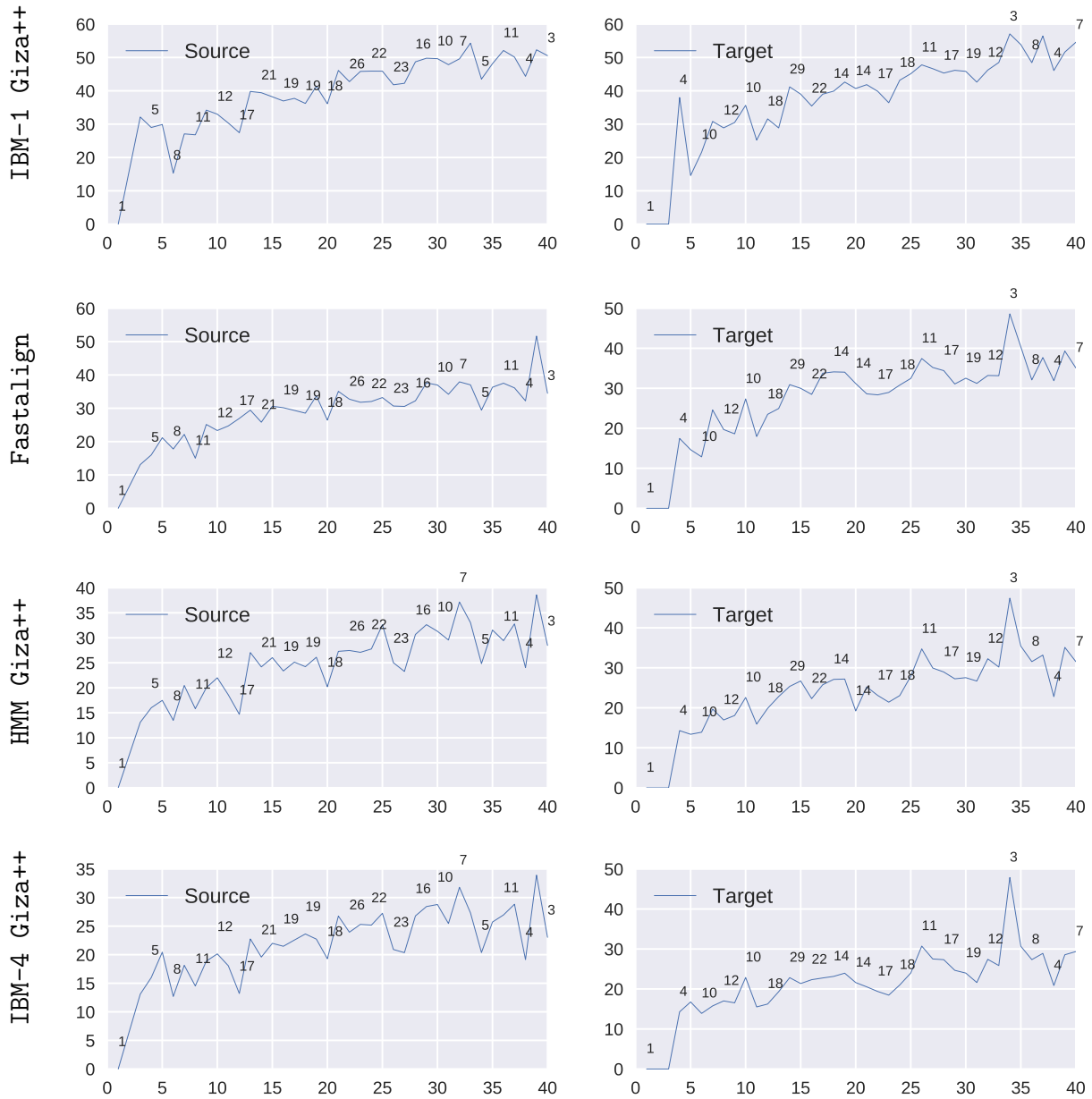


Figure A.94: Baselines for the direction German-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.

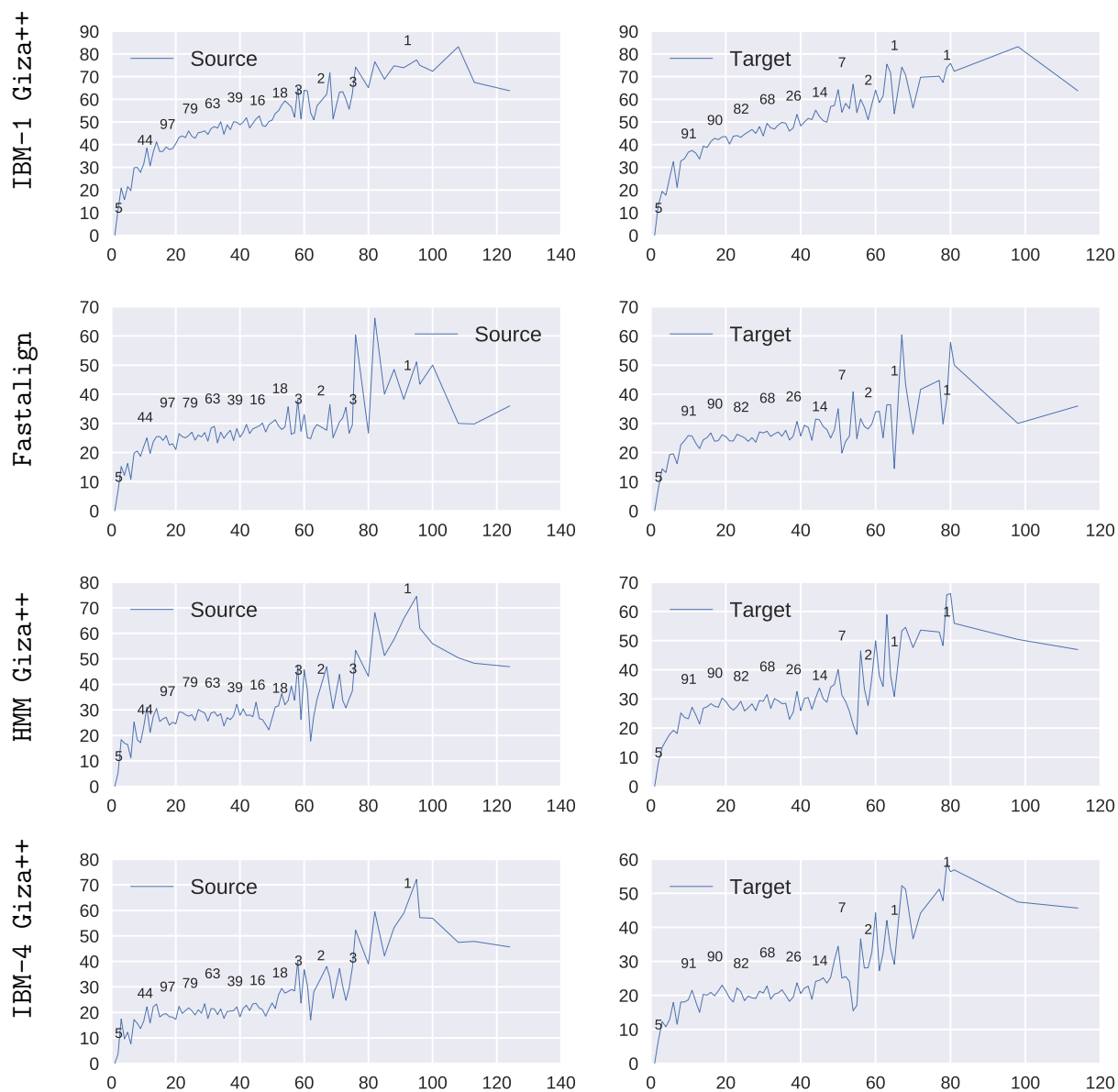


Figure A.95: Baselines for the direction English-Czech: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.

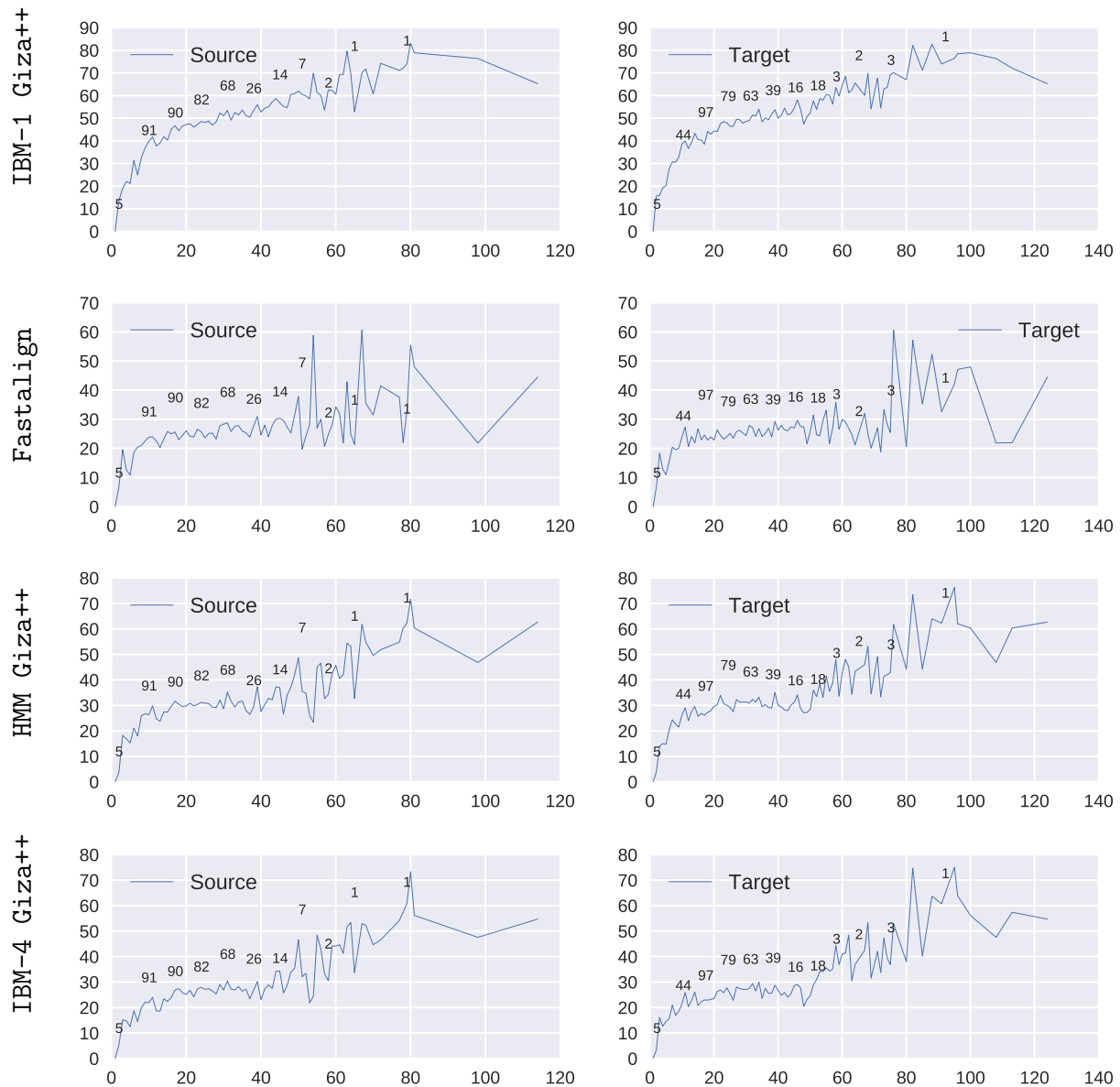


Figure A.96: Baselines for the direction Czech-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.

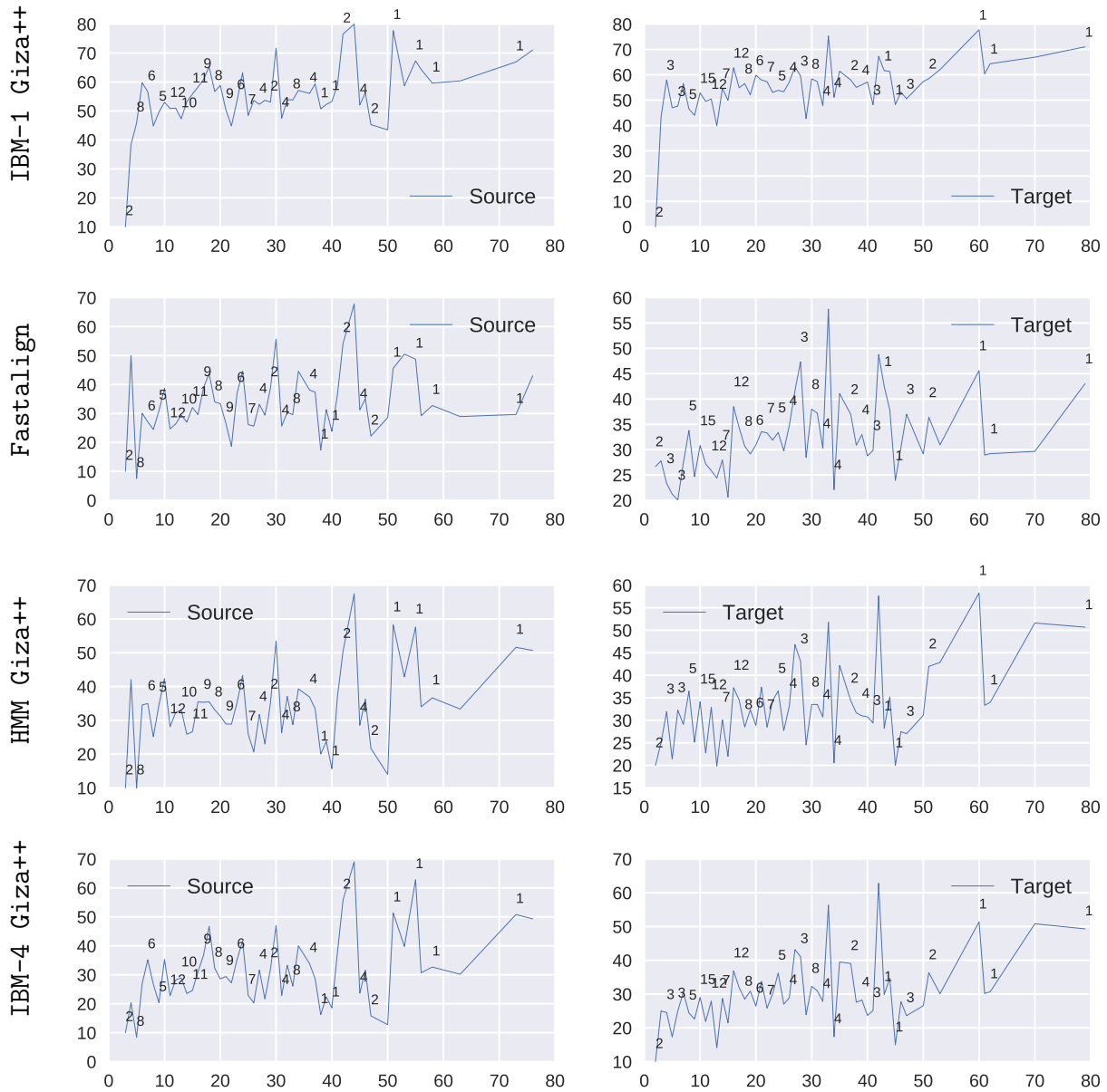


Figure A.97: Baselines for the direction English-Romanian: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.

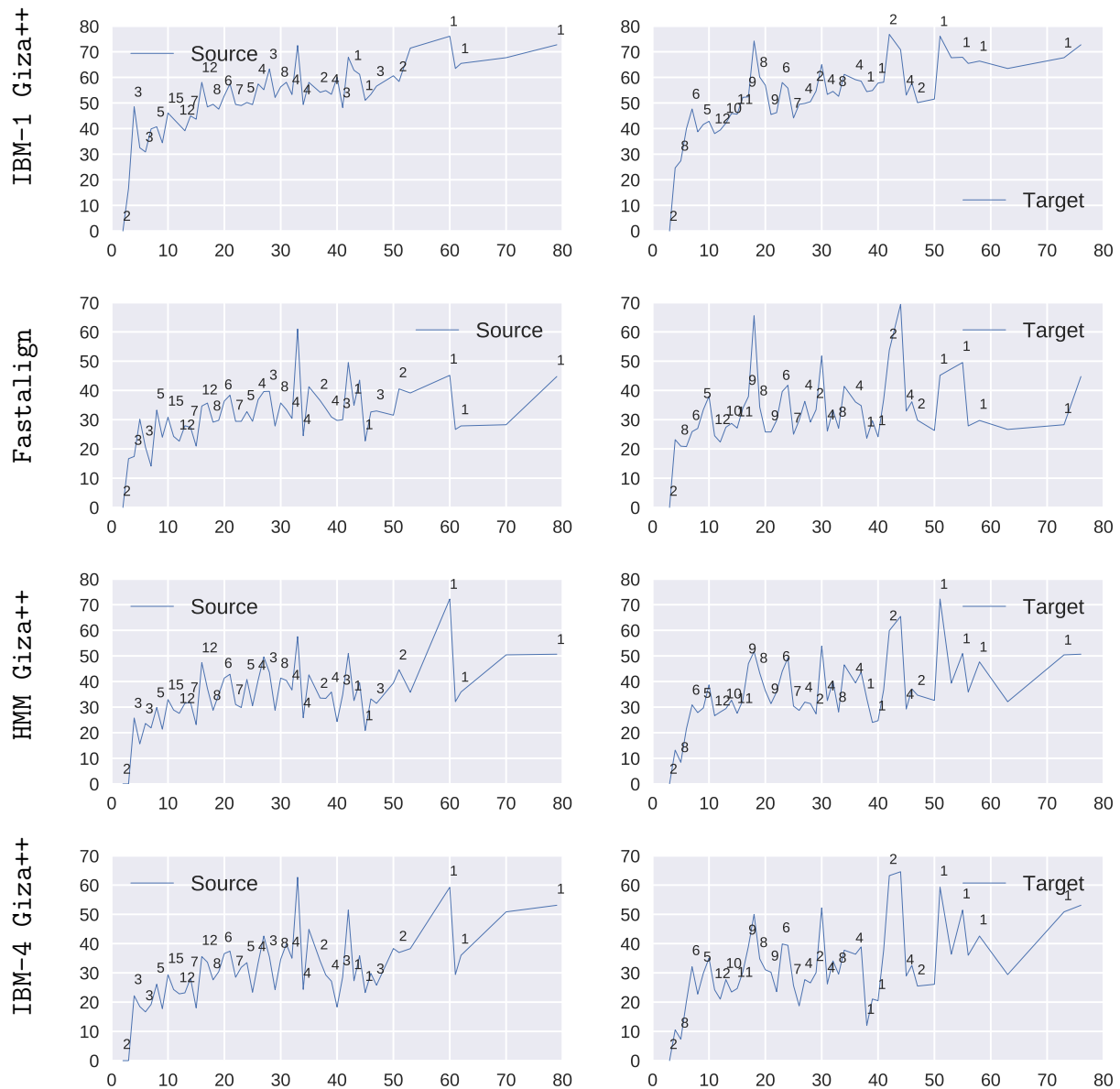


Figure A.98: Baselines for the direction Romanian-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.

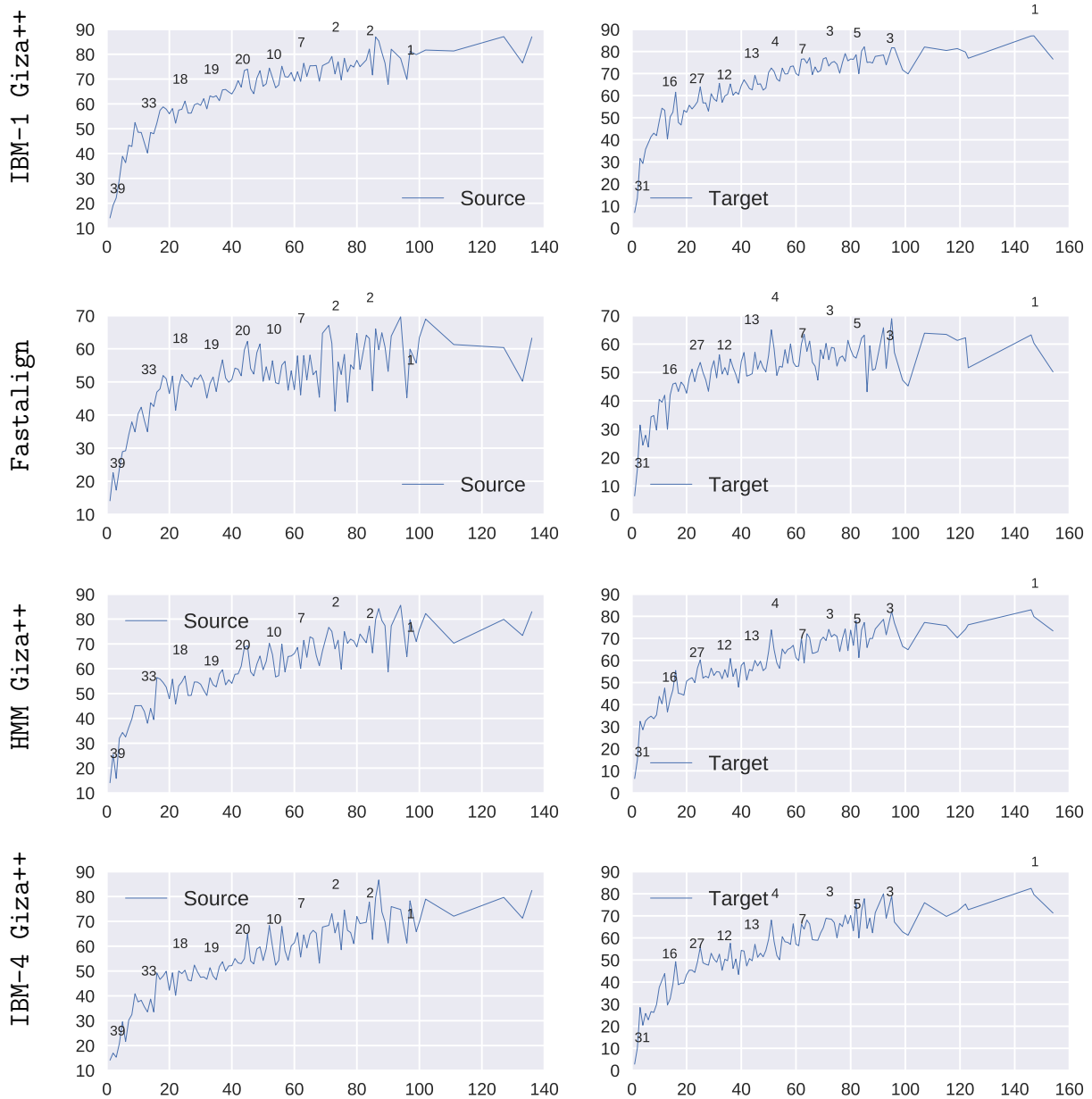


Figure A.99: Baselines for the direction English-Japanese: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.

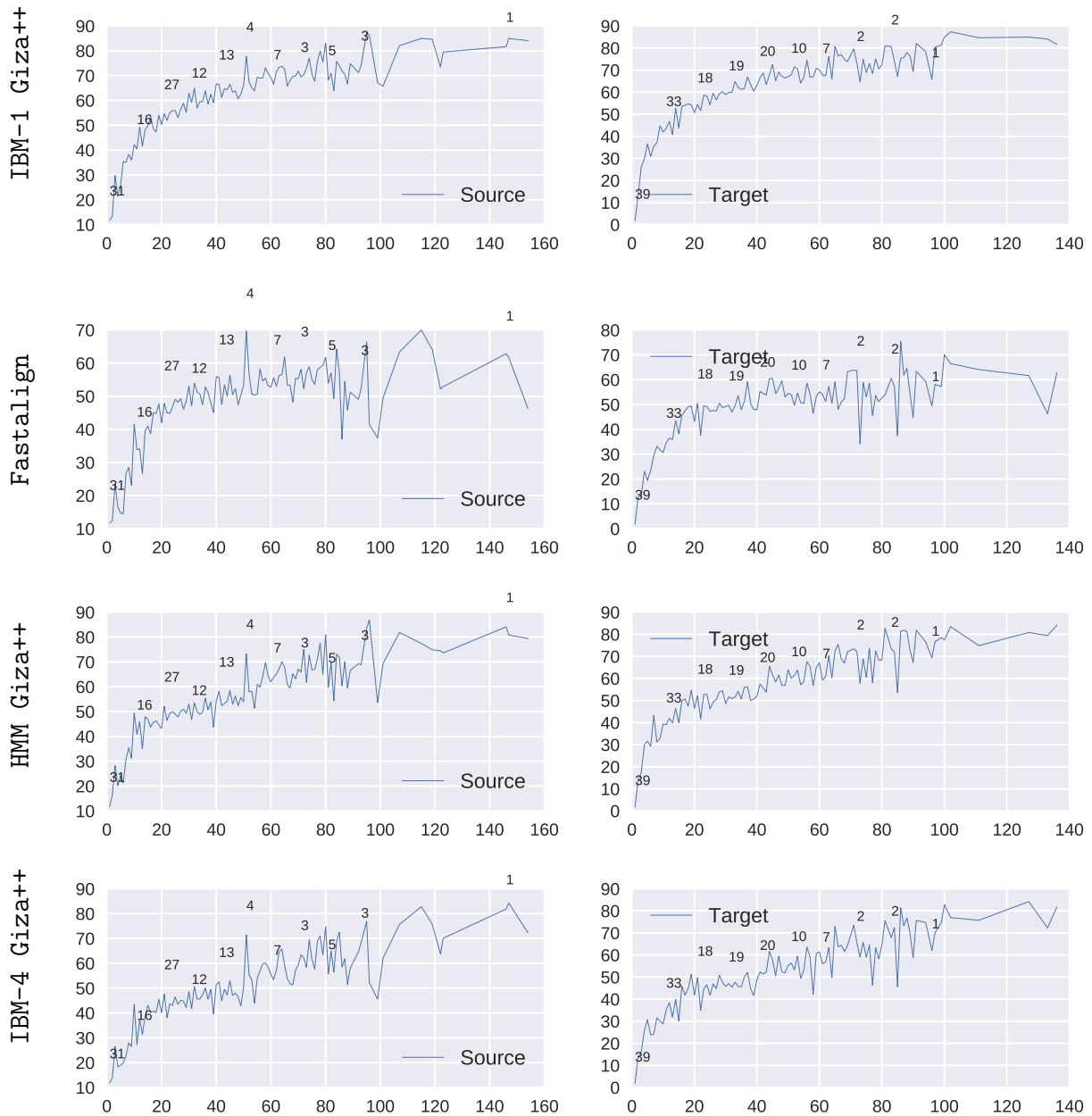


Figure A.100: Baselines for the direction Japanese-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.

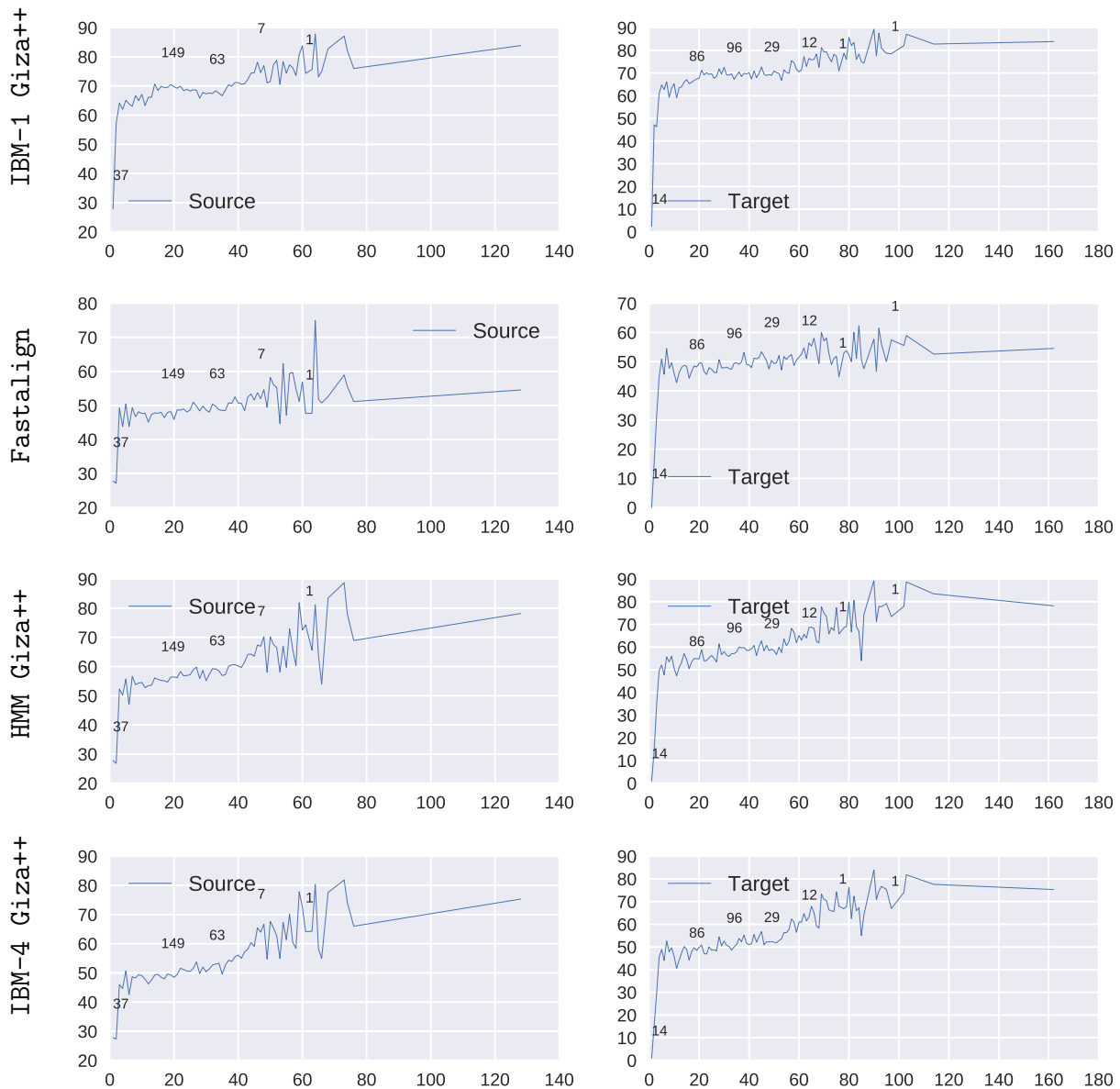


Figure A.101: Baselines for the direction English-Vietnamese: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.

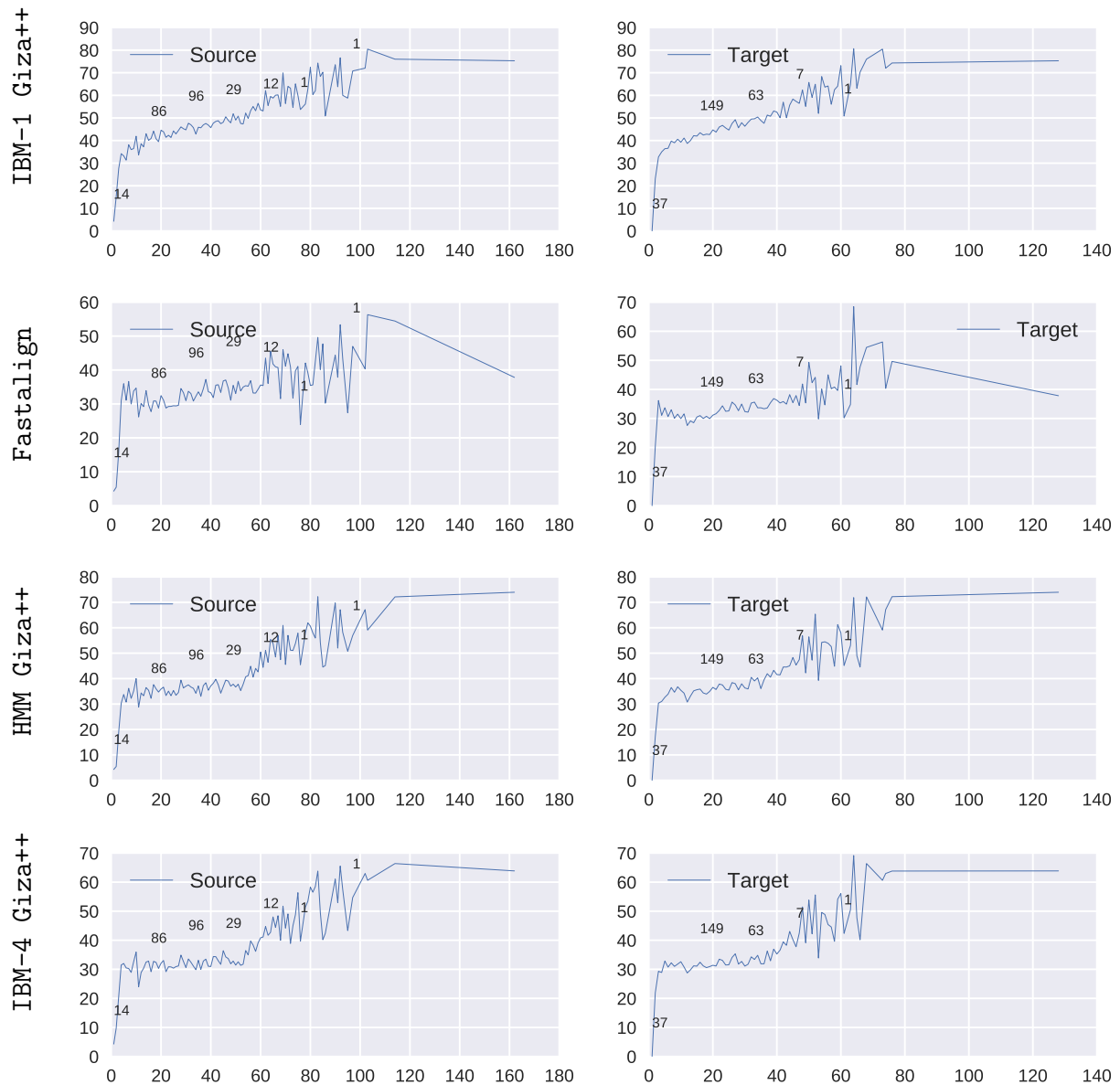


Figure A.102: Baselines for the direction Vietnamese-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.

A.11.6 Baselines: AER as a function of sentence length difference

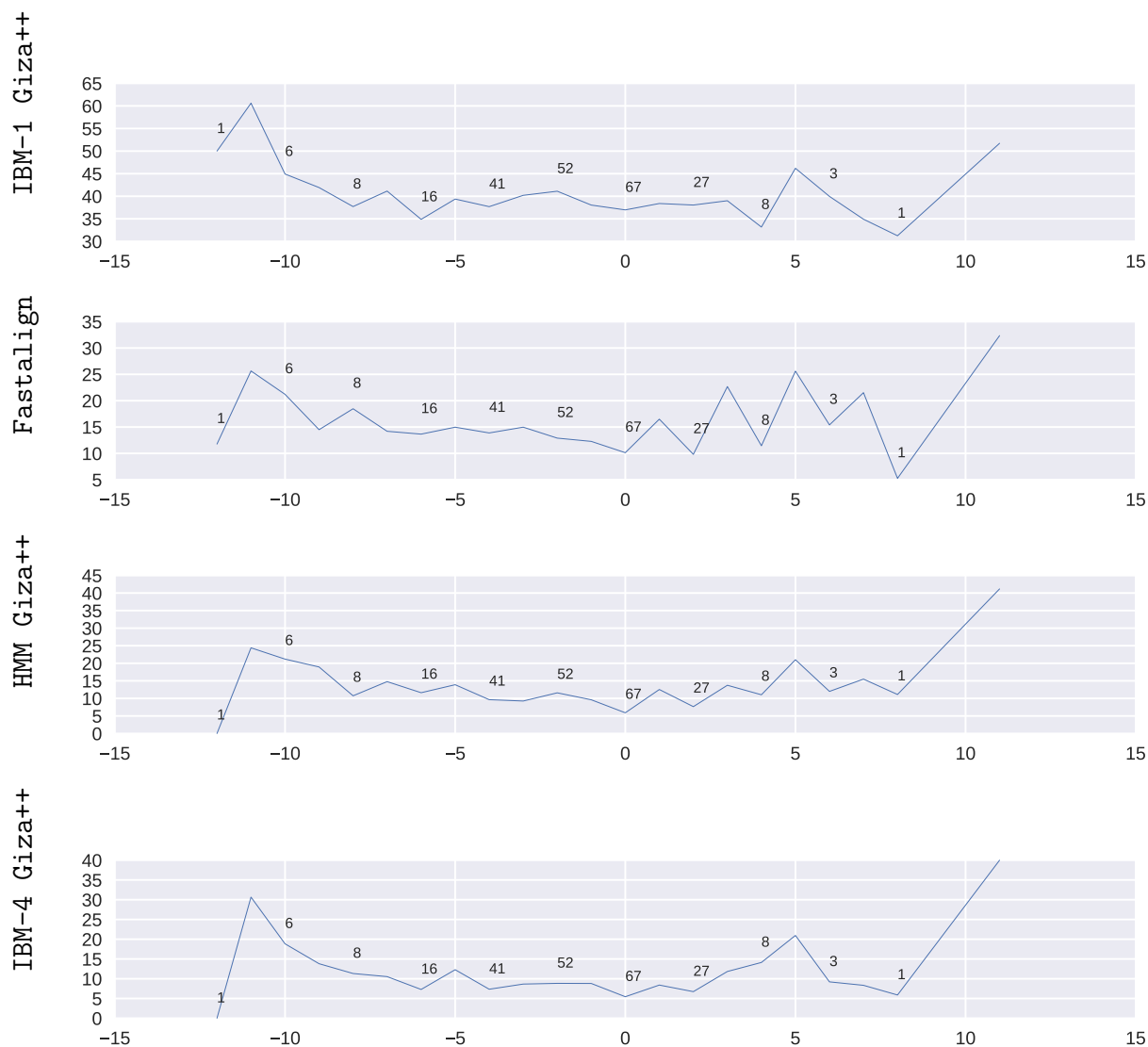
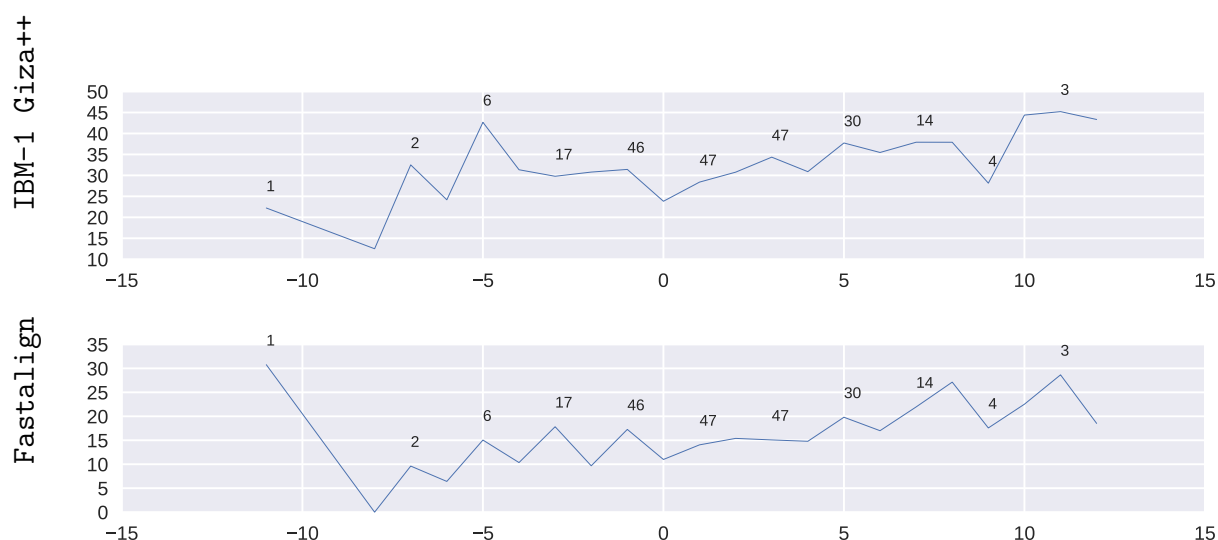


Figure A.103: Baselines for the direction English-French: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



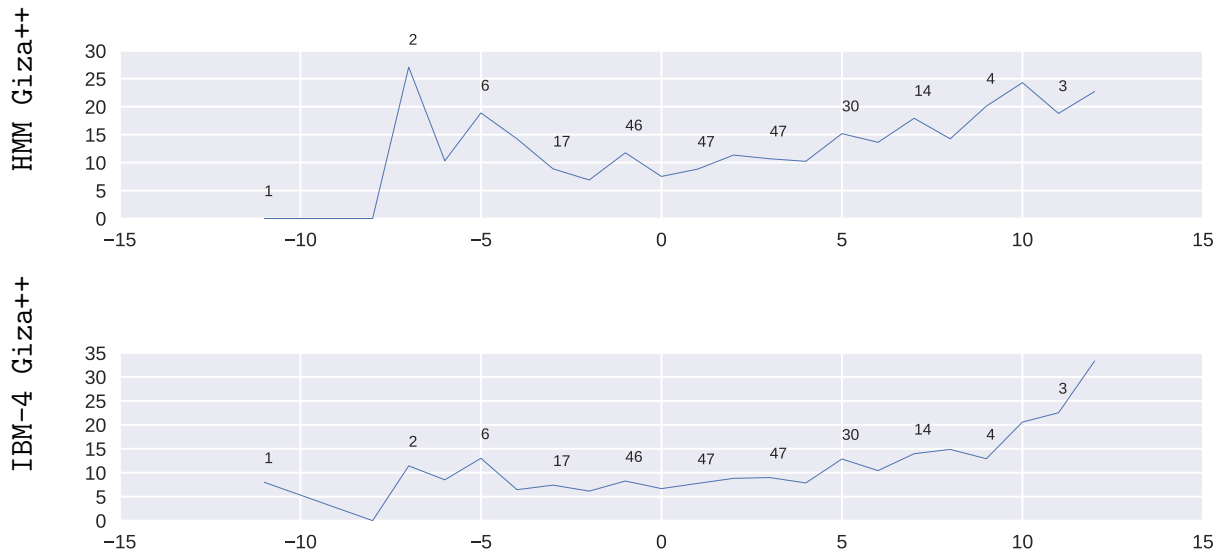


Figure A.104: Baselines for the direction French-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.

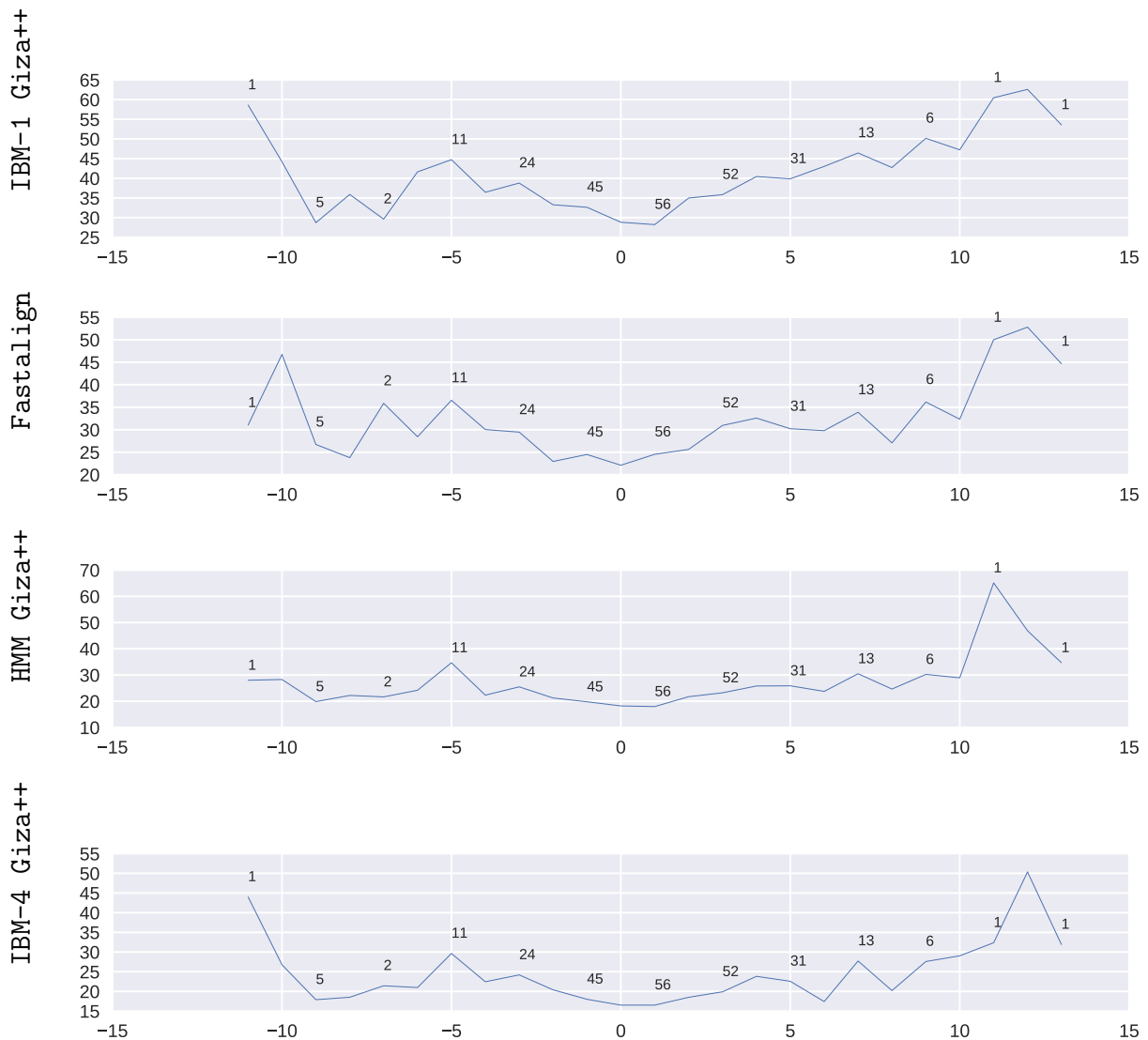


Figure A.105: Baselines for the direction English-German: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.

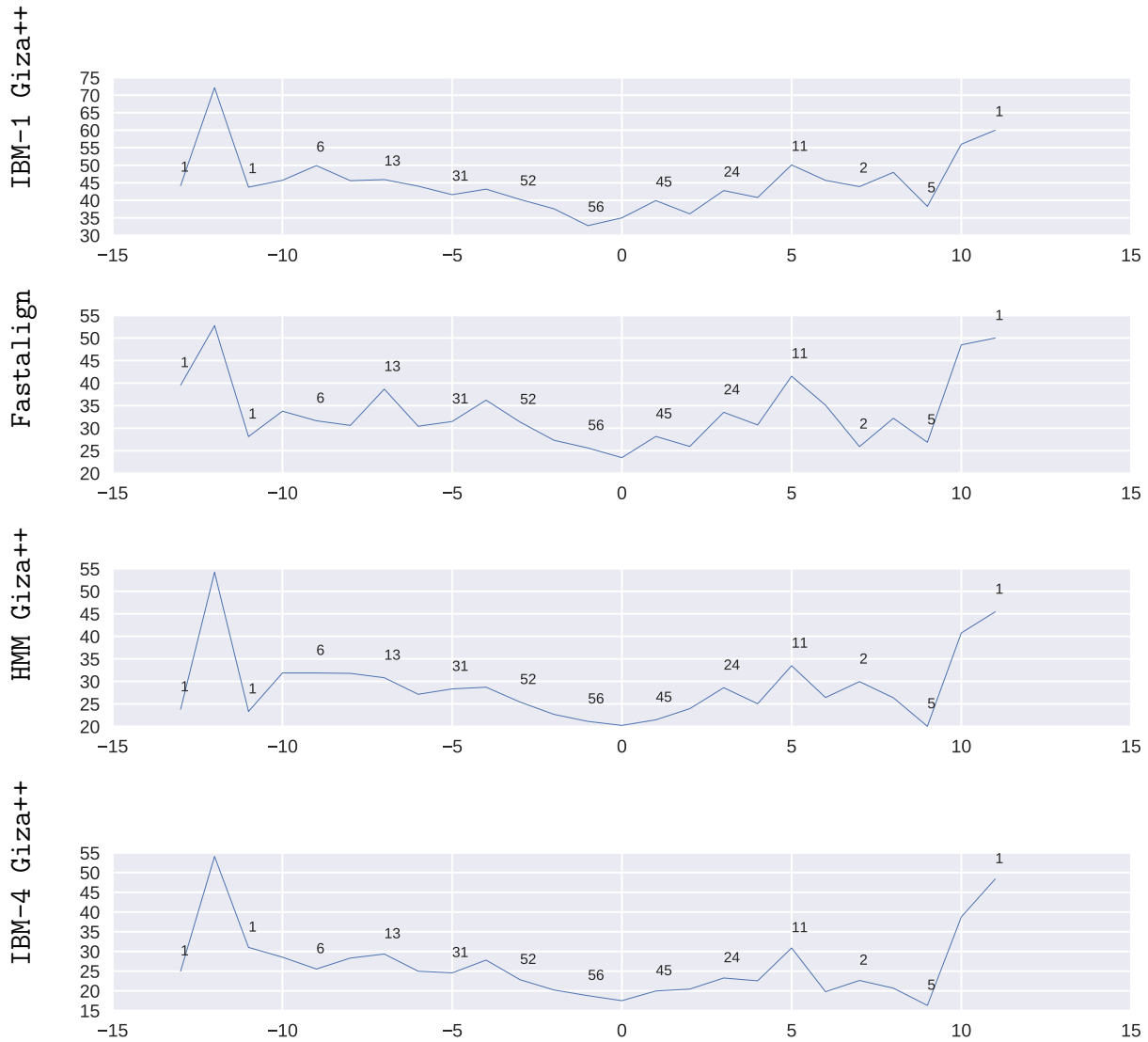


Figure A.106: Baselines for the direction German-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.

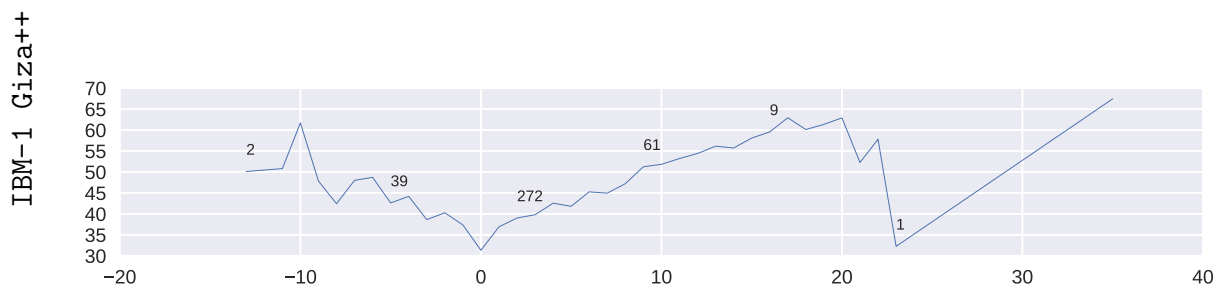
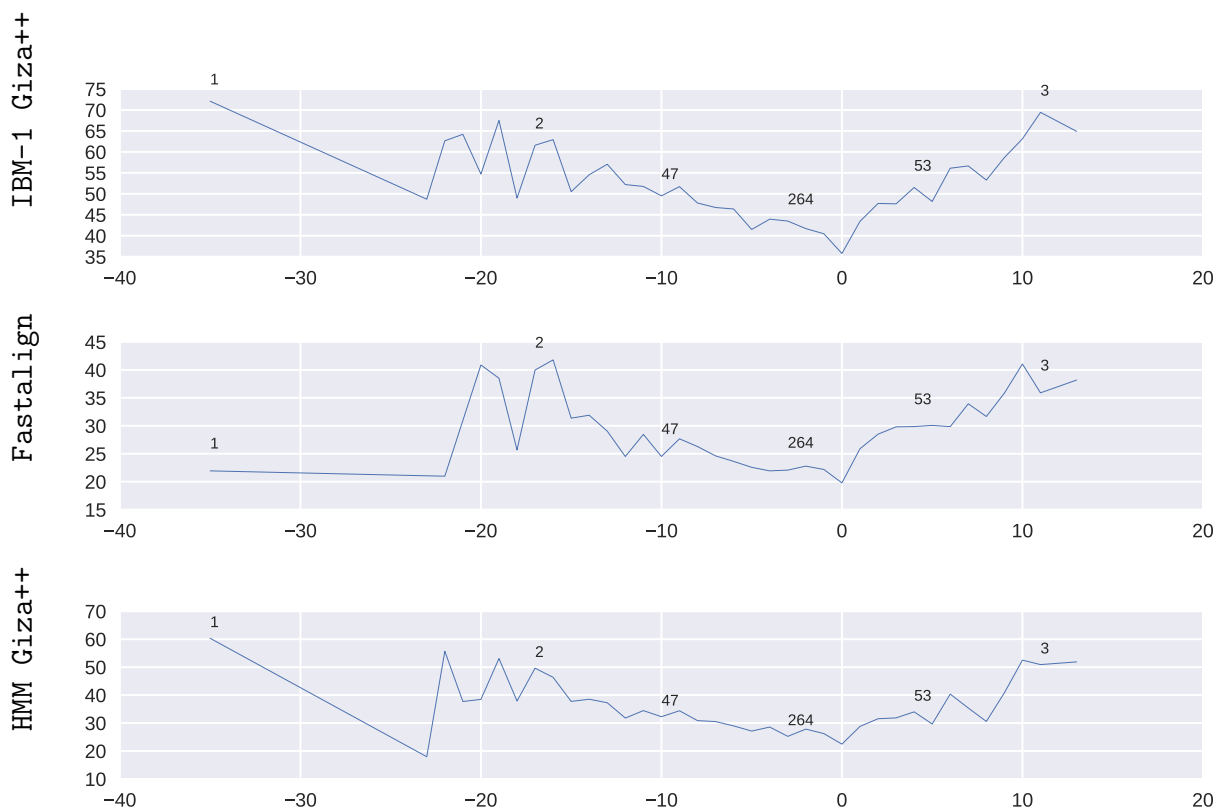




Figure A.107: Baselines for the direction English-Czech: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



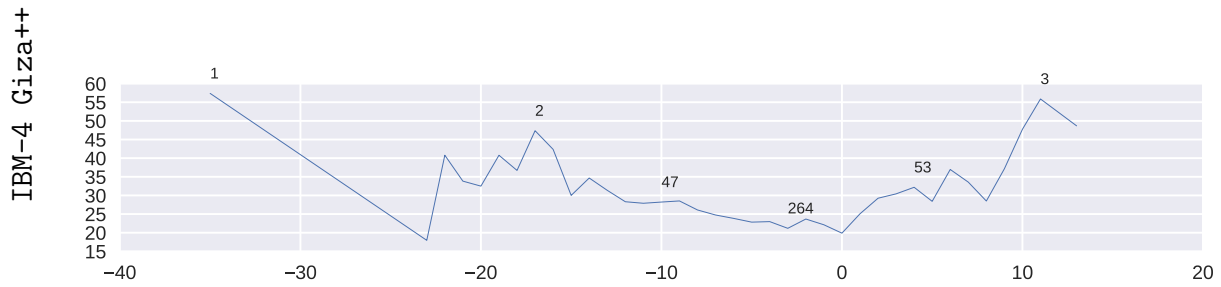


Figure A.108: Baselines for the direction Czech-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.

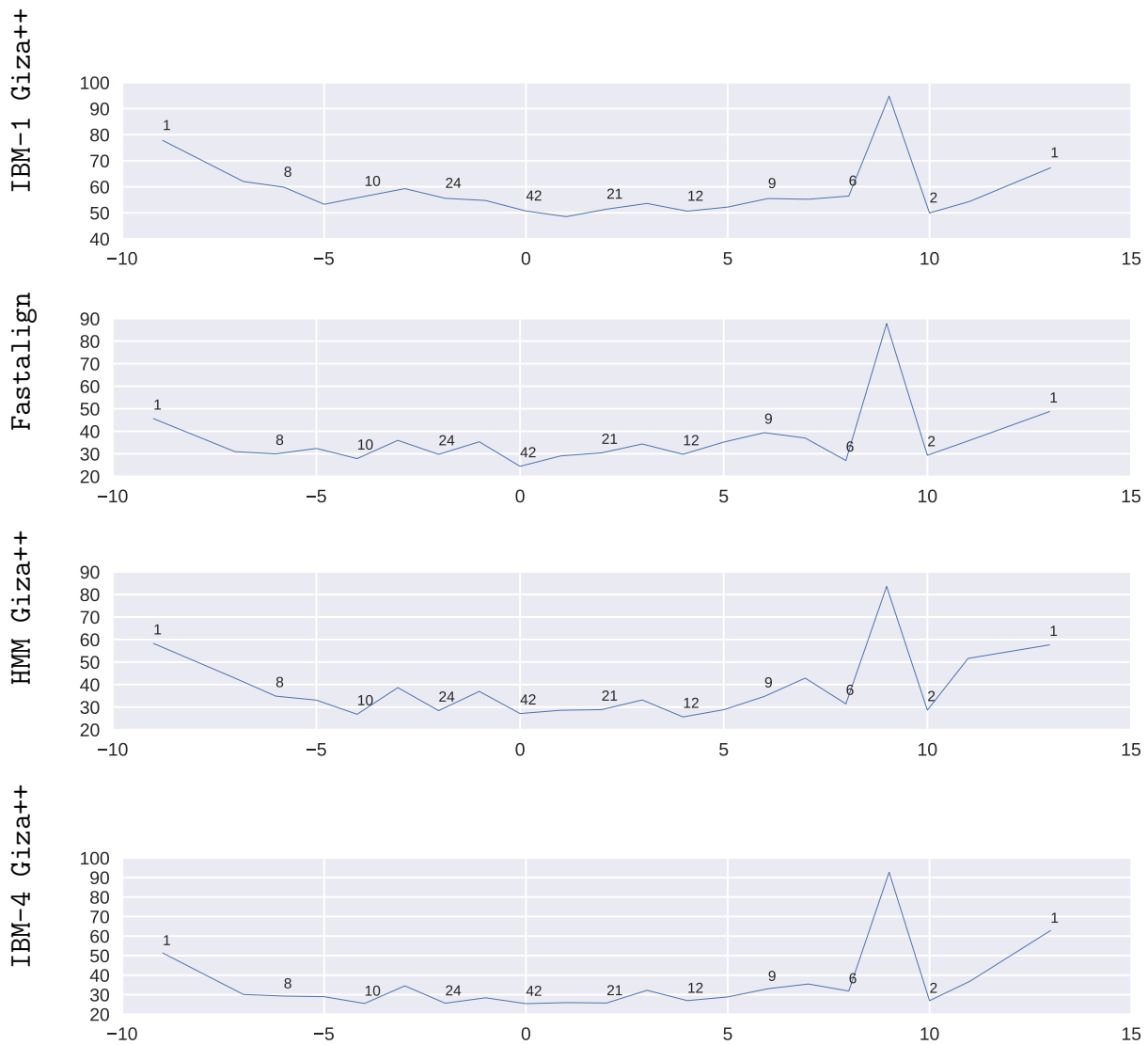


Figure A.109: Baselines for the direction English-Romanian: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.

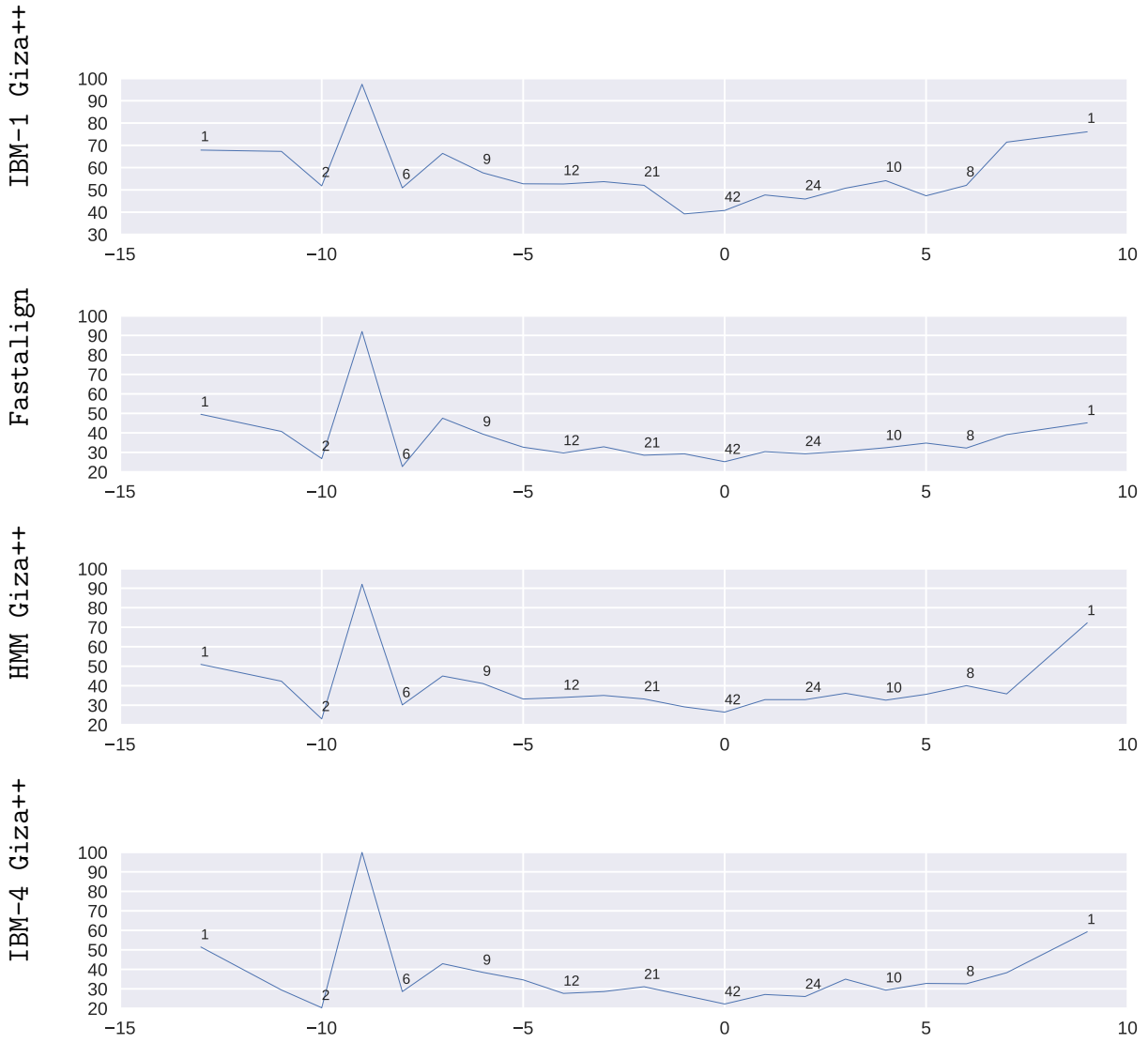


Figure A.110: Baselines for the direction Romanian-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.

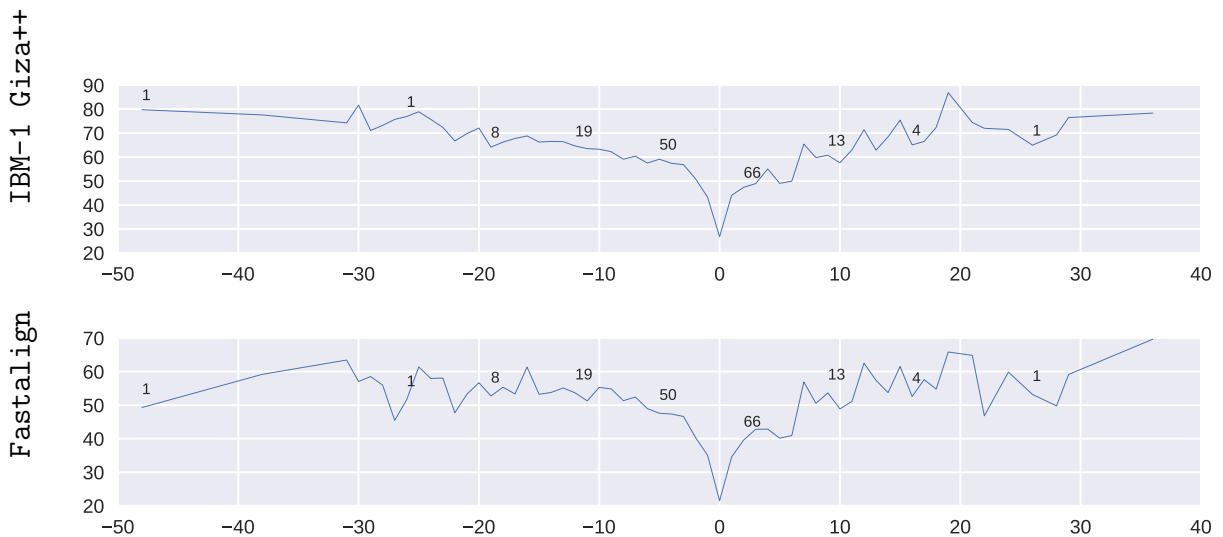




Figure A.111: Baselines for the direction English-Japanese: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.

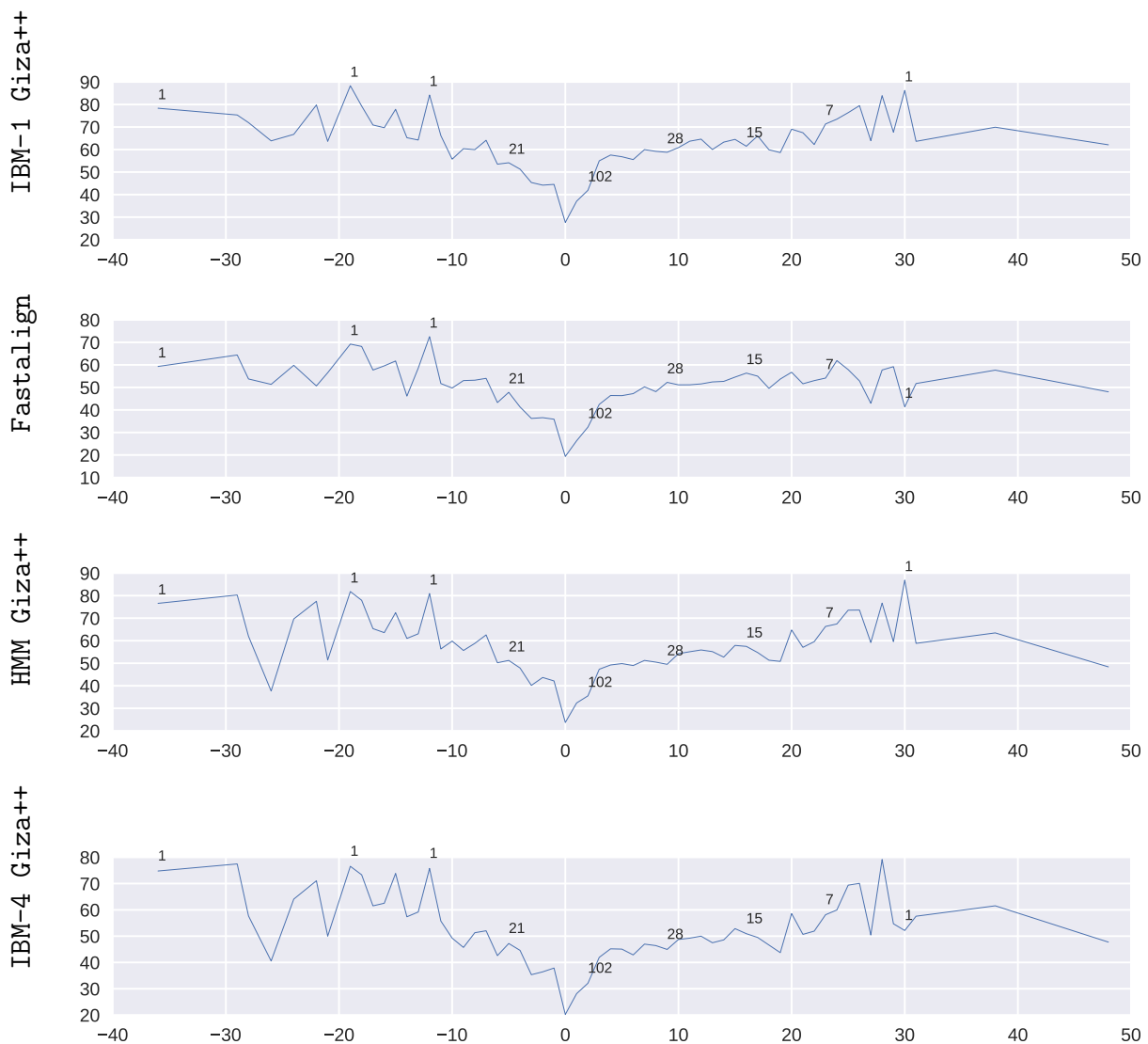


Figure A.112: Baselines for the direction Japanese-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.

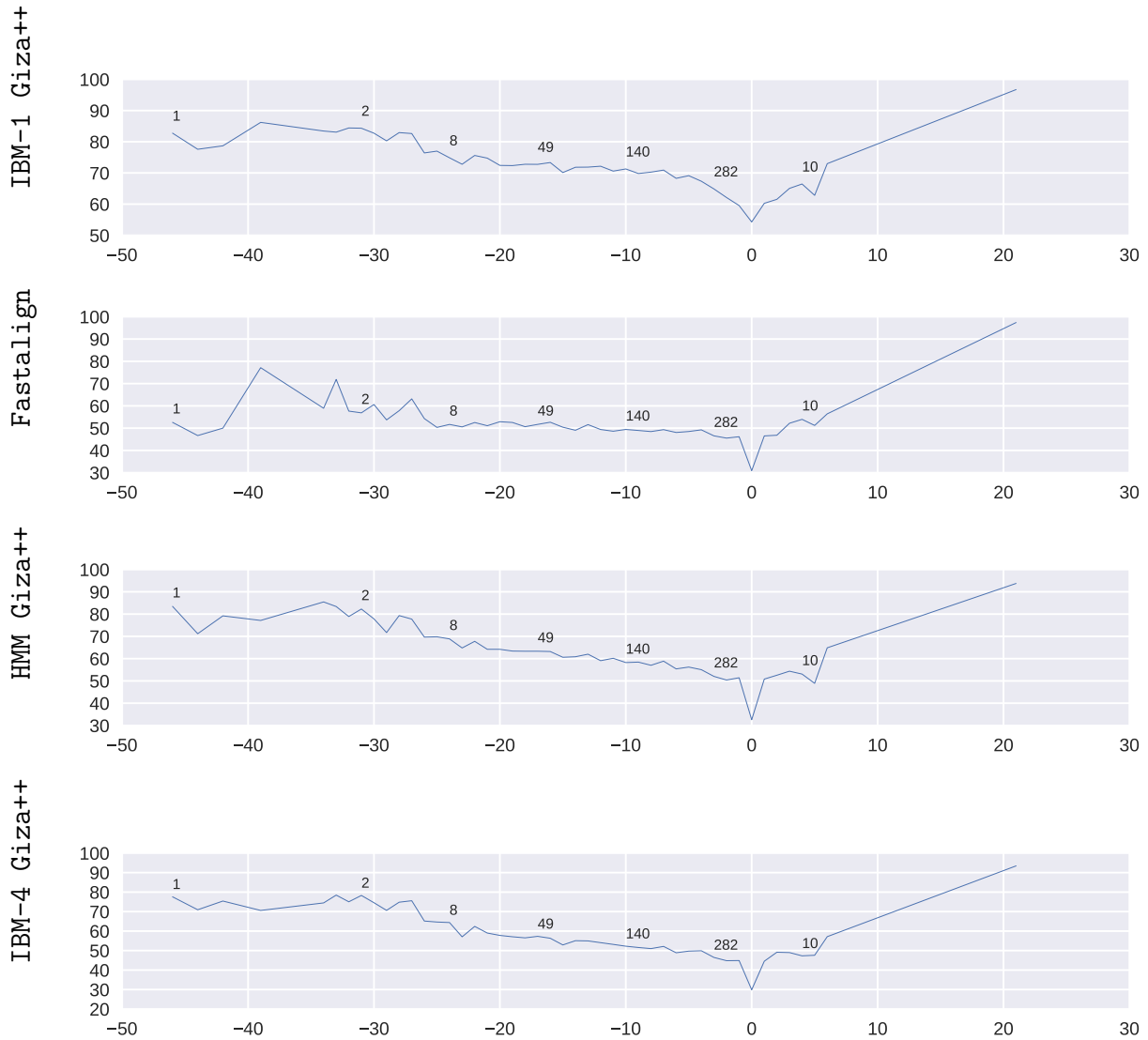
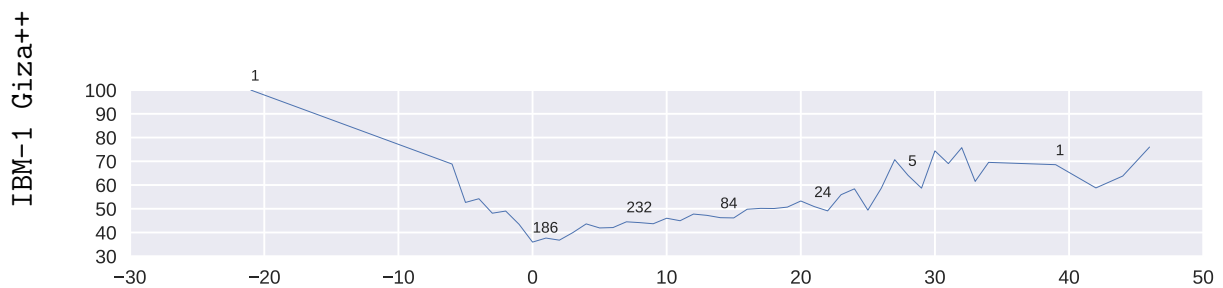


Figure A.113: Baselines for the direction English-Vietnamese: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



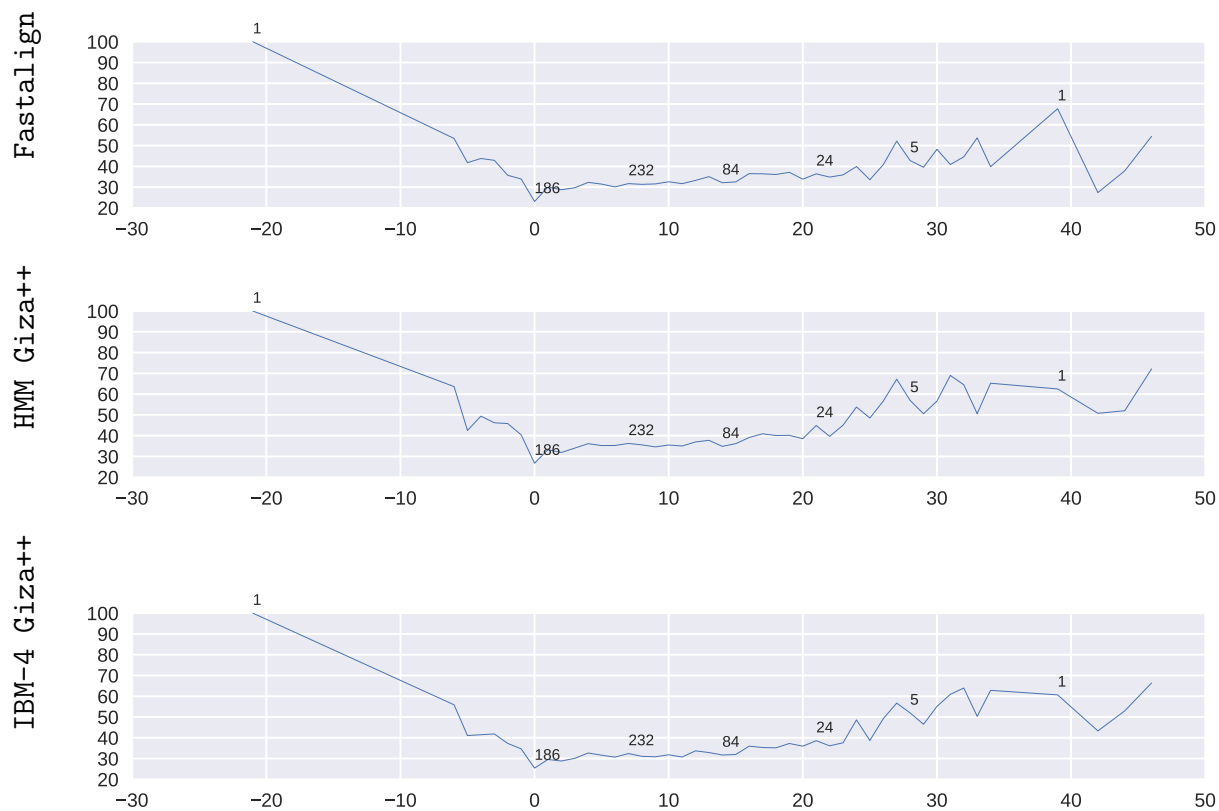


Figure A.114: Baselines for the direction Vietnamese-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.

Appendix B

Appendix: Neural word alignment models

Contents

B.1	Scores	133
B.2	Correct and incorrect alignment links	137
B.3	Unaligned words	139
B.4	Alignment types	145
B.5	Jump widths	147
B.5.1	Jump width statistics	147
B.5.2	Correct and incorrect jumps	153
B.6	Rare words	186
B.7	Unknown words	189
B.7.1	Unknown words in the case of a full vocabulary	189
B.7.2	Unknown words in the case of 50K word vocabulary	192
B.8	Part-of-Speech	196
B.8.1	Alignment links for content and function words	196
B.8.2	Unaligned words for content and function words	202
B.9	Symmetrization	208
B.9.1	Intersection	208
B.9.2	Grow-diag-final (GDF)	211
B.10	AER as a function of sentence length	214
B.11	AER as a function of sentence length difference	242

B.1 Scores

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
IBM-1 Giza++	40.1	26.7	71.55	16.41	89.01	33.9	36.49	59.24	26.37	88.81
IBM-1+NN	27.96	36.42	69.66	24.65	89.5	27.21	38.08	68.12	26.43	89.52
IBM-1+NNCtxCc	27.42	38.66	68.65	26.91	89.59	28.86	39.13	65.26	27.94	89.4
IBM-1+NNCtxCNN	27.86	38.2	68.4	26.5	89.54	27.99	39.07	66.89	27.59	89.5
IBM-1+NNChar	28.76	37.5	67.13	26.01	89.42	31.4	37.21	62.64	26.47	89.11
IBM-1+NNCharWord	27.03	38.69	68.96	26.89	89.61	28.33	38.97	66.1	27.63	89.45
Fastalign	15.19	44.98	82.5	30.92	90.78	16.23	46.32	80.08	32.58	90.79
HMM Giza++	11.99	45.18	86.12	30.62	90.94	11.97	45.98	85.2	31.49	90.98
HMM+NN	11.84	45.57	86.68	30.91	91	11.15	46.86	86.17	32.18	91.1
HMM+NNCtxCc	10.39	45.83	88.65	30.9	91.09	11.57	46.6	86.33	31.91	91.08
HMM+NNCtxCNN	11.64	45.16	87.46	30.44	90.99	12.71	46.42	85.78	31.82	91.04
HMM+NNCharTgt	9.17	47.22	89.53	32.07	91.26	9.56	47.87	88.15	32.86	91.27
HMM+NNCharWord	10.45	47.33	87.94	32.38	91.21	10.27	48.56	86.92	33.69	91.3
HMM+NNCharBoth	10.9	46.74	87.41	31.9	91.13	11.17	47.5	86.1	32.8	91.16
HMM+NNCharJT	8.41	44.71	91.58	29.57	91.08	7.7	44.45	92.82	29.22	91.09
HMM+NNCharJB	8.47	44.38	91.8	29.26	91.06	7.74	46.26	91.42	30.96	91.23
IBM-4 Giza++	10	44.43	90.61	29.43	91.02	9.64	45.43	89.58	30.43	91.08

Table B.1: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-French

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
IBM-1 Giza++	39.03	58.76	59.1	58.43	96.4	42.66	55.39	57.02	53.84	96.2
IBM-1+NN	37.64	60.07	62.98	57.41	96.65	39.22	58.53	62.22	55.25	96.57
IBM-1+NNCtxCc	36.4	61.36	63.07	59.74	96.7	36.31	61.36	65.4	57.79	96.81
IBM-1+NNCtxCNN	37.16	60.67	62.23	59.19	96.64	36.02	61.68	64.95	58.73	96.8
IBM-1+NNChar	36.22	61.55	62.76	60.39	96.69	40.88	56.99	59.75	54.48	96.4
IBM-1+NNCharWord	35.31	62.41	63.79	61.08	96.77	40.48	57.44	60.15	54.97	96.43
Fastalign	28.98	68.75	71.11	66.54	97.35	31.28	66.47	70.73	62.69	97.23
HMM Giza++	23.92	73.3	79.23	68.2	97.82	26.33	71.04	79.47	64.23	97.7
HMM+NN	26.78	70.95	73.94	68.2	97.55	29.44	68.21	74.69	62.76	97.44
HMM+NNCtxCc	34.36	63.19	70.54	57.23	97.08	32.31	65.34	72.92	59.19	97.25
HMM+NNCtxCNN	30.64	67.18	70.28	64.33	97.24	30.34	67.41	73.13	62.52	97.35
HMM+NNCharTgt	26.04	71.57	75.99	67.64	97.64	28.11	69.48	75.59	64.29	97.52
HMM+NNCharWord	24.98	72.64	76.53	69.13	97.72	29.77	67.76	74.12	62.4	97.4
HMM+NNCharBoth	27.14	70.6	73.65	67.79	97.52	29.31	68.34	74.11	63.41	97.42
HMM+NNCharJT	23.79	73.15	82.8	65.52	97.89	25.21	71.85	83.64	62.98	97.84
HMM+NNCharJB	23.69	73.38	82.38	66.15	97.9	24.9	72.16	83.36	63.61	97.85
IBM-4 Giza++	21.46	75.48	85.79	67.39	98.08	23.31	73.63	86.56	64.06	97.99

Table B.2: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-German

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
IBM-1 Giza++	45.09	46.75	50.4	43.59	95.97	48.47	42.88	49.17	38.02	95.89
IBM-1+NN	42.29	48.64	54.32	44.04	96.22	40.97	49.08	56.81	43.2	96.36
IBM-1+NNCtxCc	44.63	46.86	51.31	43.12	96.03	42.33	48.22	57.14	41.71	96.36
IBM-1+NNCtxCNN	45.47	46.13	50.35	42.55	95.96	41.92	48.4	56.42	42.38	96.33
IBM-1+NNChar	40.85	50.35	54.2	47.01	96.24	42.35	47.99	55.68	42.17	96.29
IBM-1+NNCharWord	40.28	50.82	55.18	47.11	96.3	46.2	44.94	51.87	39.65	96.06
Fastalign	25.75	64.09	70.98	58.42	97.34	25.3	62.86	73.13	55.13	97.36
HMM Giza++	27.86	61.22	70.81	53.92	97.23	30.38	57.28	69.26	48.83	97.04
HMM+NN	23.5	65.45	74.39	58.43	97.5	24.06	64.03	75.48	55.6	97.46
HMM+NNCtxCc	24.31	63.74	76.38	54.69	97.47	25.02	62.66	75.88	53.36	97.42
HMM+NNCtxCNN	24.18	65.38	73.64	58.79	97.47	22.92	64.63	78.32	55.01	97.55
HMM+NNCharTgt	16.74	69.36	83.82	59.15	97.88	24.61	63.94	73.42	56.63	97.41
HMM+NNCharWord	16.04	70.34	83.18	60.93	97.91	22.8	64.94	77.4	55.94	97.55
HMM+NNCharBoth	17.38	69.09	81.89	59.76	97.83	28.26	61.04	70.15	54.02	97.2
HMM+NNCharJT	15.94	68.31	88.7	55.55	97.91	25.82	62.11	74.22	53.39	97.35
HMM+NNCharJB	16.38	69.35	83.84	59.13	97.88	23.88	62.72	78.56	52.19	97.48
IBM-4 Giza++	20.92	65.7	79.48	56	97.63	26.5	59.81	75.58	49.48	97.3

Table B.3: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Czech

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
IBM-1 Giza++	56.02	43.99	58.8	35.14	96.66	53.52	46.49	49.92	43.5	96.26
IBM-1+NN	46.4	53.62	57.71	50.07	96.77	44.9	55.11	60.08	50.9	96.91
IBM-1+NNCtxCc	49.93	50.09	54.28	46.49	96.54	43.95	56.07	61.32	51.64	96.98
IBM-1+NNCtxCNN	46.15	53.87	60.08	48.81	96.88	43.93	56.09	62.73	50.72	97.04
IBM-1+NNChar	50.16	49.85	54.28	46.09	96.54	48.28	51.73	56.08	48.01	96.66
IBM-1+NNCharWord	46.54	53.47	56.91	50.43	96.73	43.94	56.08	60.71	52.1	96.96
Fastalign	33.36	66.65	72.77	61.49	97.7	32.91	67.1	73.7	61.59	97.75
HMM Giza++	33.36	66.65	75.28	59.8	97.77	36.38	63.64	72.9	56.46	97.59
HMM+NN	30.69	69.33	76.93	63.09	97.92	40.12	59.89	63.85	56.4	97.18
HMM+NNCtxCc	33.84	66.18	75.21	59.08	97.75	34.83	65.19	73.24	58.73	97.66
HMM+NNCtxCNN	30.87	69.15	80.01	60.89	97.97	31.03	68.99	77.58	62.11	97.92
HMM+NNCharTgt	27.54	72.47	78.13	67.59	98.08	28.01	72.01	82.79	63.71	98.15
HMM+NNCharWord	25.51	74.51	82.68	67.82	98.27	29.19	70.83	79.3	63.99	98.03
HMM+NNCharBoth	28.14	71.88	77.39	67.1	98.04	31.53	68.49	77.29	61.49	97.89
HMM+NNCharJT	28.27	71.75	77.83	66.55	98.04	29.02	71	83.72	61.64	98.12
HMM+NNCharJB	26.85	73.17	83.55	65.08	98.22	29.76	70.26	80.7	62.21	98.04
IBM-4 Giza++	31.04	68.98	79.28	61.04	97.95	32.3	67.72	80.97	58.2	97.93

Table B.4: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Romanian

Baselines	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
BM-1 Giza++	63.12	36.88	40.24	34.04	97.22	61.55	38.45	40.13	36.9	97.18
IBM-1+NN	62.64	37.36	48.83	30.25	97.58	56.9	43.1	49.95	37.91	97.61
IBM-1+NNCtxCc	61.62	38.38	44.94	33.49	97.43	55.69	44.31	47.39	41.61	97.5
IBM-1+NNCtxCNN	60.21	39.79	45.2	35.54	97.43	54.01	45.99	48.17	43.99	97.53
IBM-1+NNChar	57.96	42.04	57.54	33.12	97.82	54.91	45.09	47.76	42.7	97.52
IBM-1+NNCharWord	60.92	39.08	45.78	34.08	97.46	53.54	46.46	48.62	44.49	97.55
Fastalign	50.67	49.33	55.79	44.21	97.83	49.37	50.63	54.59	47.2	97.8
HMM Giza++	57.01	42.99	52.22	36.53	97.69	54.41	45.59	52.45	40.32	97.7
HMM+NN	49.68	50.32	70.46	39.13	98.16	47.09	52.91	56.81	49.51	97.9
HMM+NNCtxCc	47.93	52.07	70.37	41.32	98.18	45.67	54.33	56.82	52.05	97.91
HMM+NNCtxCNN	46.82	53.18	71.49	42.33	98.22	44.79	55.21	72.7	44.51	98.28
HMM+NNCharTgt	43.92	56.08	65.13	49.25	98.16	43.6	56.4	67.58	48.39	98.21
HMM+NNCharWord	45.35	54.65	65.28	46.99	98.14	40.16	59.84	66.91	54.12	98.27
HMM+NNCharBoth	43.71	56.29	65.37	49.43	98.17	40.26	59.74	72.73	50.68	98.37
HMM+NNCharJT	39.69	60.31	84.87	46.78	98.53	39.22	60.78	72.74	52.19	98.39
HMM+NNCharJB	40.92	59.08	84.77	45.35	98.5	37.71	62.29	67.77	57.62	98.33
IBM-4 Giza++	52.52	47.48	61.27	38.76	97.95	49.23	50.77	61.46	43.24	98

Table B.5: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Japanese

Baselines	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
BM-1 Giza++	69.43	30.58	43.42	23.59	96.51	46.45	53.56	50.79	56.64	96.8
IBM-1+NN	58.87	41.13	57.25	32.1	97.01	42.25	57.75	57.33	58.19	97.23
IBM-1+NNCtxCc	59.4	40.6	49.98	34.19	96.74	37.5	62.5	61.2	63.87	97.5
IBM-1+NNCtxCNN	54.55	45.45	62.31	35.77	97.2	38.45	61.56	60.89	62.25	97.47
IBM-1+NNChar	55.06	44.94	62.54	35.07	97.2	41.15	58.86	63.9	54.56	97.51
IBM-1+NNCharWord	53.2	46.8	69.3	35.33	97.38	35.45	64.55	73.35	57.64	97.94
Fastalign	48.89	51.11	56.87	46.42	97.11	32.82	67.19	64.5	70.11	97.77
HMM Giza++	57.86	42.14	49.18	36.86	96.7	37.57	62.43	60.28	64.74	97.46
HMM+NN	49.27	50.73	63.52	42.24	97.33	31.45	68.56	67.75	69.39	97.93
HMM+NNCtxCc	49.56	50.44	67.34	40.32	97.42	33	67	67.75	66.27	97.87
HMM+NNCtxCNN	48.62	51.39	64.29	42.8	97.36	29.23	70.78	69.94	71.64	98.07
HMM+NNCharTgt	47.52	52.49	67.71	42.86	97.47	30.94	69.06	75.69	63.51	98.15
HMM+NNCharWord	44.37	55.64	83.68	41.67	97.83	30.92	69.09	70.04	68.16	98.01
HMM+NNCharBoth	48.1	51.91	74.95	39.7	97.6	30.72	69.28	75.79	63.81	98.16
HMM+NNCharJT	43.78	56.22	84.2	42.2	97.86	28.86	71.15	70.35	71.97	98.1
HMM+NNCharJB	43.28	56.73	84.49	42.7	97.88	27.59	72.42	72.7	72.14	98.21
IBM-4 Giza++	51.91	48.1	58.38	40.89	97.12	33.19	66.82	67.17	66.47	97.85

Table B.6: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Vietnamese

B.2 Correct and incorrect alignment links

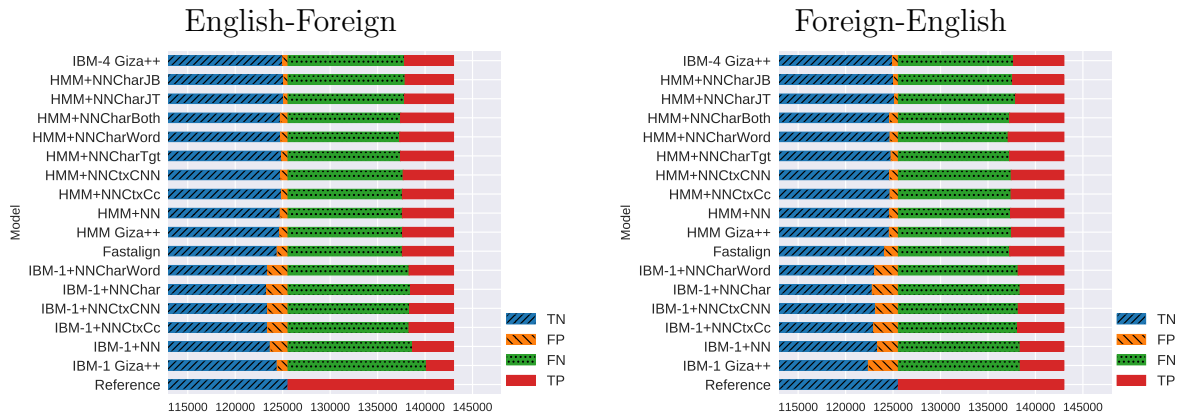


Figure B.1: Results of our neural models: Alignment links for English-French

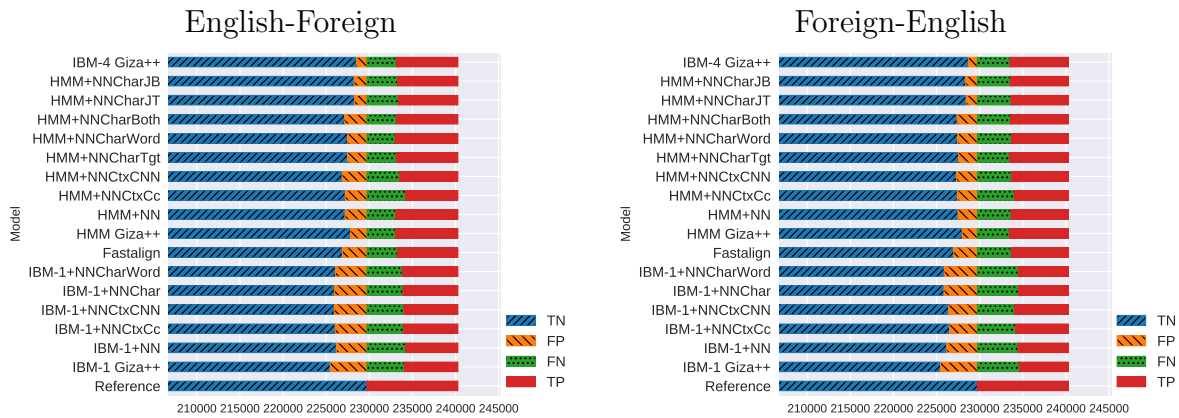


Figure B.2: Results of our neural models: Alignment links for English-German

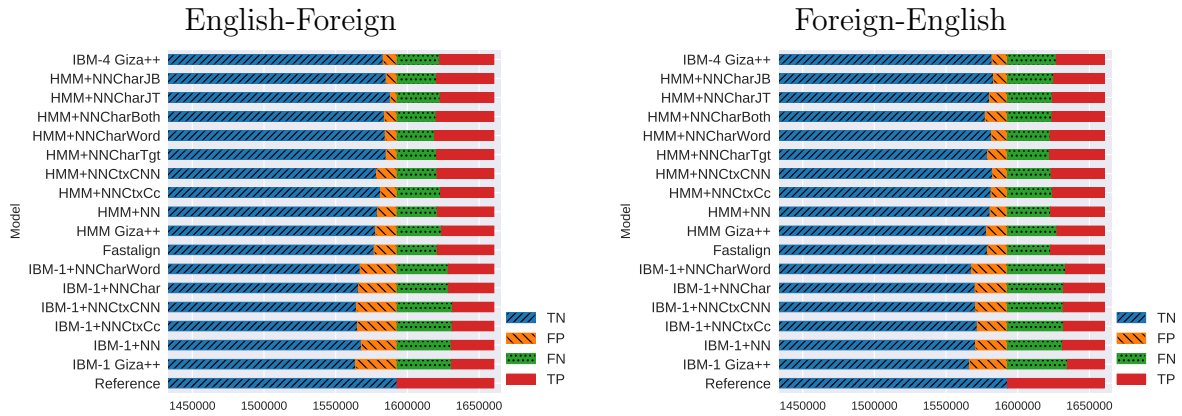


Figure B.3: Results of our neural models: Alignment links for English-Czech

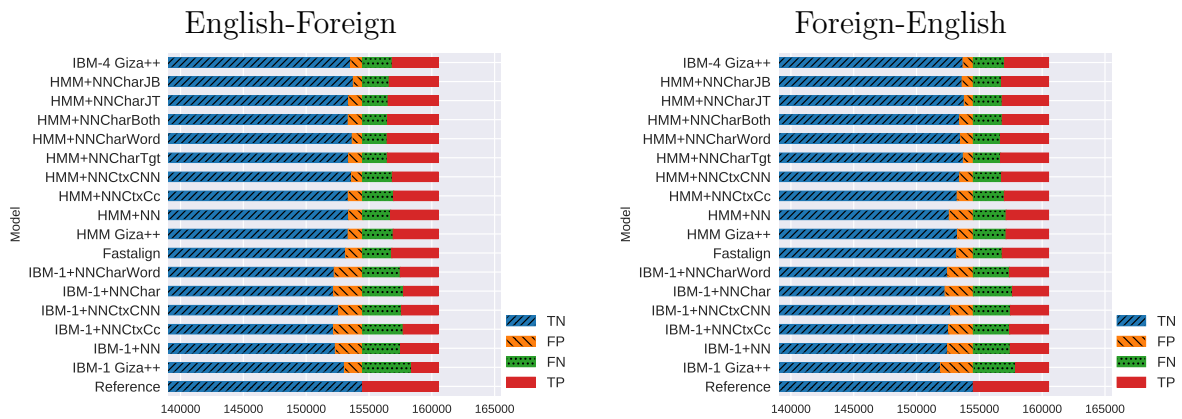


Figure B.4: Results of our neural models: Alignment links for English-Romanian

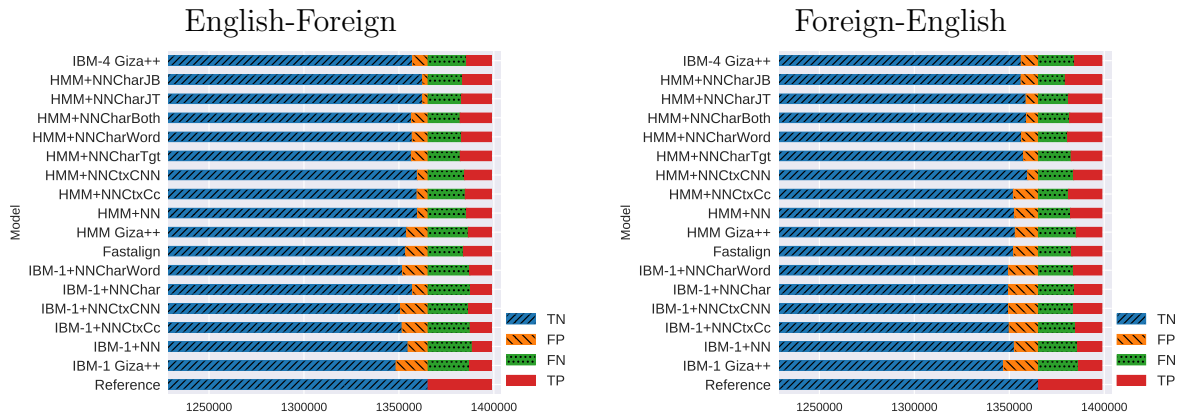


Figure B.5: Results of our neural models: Alignment links for English-Japanese

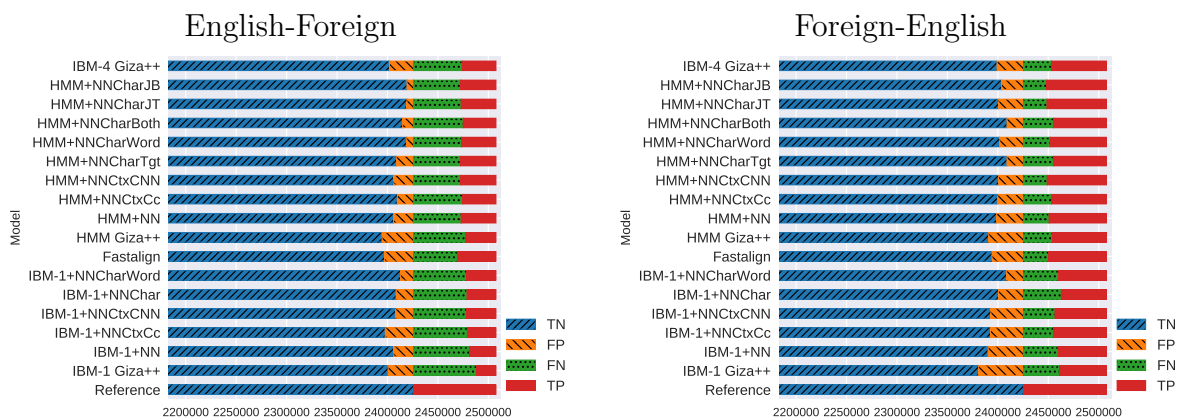


Figure B.6: Results of our neural models: Alignment links for English-Vietnamese

B.3 Unaligned words

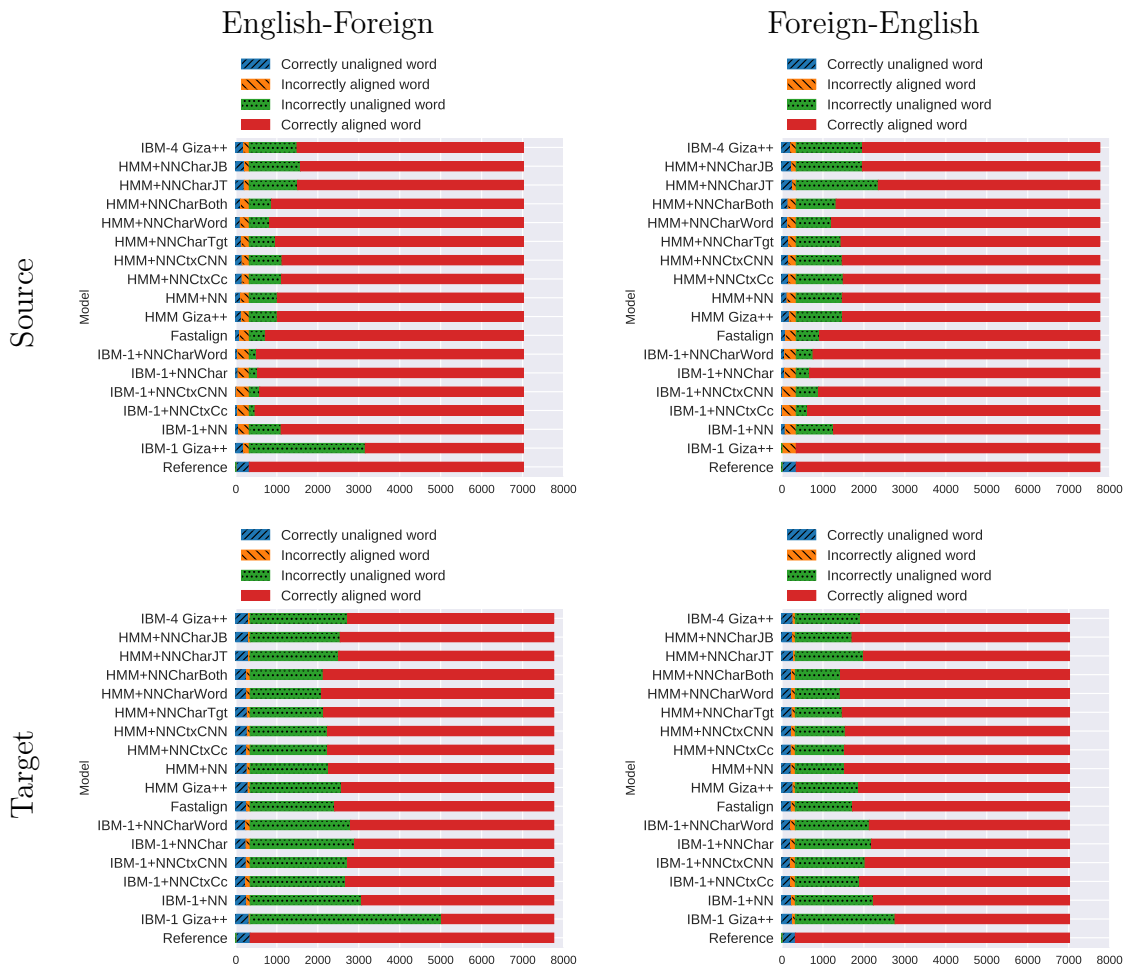


Figure B.7: Results of our neural models: Unaligned words for English-French

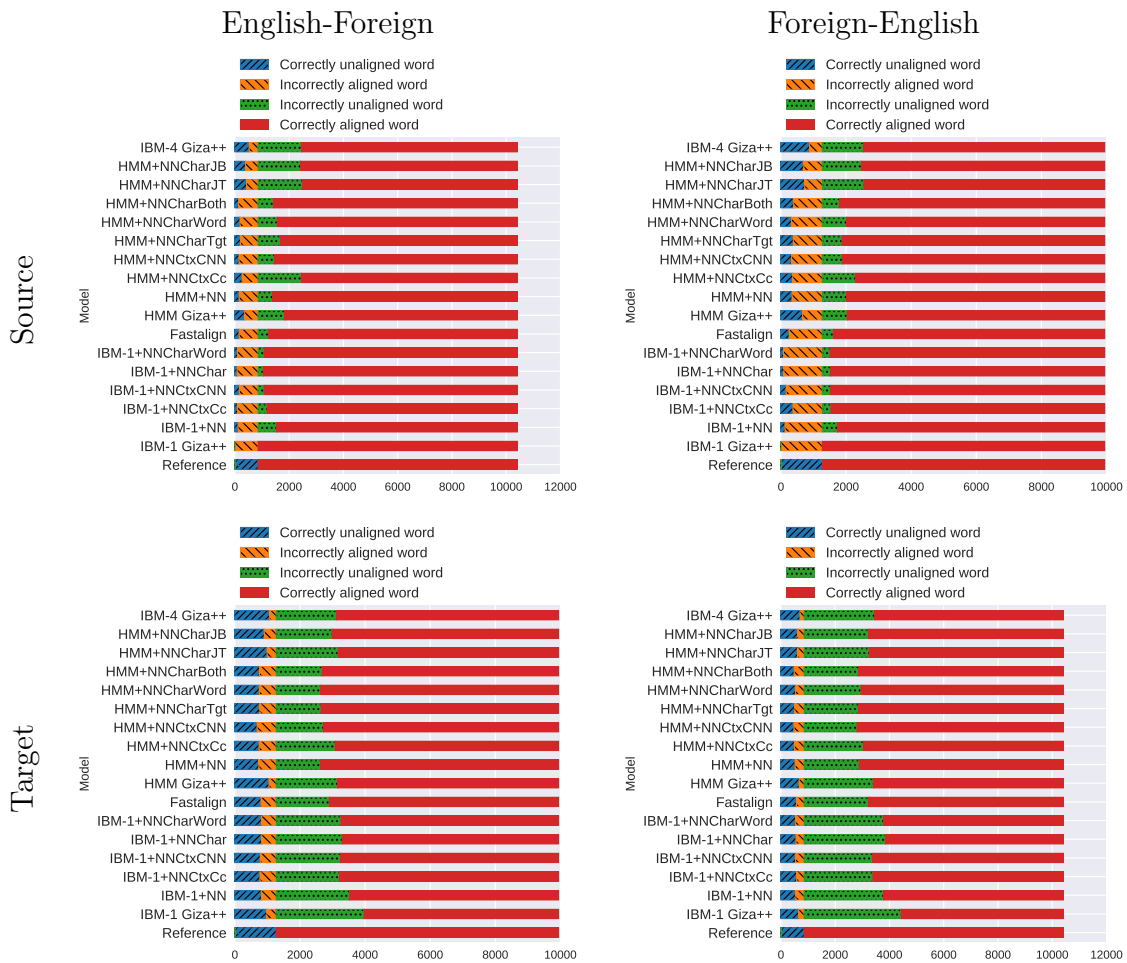


Figure B.8: Results of our neural models: Unaligned words for English-German

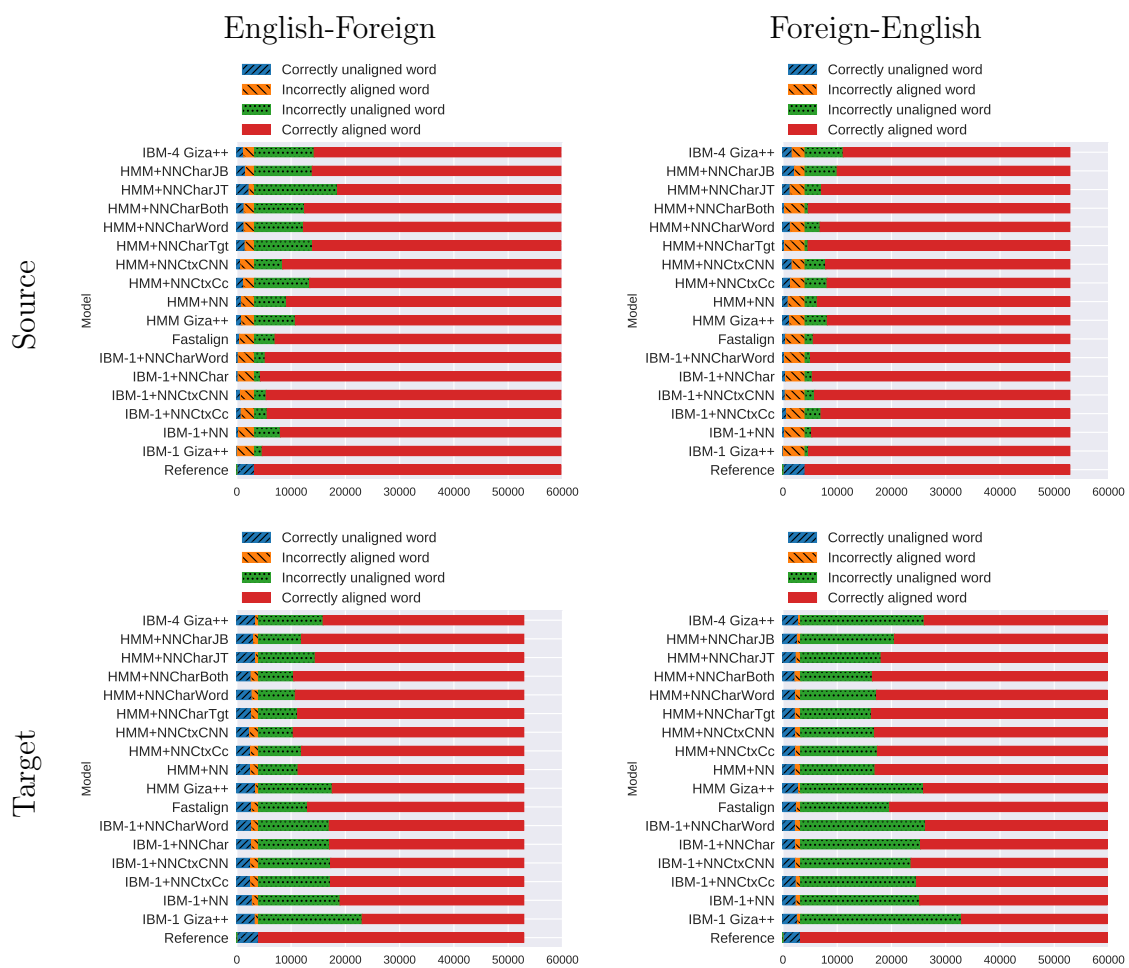


Figure B.9: Results of our neural models: Unaligned words for English-Czech

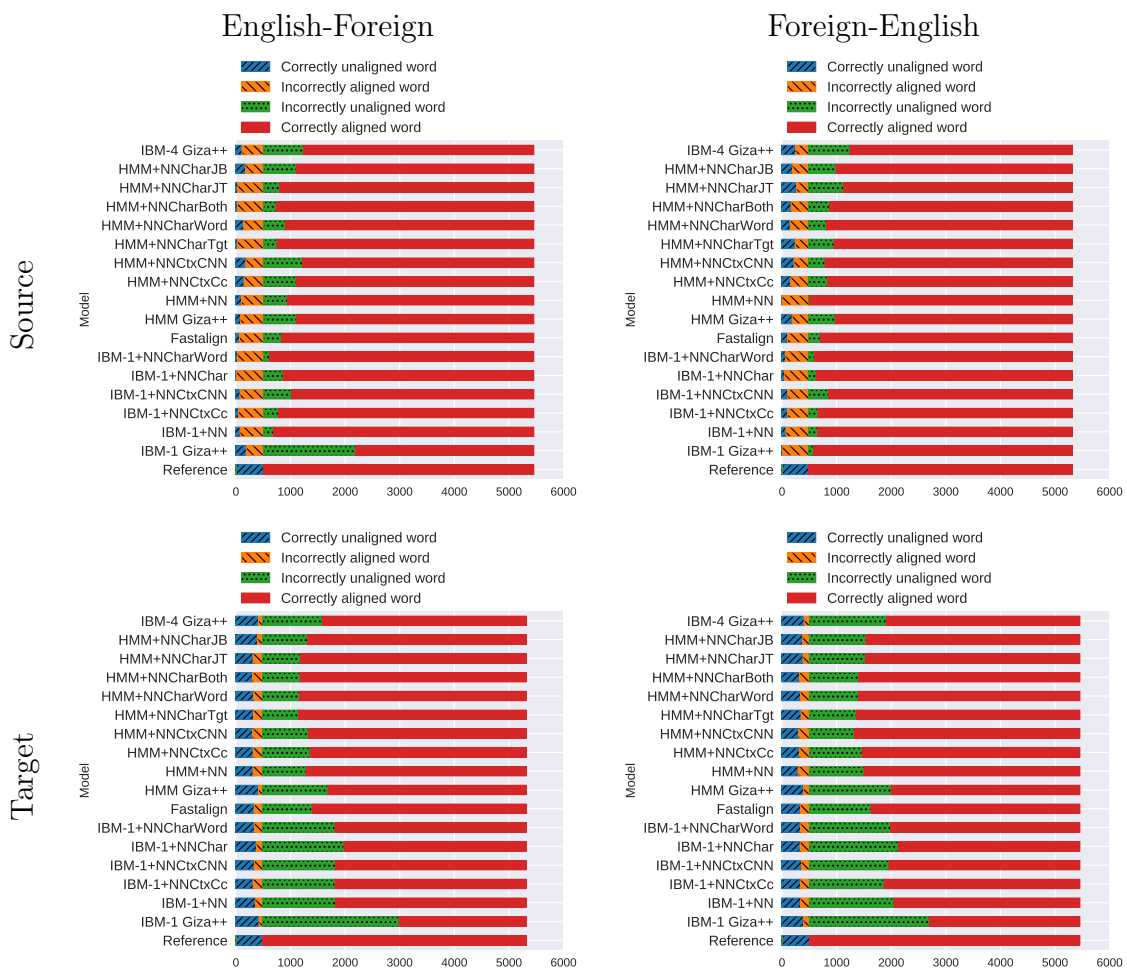


Figure B.10: Results of our neural models: Unaligned words for English-Romanian

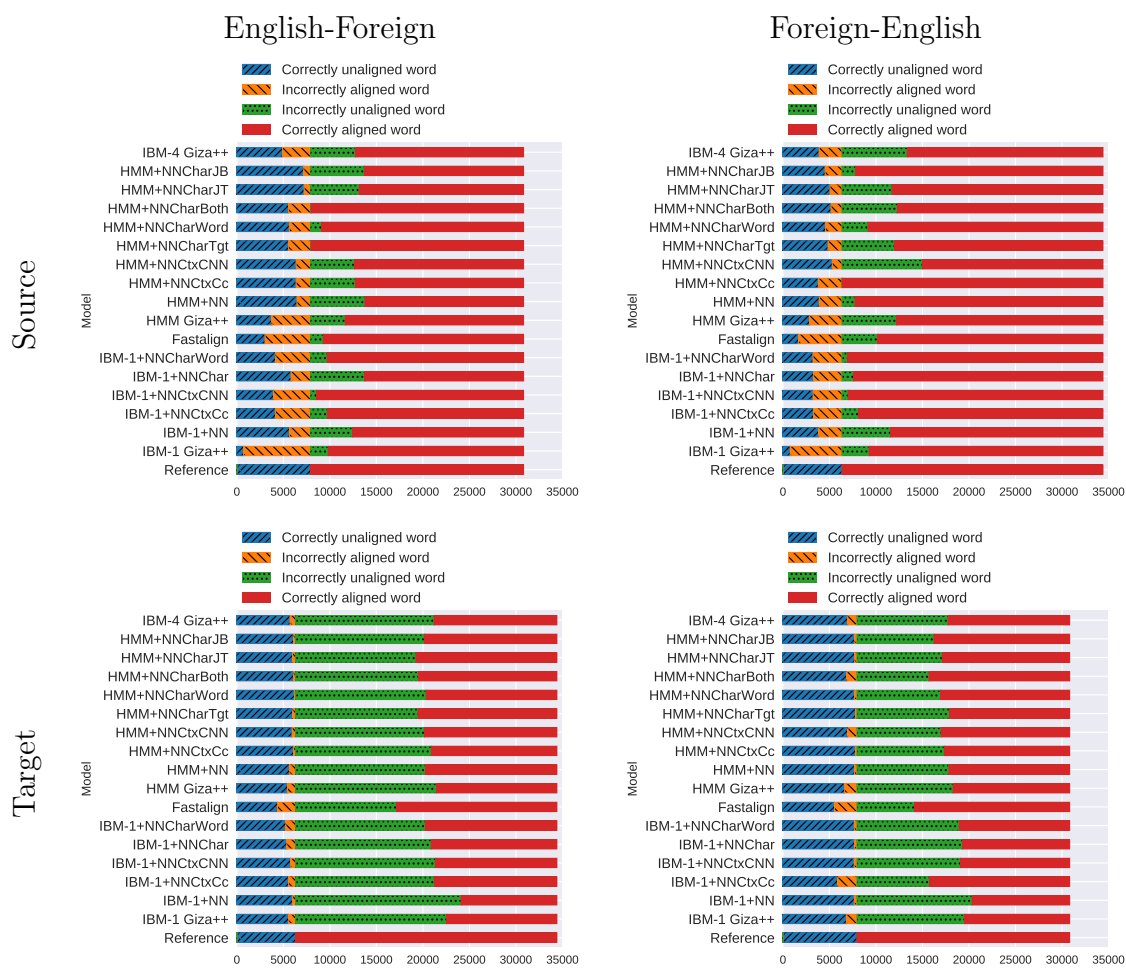


Figure B.11: Results of our neural models: Unaligned words for English-Japanese

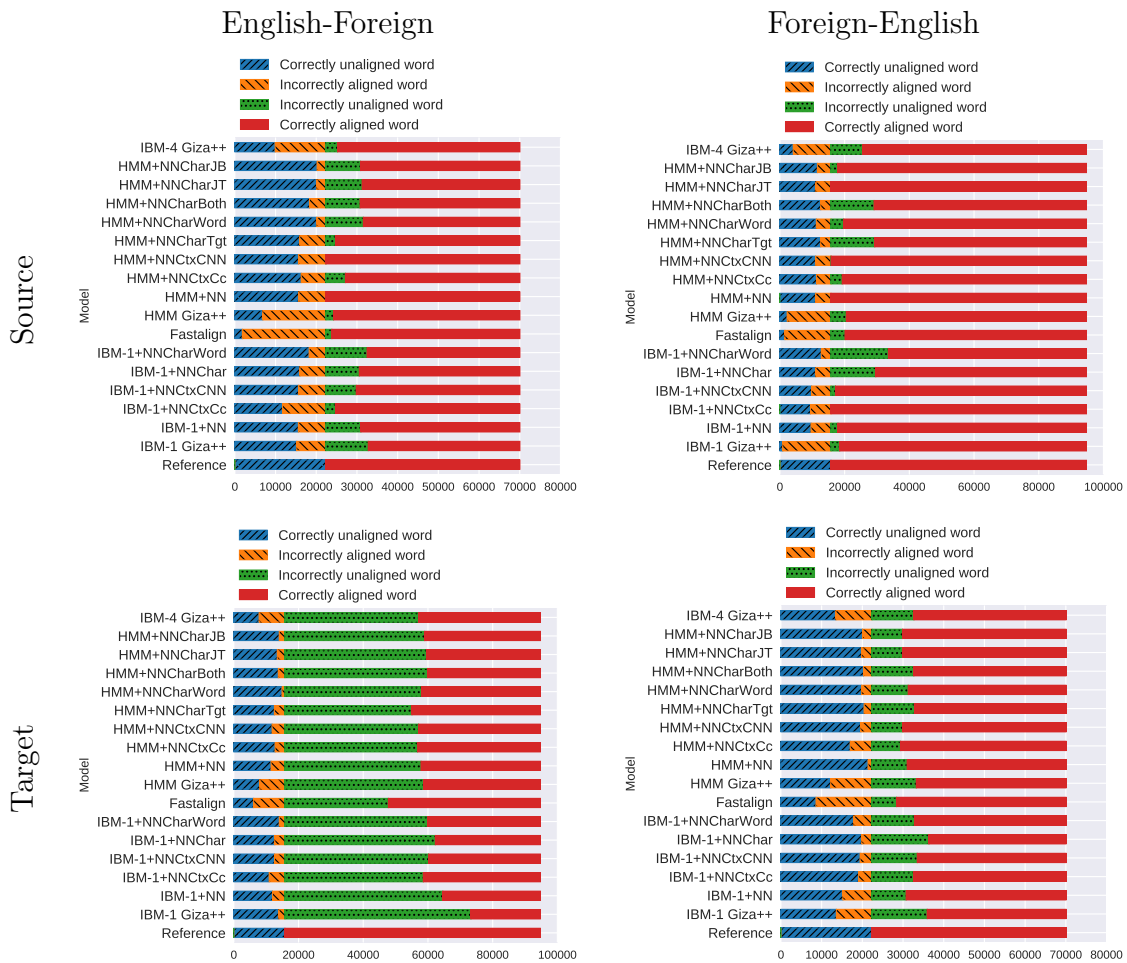


Figure B.12: Results of our neural models: Unaligned words for English-Vietnamese

B.4 Alignment types

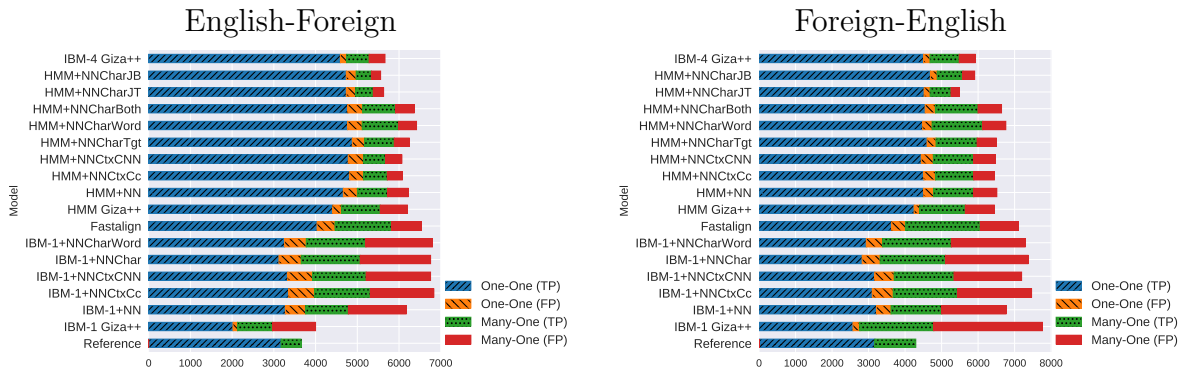


Figure B.13: Results of our neural models: Alignment types for English-French

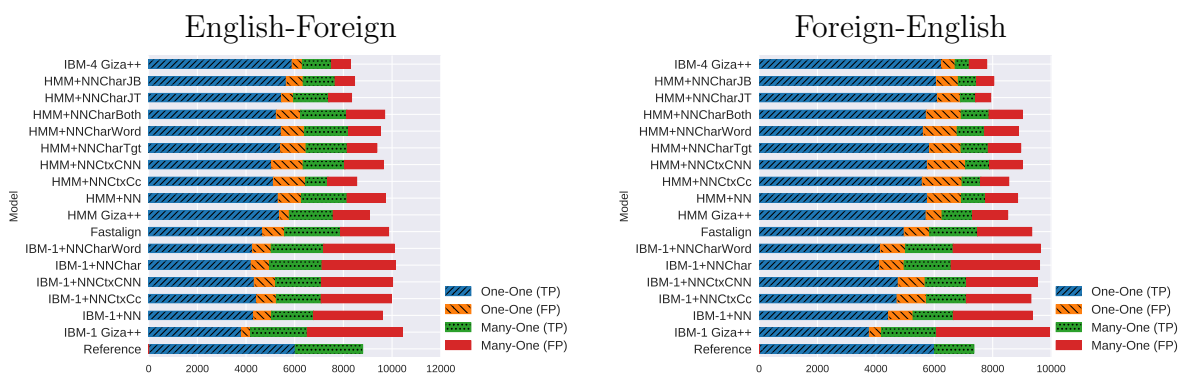


Figure B.14: Results of our neural models: Alignment types for English-German

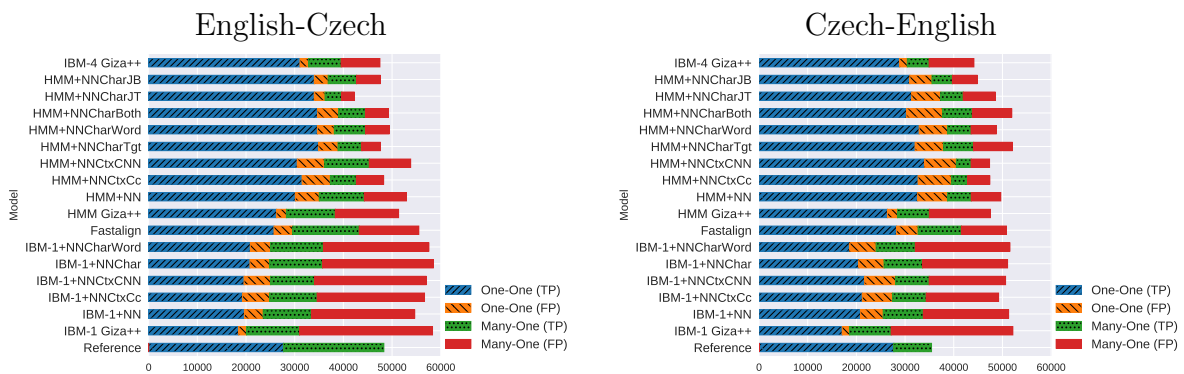


Figure B.15: Results of our neural models: Alignment types for English-Czech



Figure B.16: Results of our neural models: Alignment types for English-Romanian

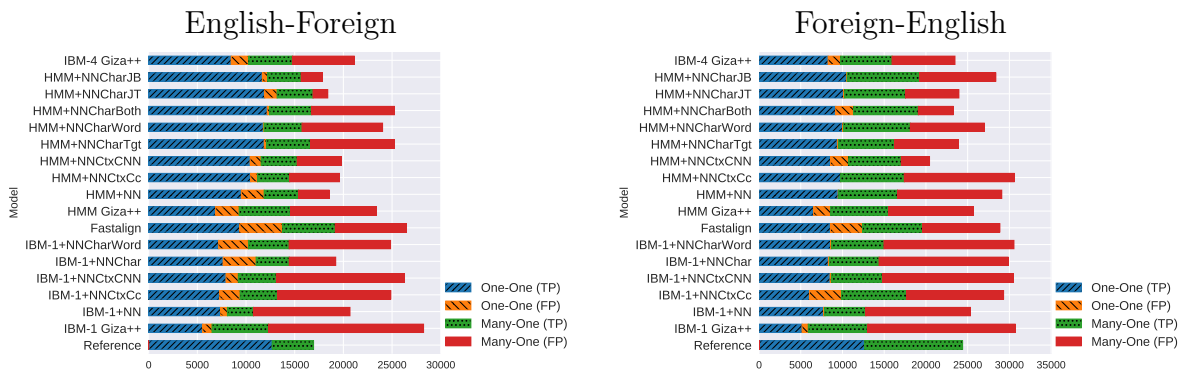


Figure B.17: Results of our neural models: Alignment types for English-Japanese

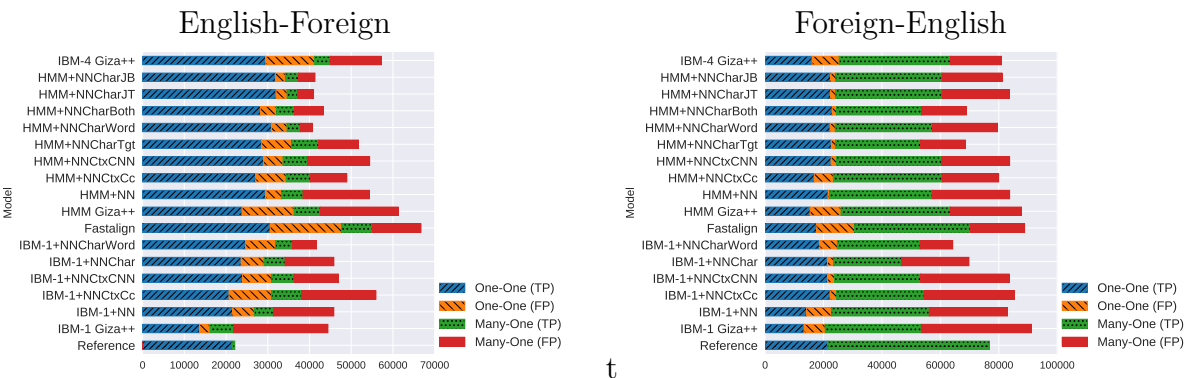


Figure B.18: Results of our neural models: Alignment types for English-Vietnamese

B.5 Jump widths

B.5.1 Jump width statistics

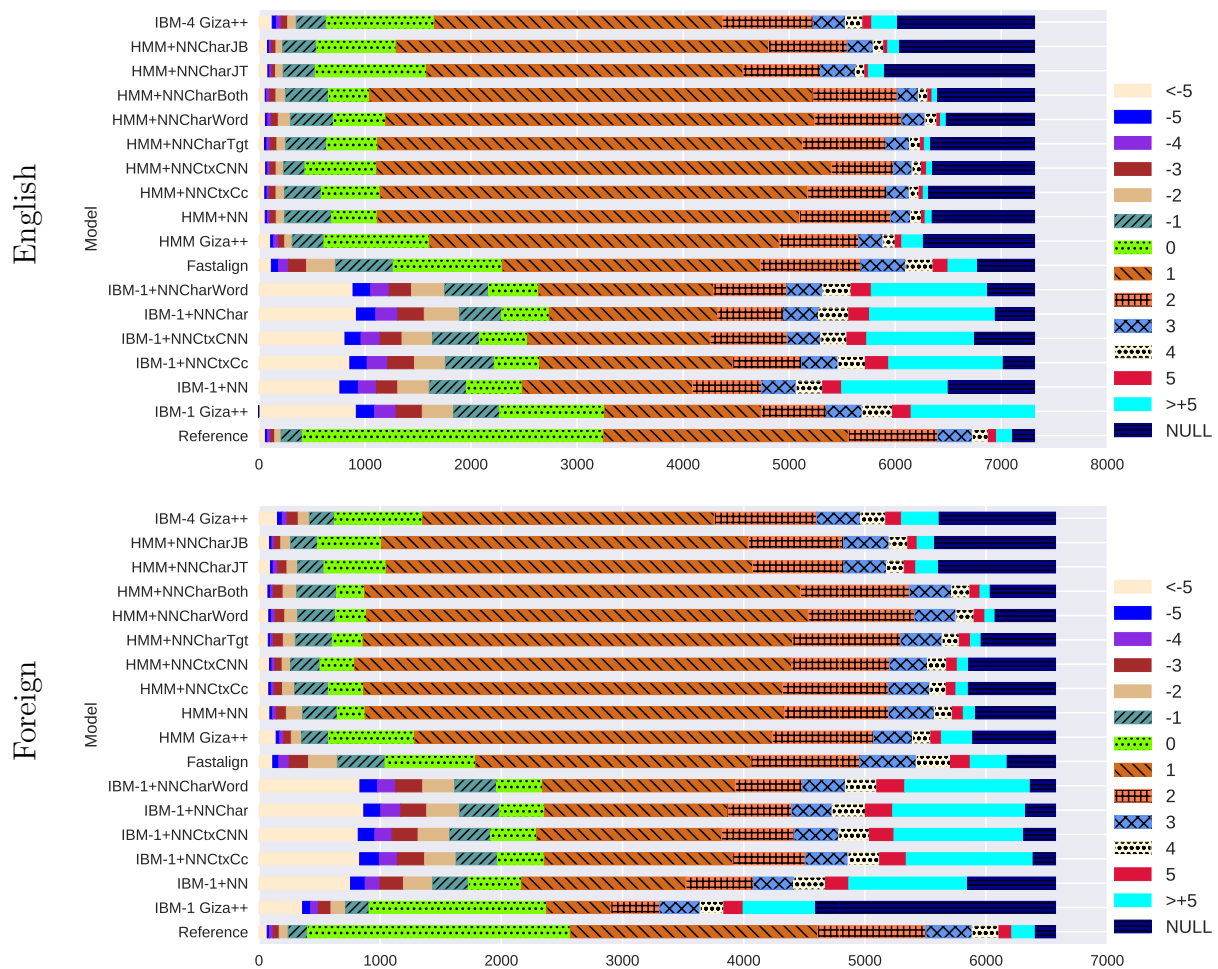


Figure B.19: Neural models: Jump widths for English words in the direction French-English and jump widths of foreign words in direction English-French

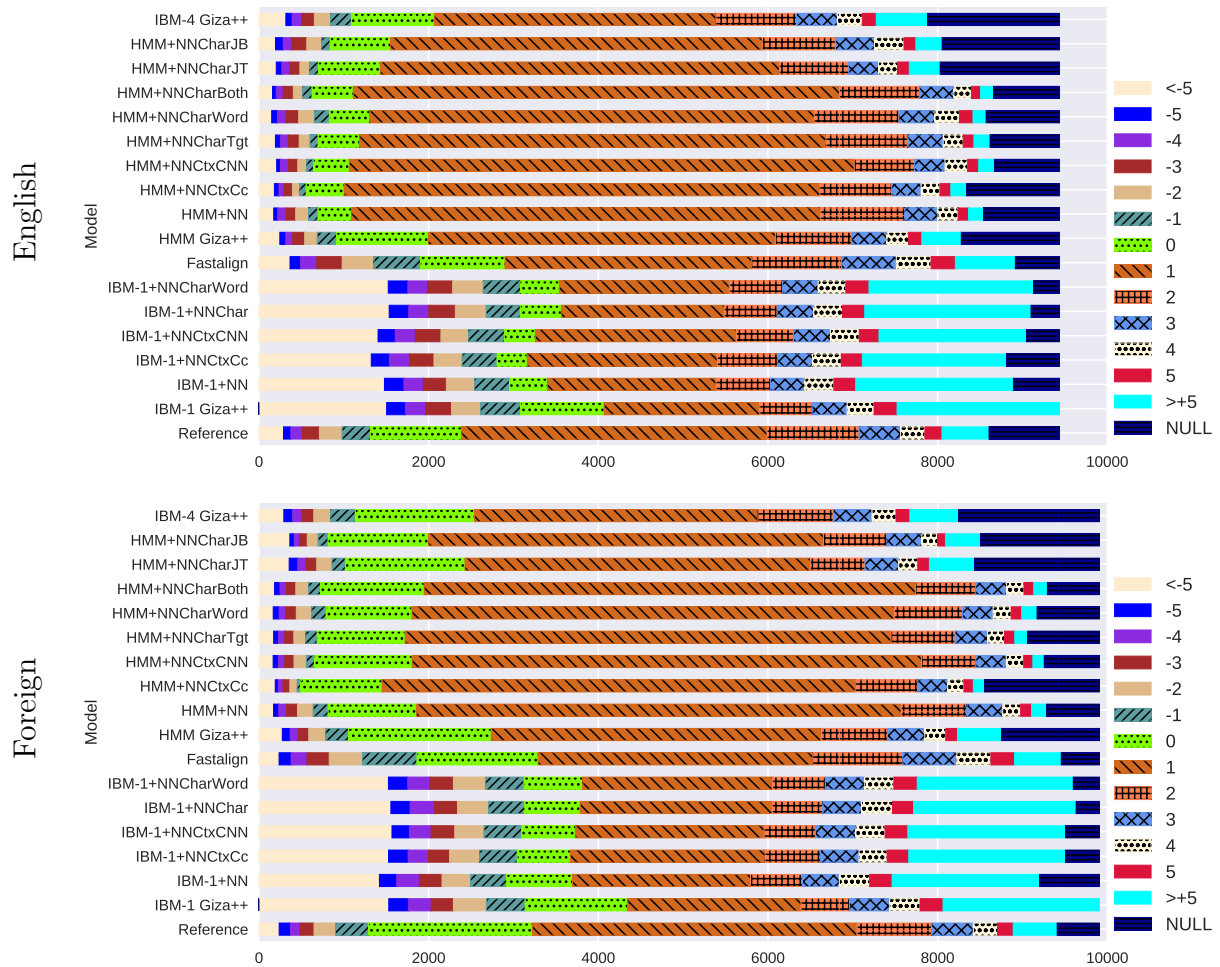


Figure B.20: Neural models: Jump widths for English words in the direction German-English and jump widths of foreign words in direction English-German

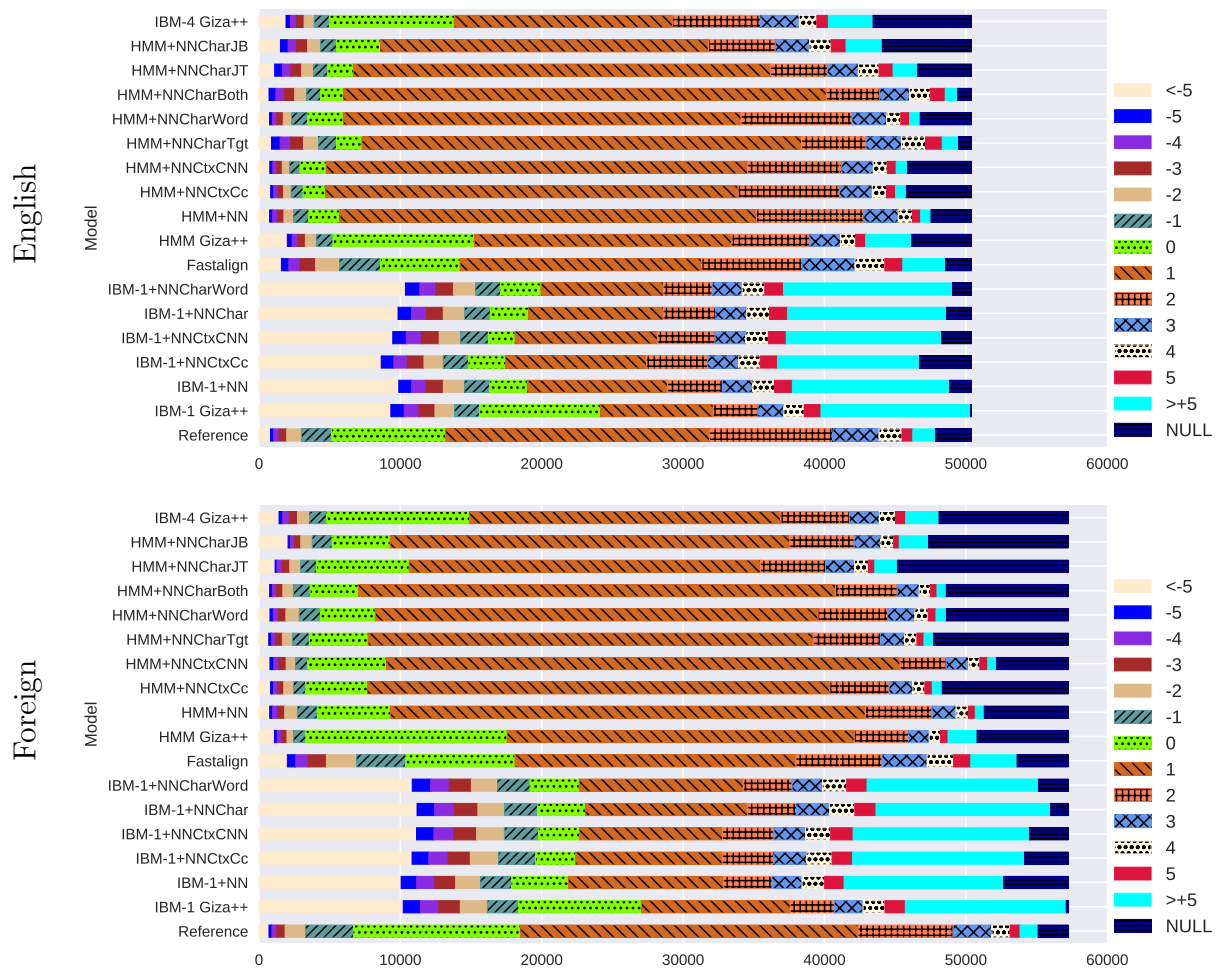


Figure B.21: Neural models: Jump widths for English words in the direction Czech-English and jump widths of foreign words in direction English-Czech

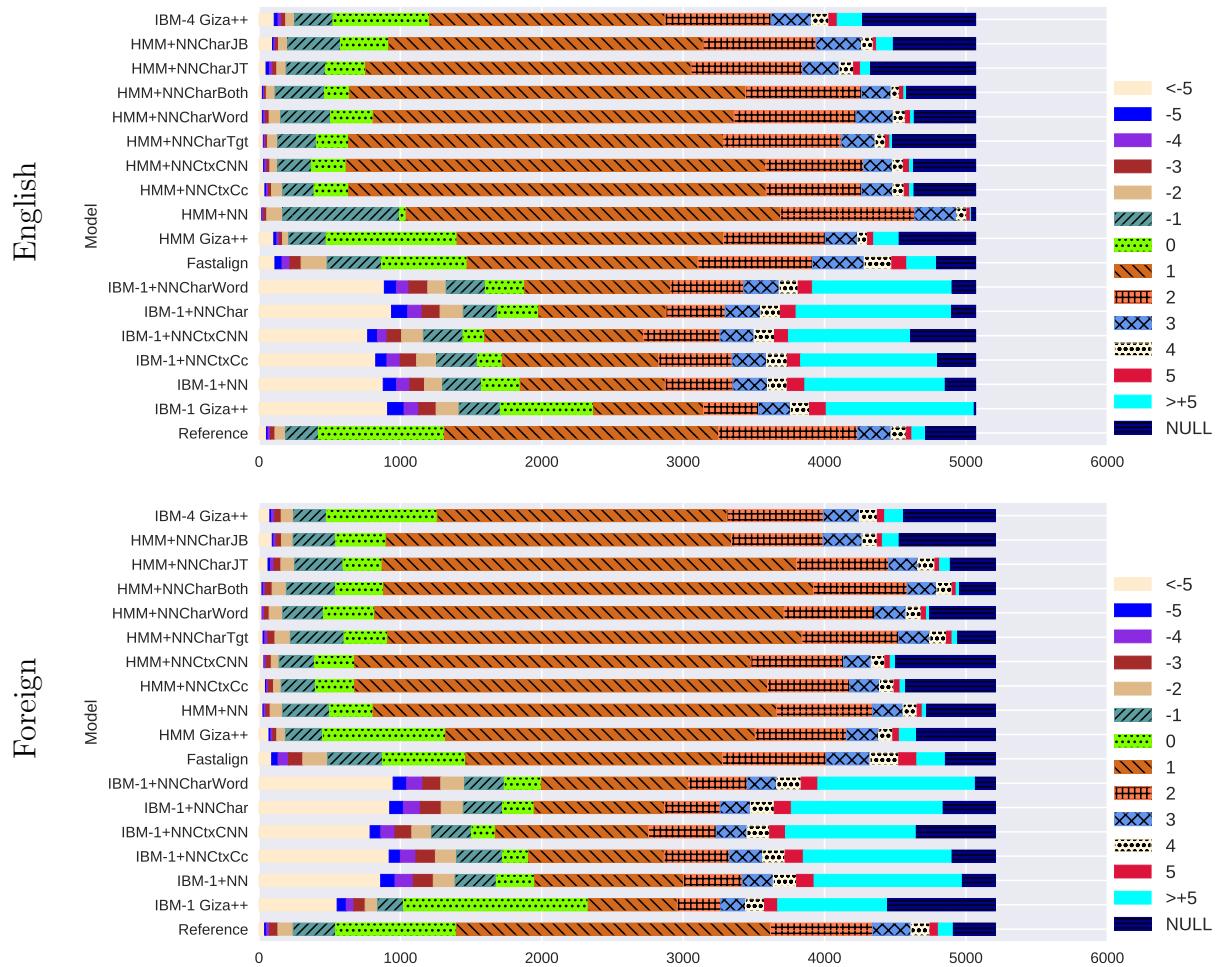


Figure B.22: Neural models: Jump widths for English words in the direction Romanian-English and jump widths of foreign words in direction English-Romanian

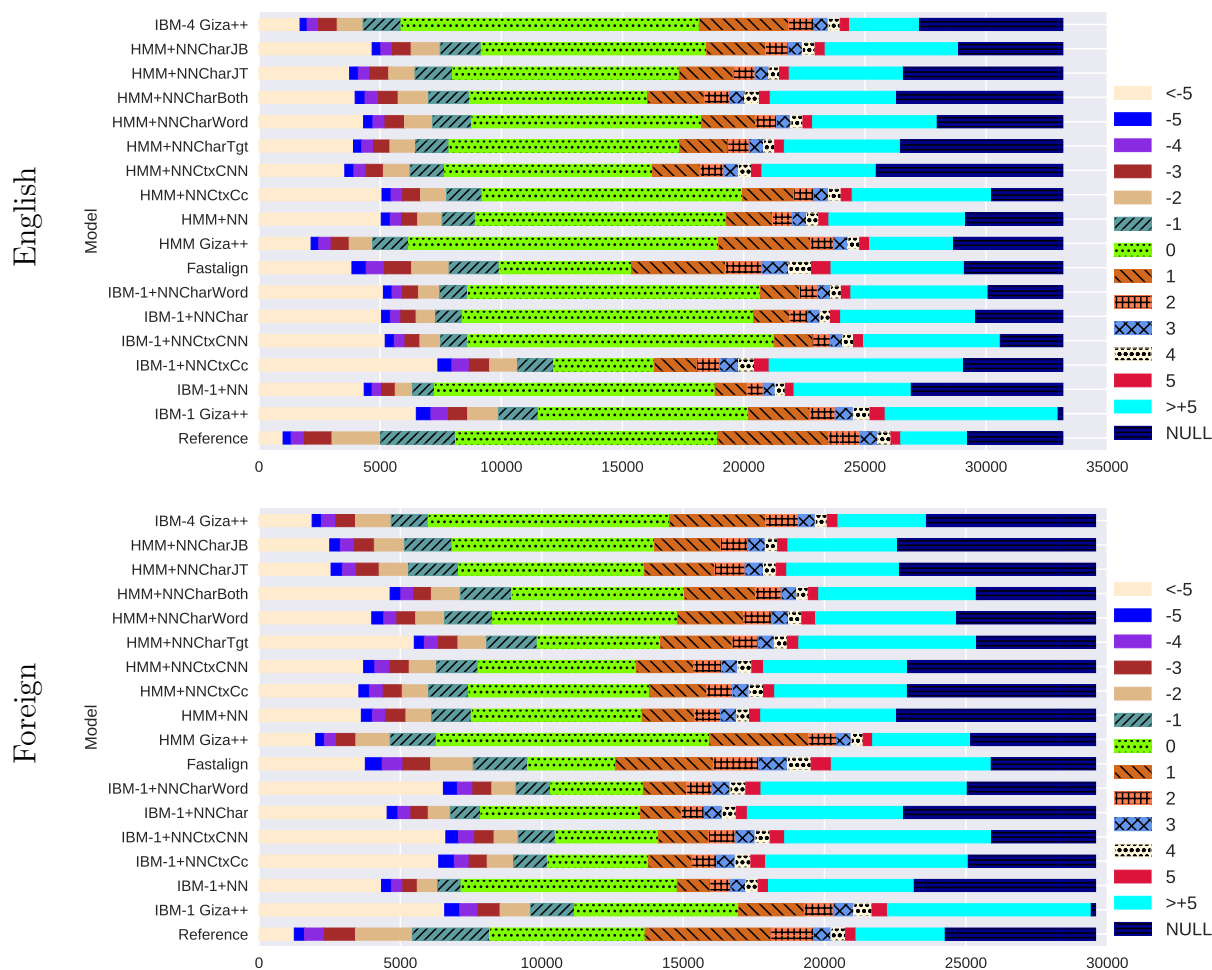


Figure B.23: Neural models: Jump widths for English words in the direction Japanese-English and jump widths of foreign words in direction English-Japanese

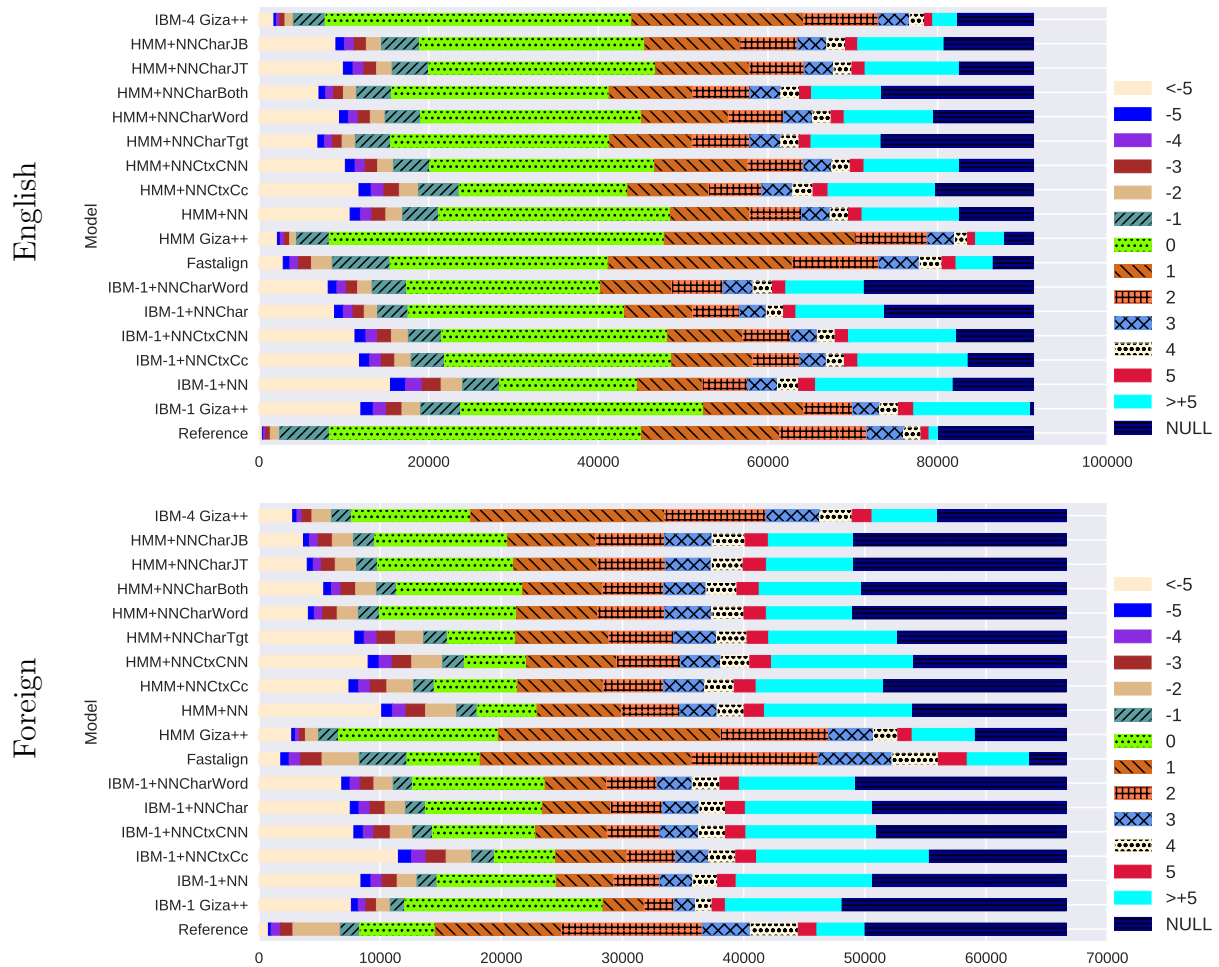
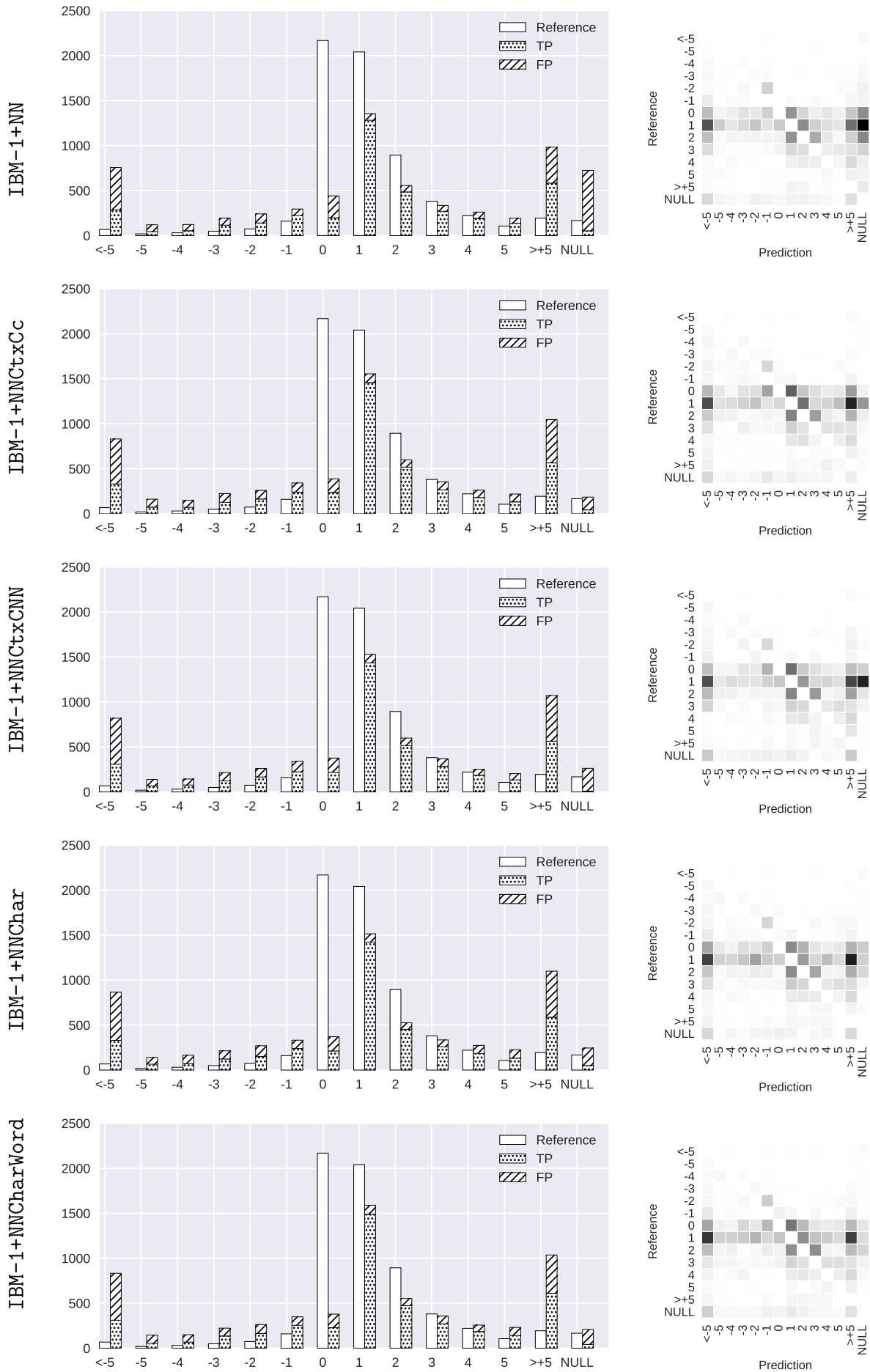
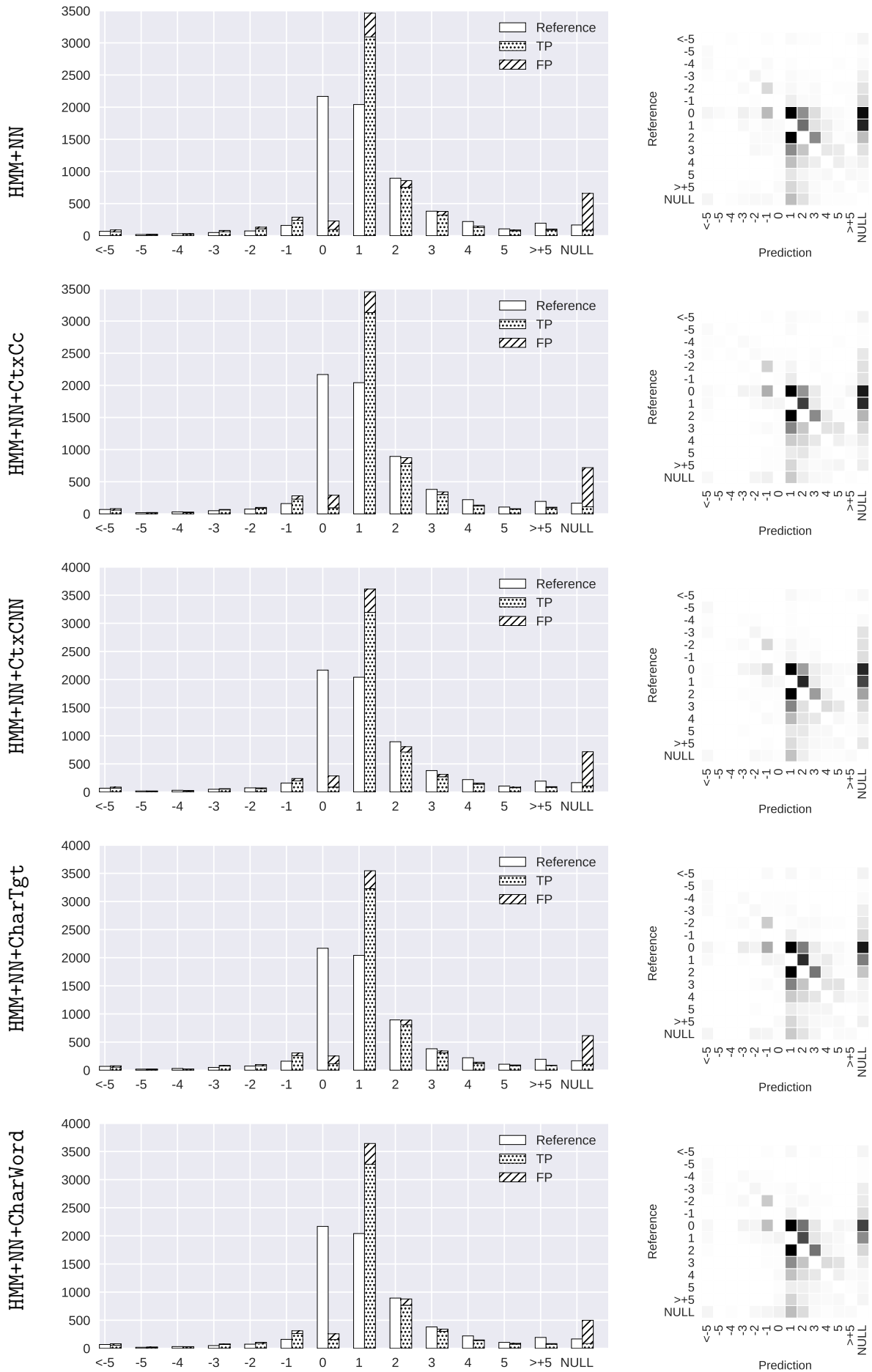


Figure B.24: Neural models: Jump widths for English words in the direction Vietnamese-English and jump widths of foreign words in direction English-Vietnamese

B.5.2 Correct and incorrect jumps





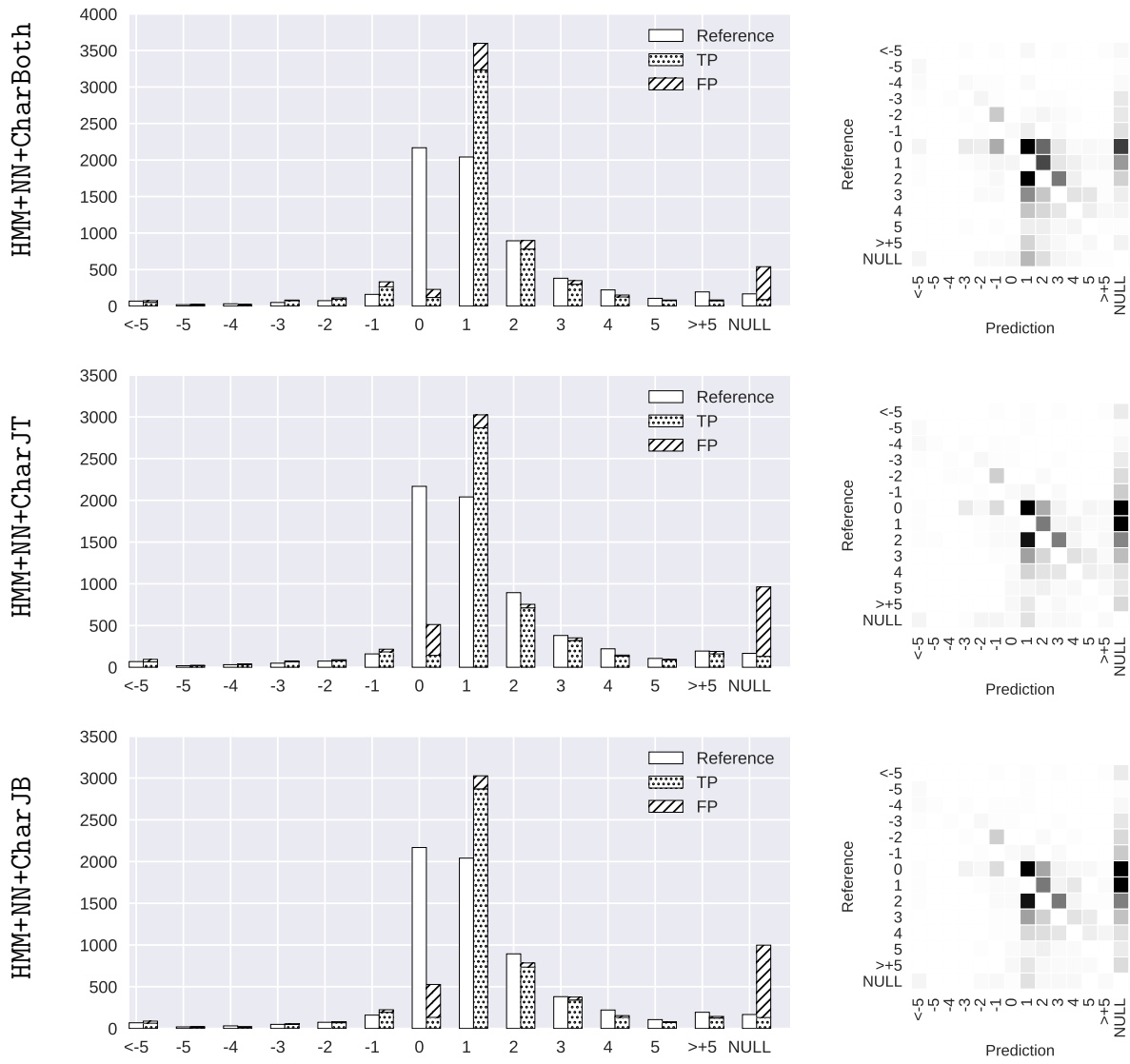
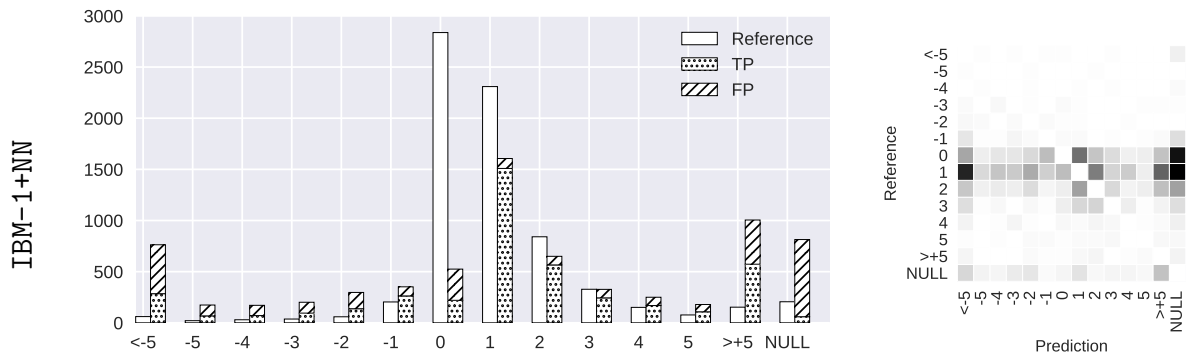
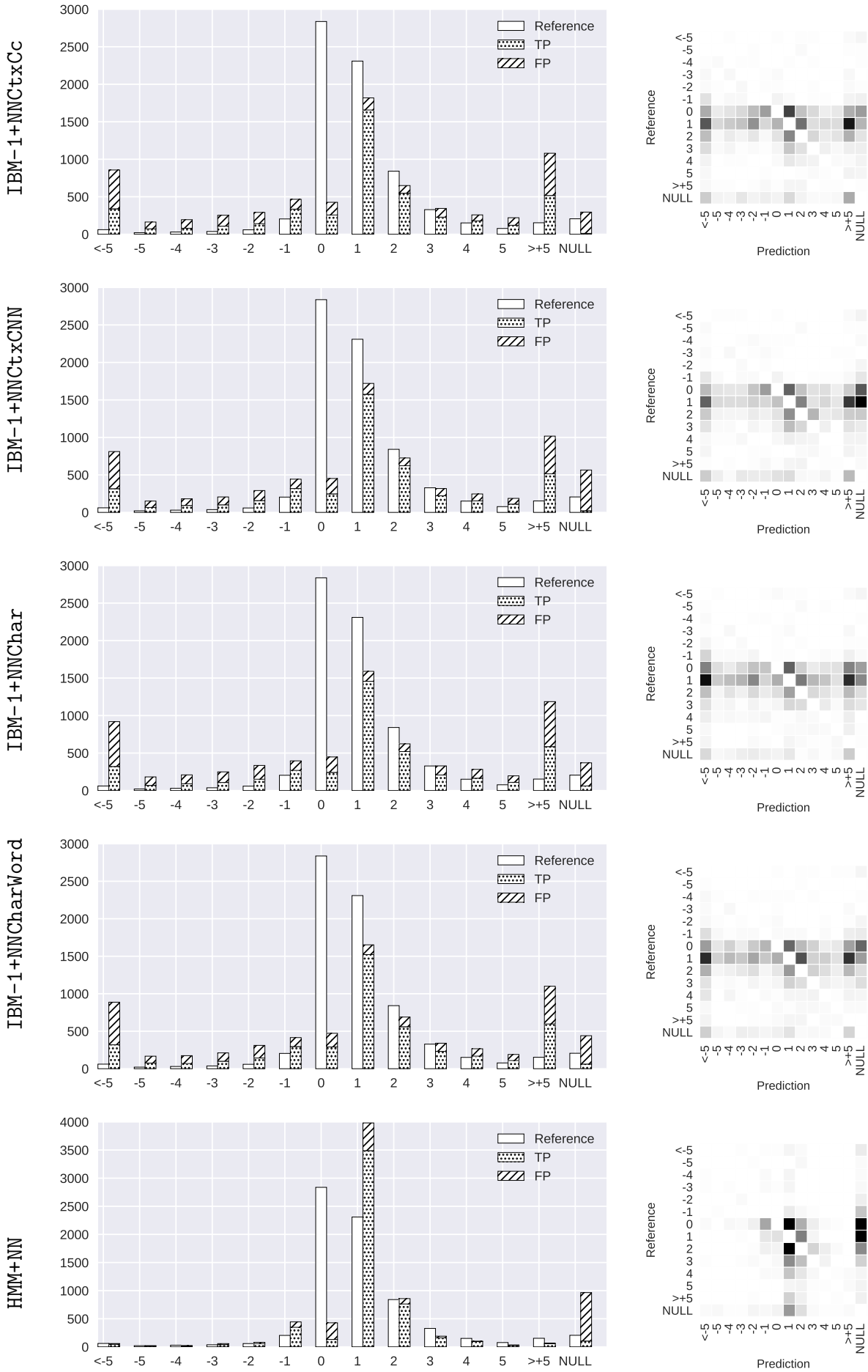
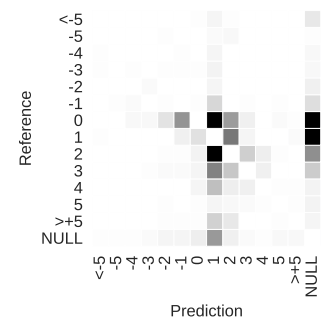
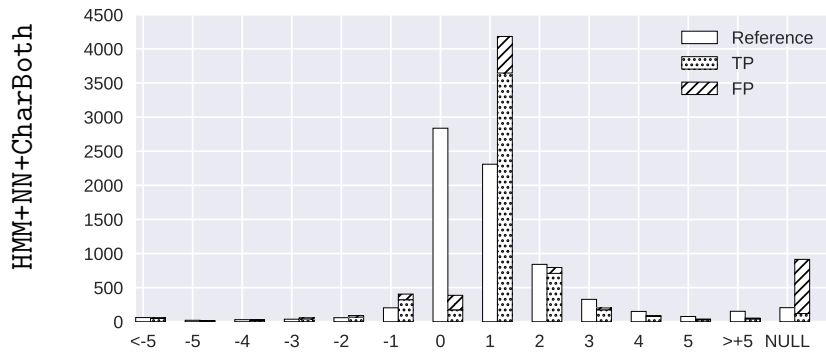
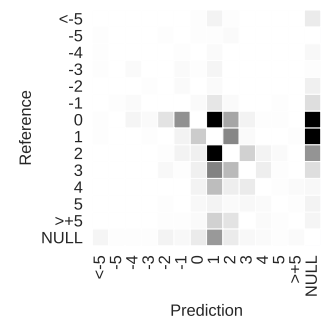
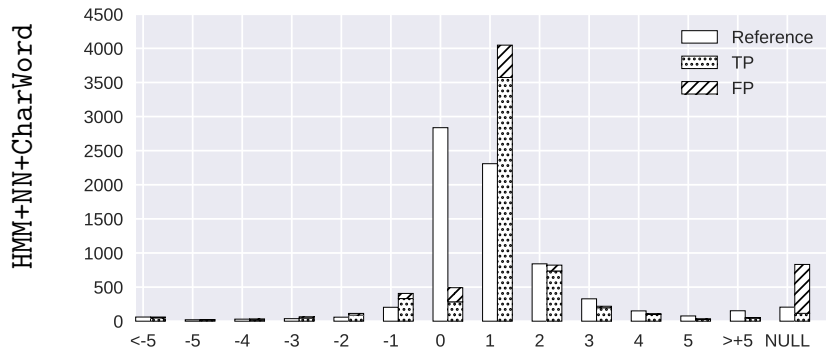
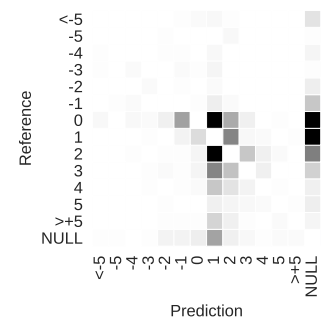
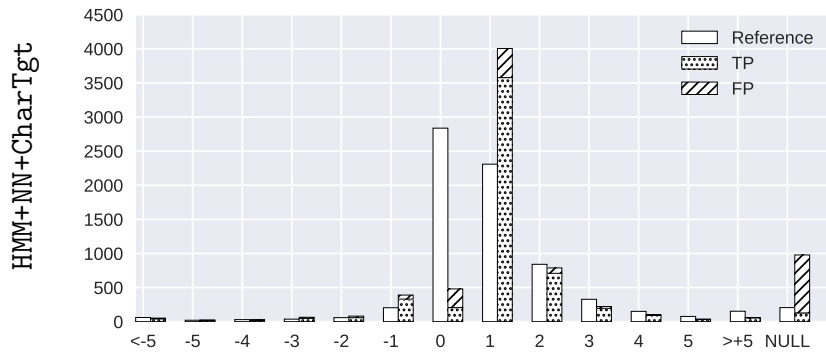
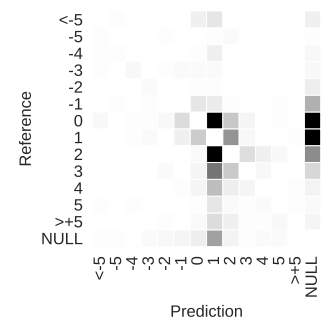
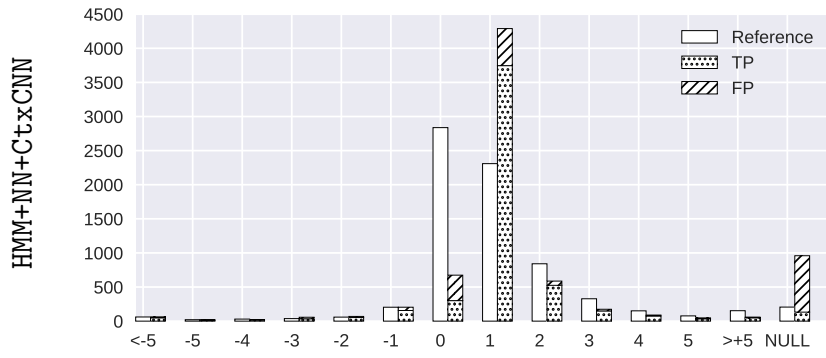
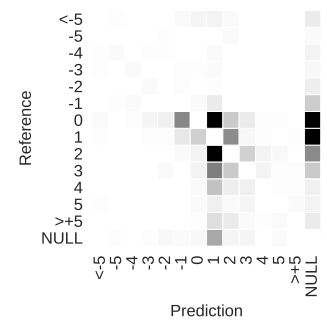
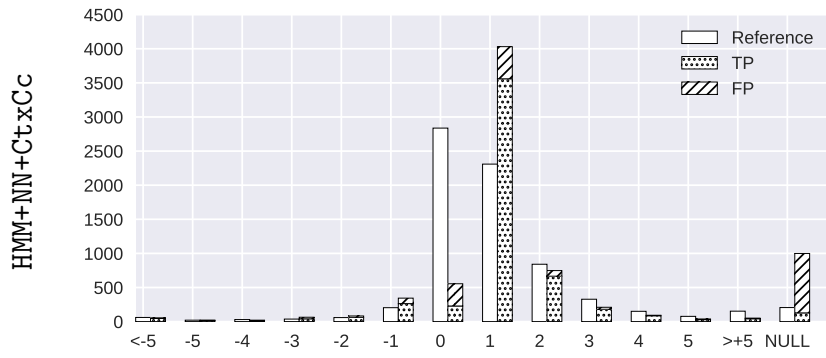


Figure B.25: Neural models for the direction English-French: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.







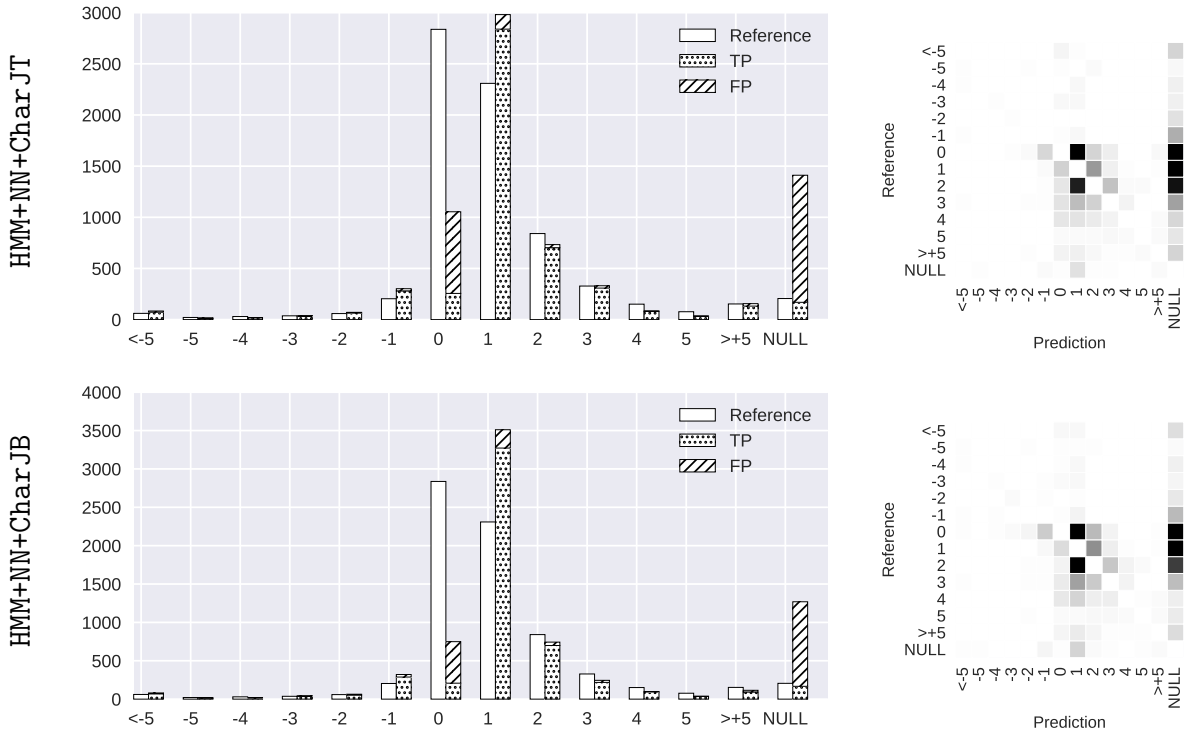
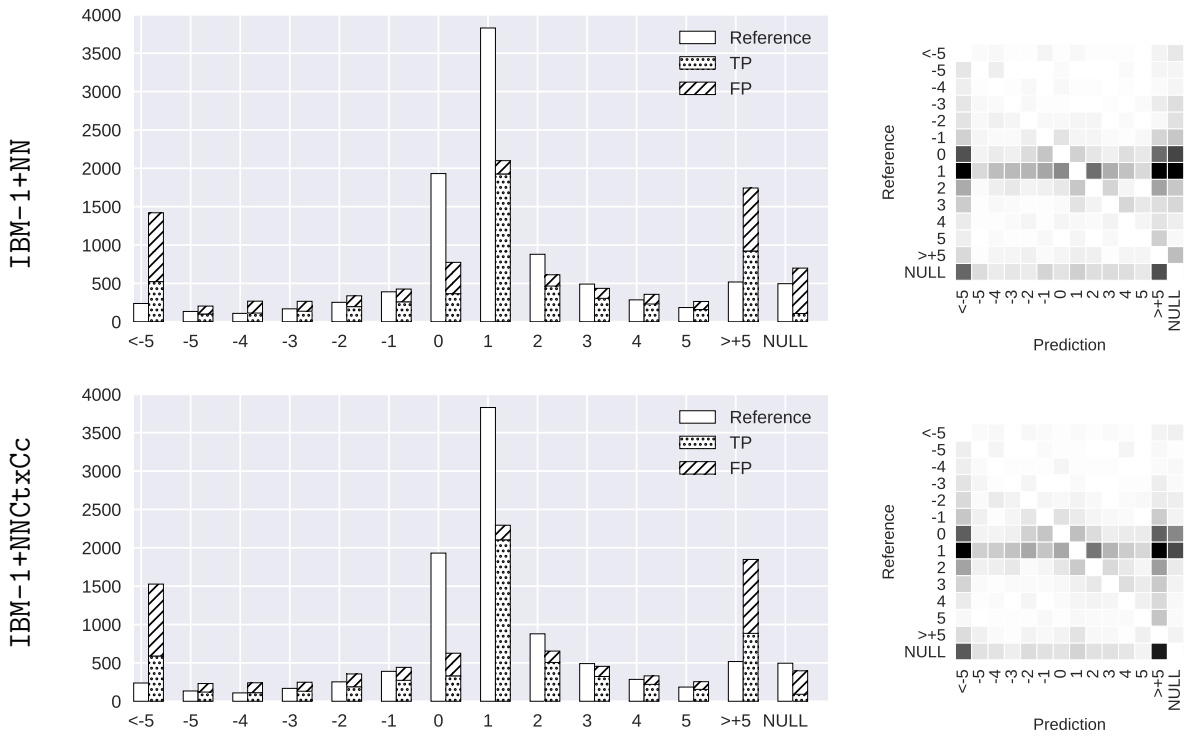
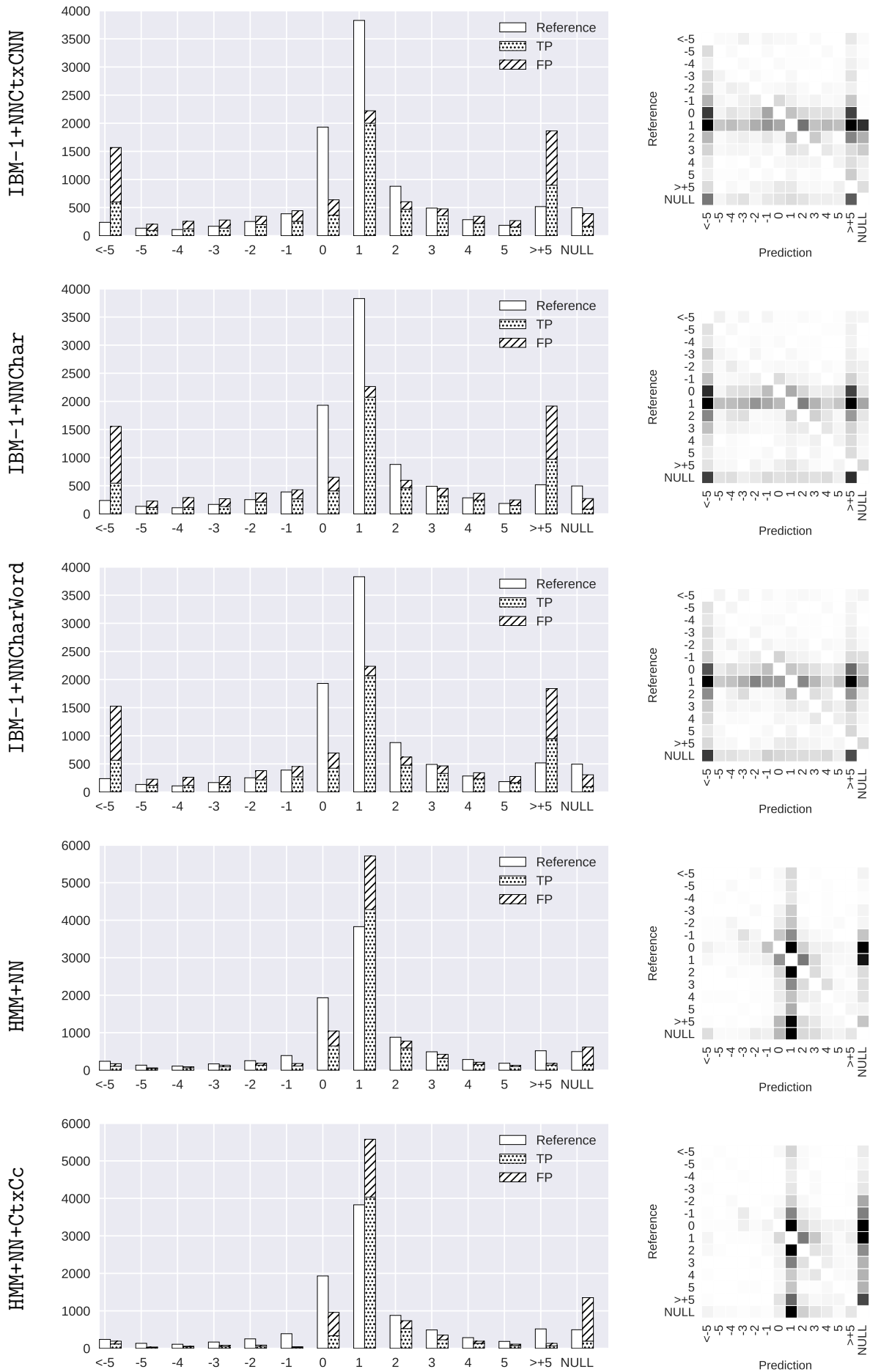
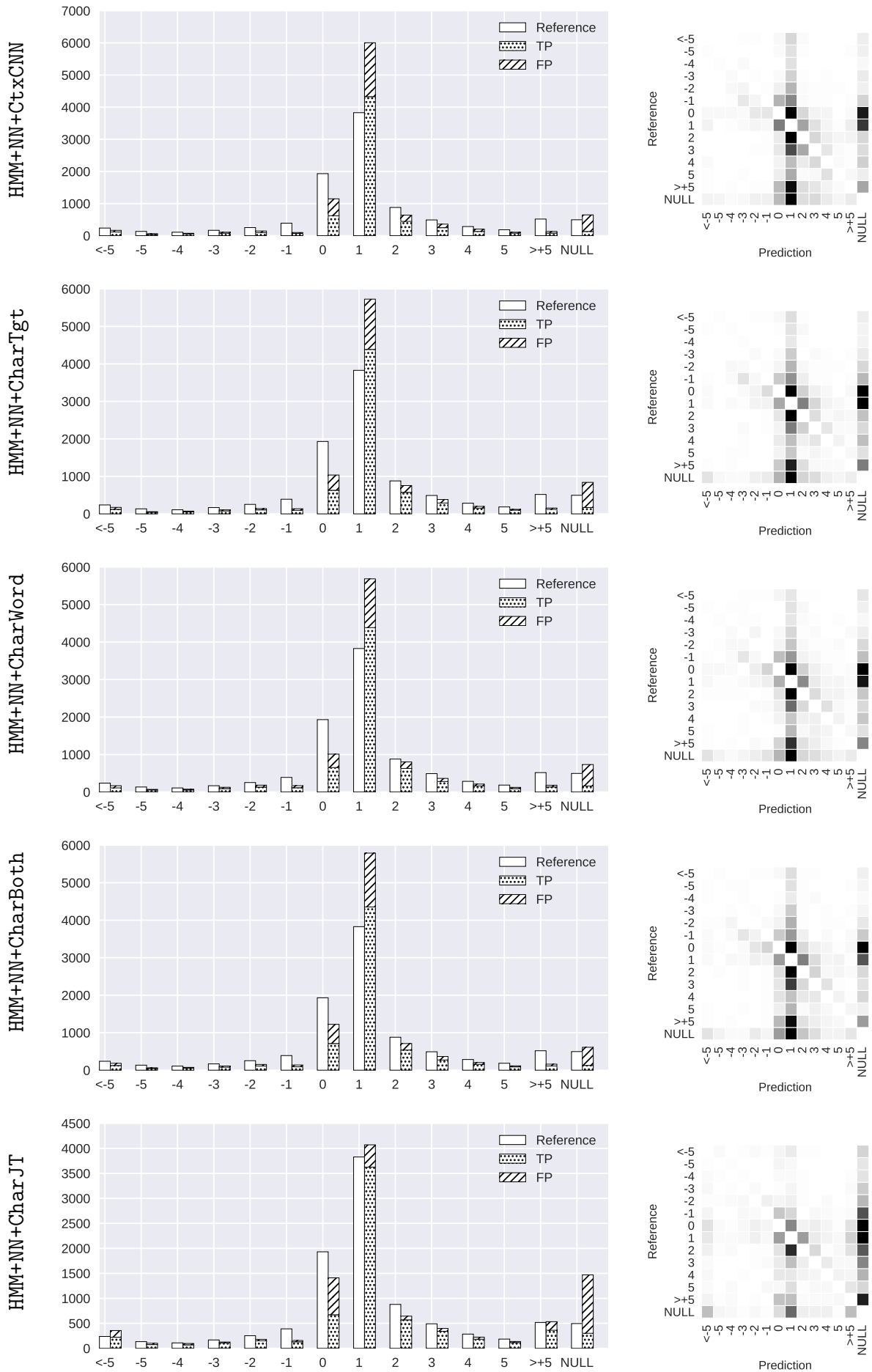


Figure B.26: Neural models for the direction French-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.







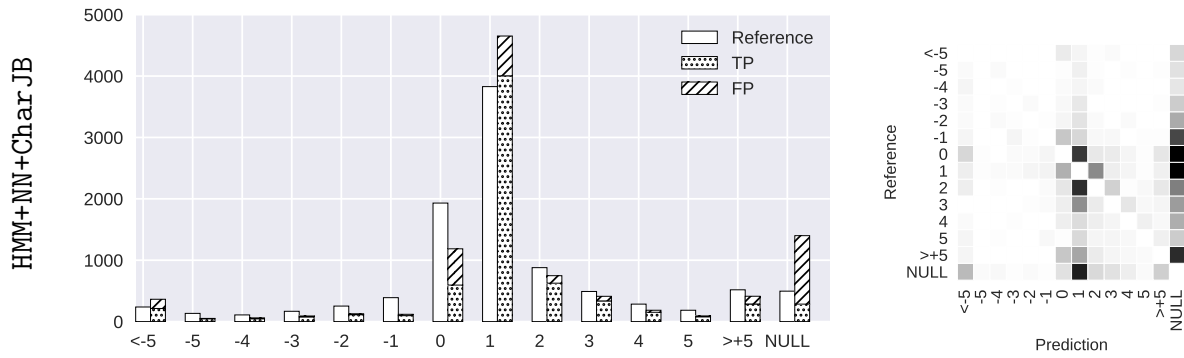
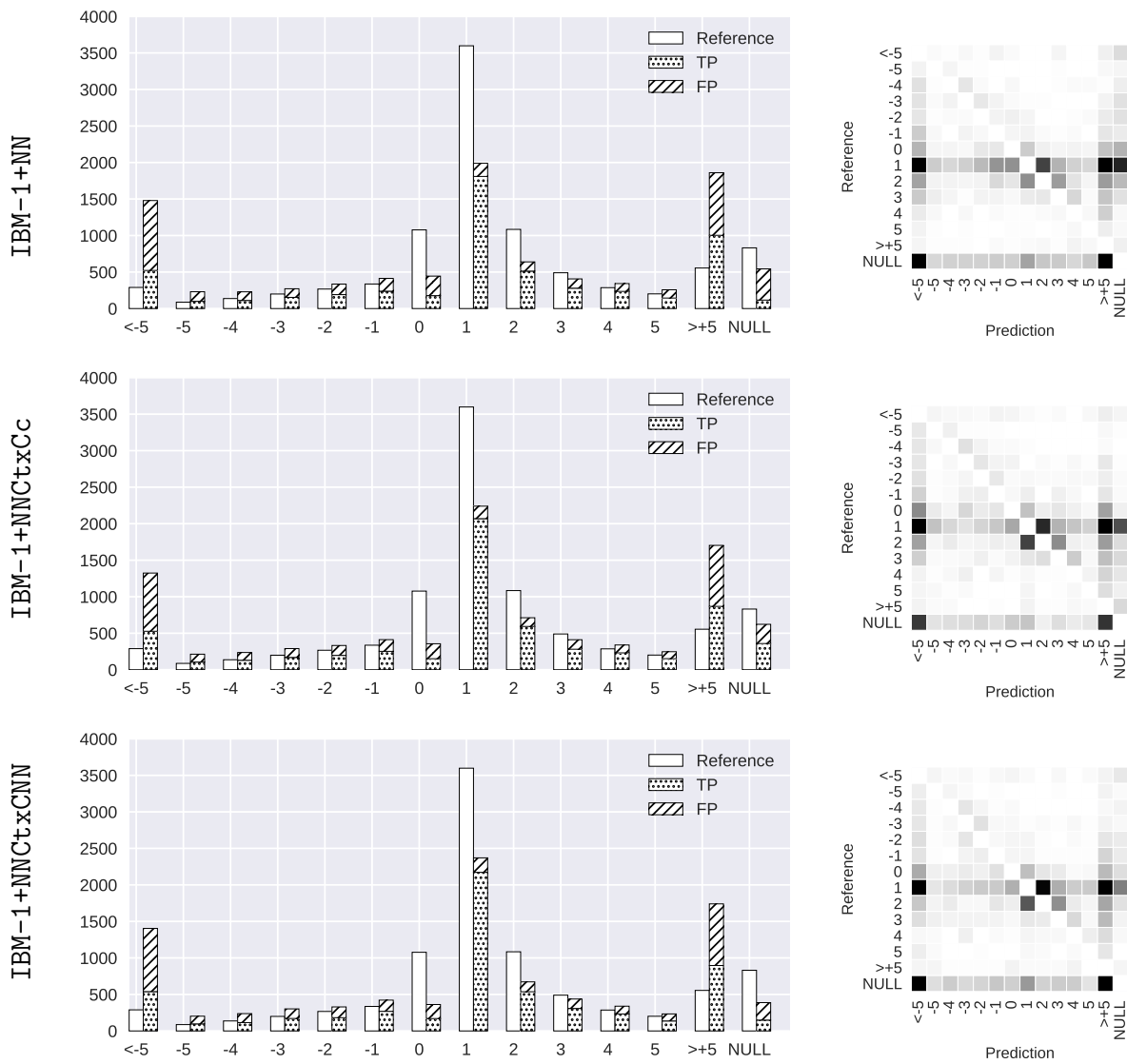
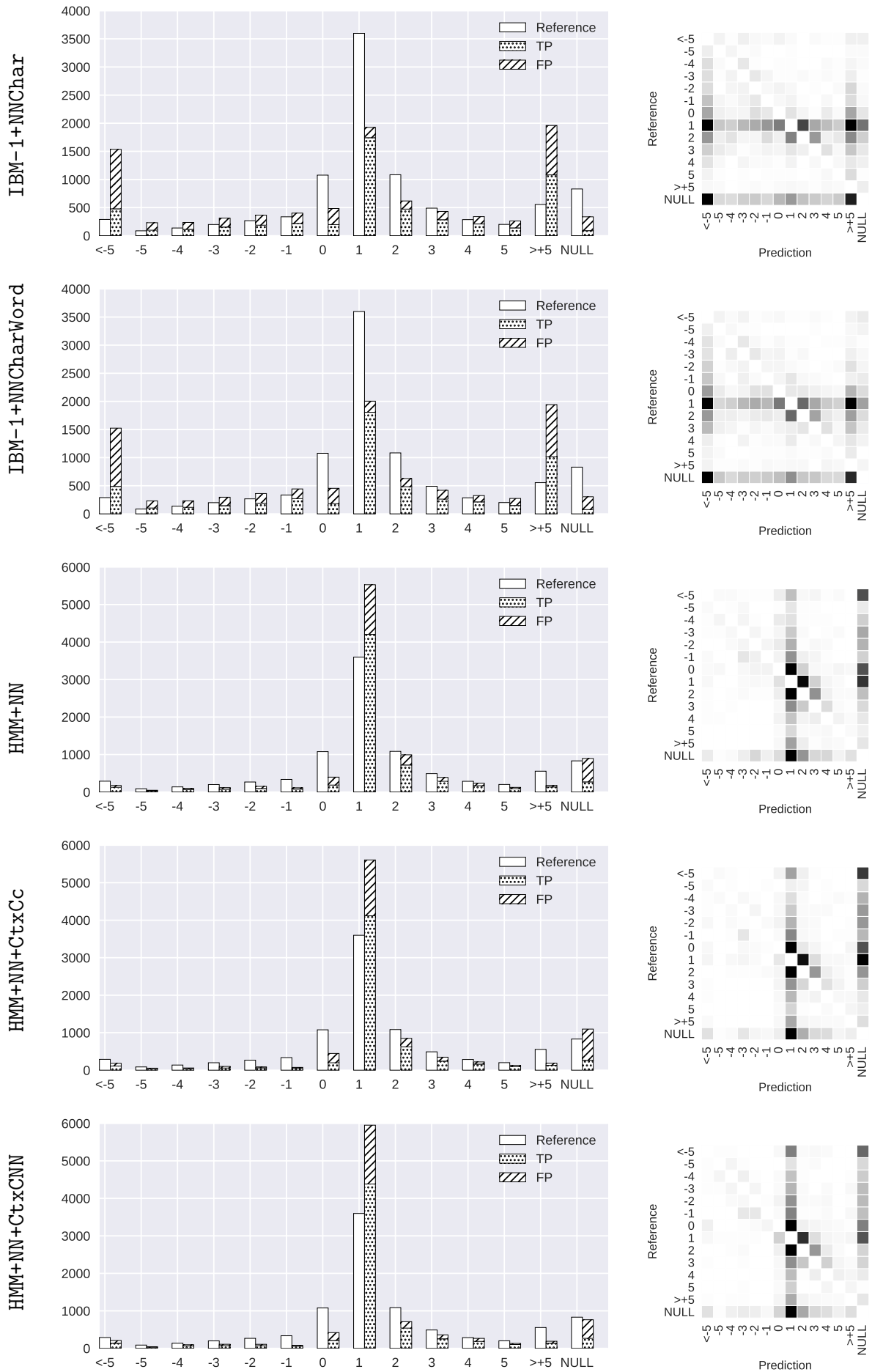


Figure B.27: Neural models for the direction English-German: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.





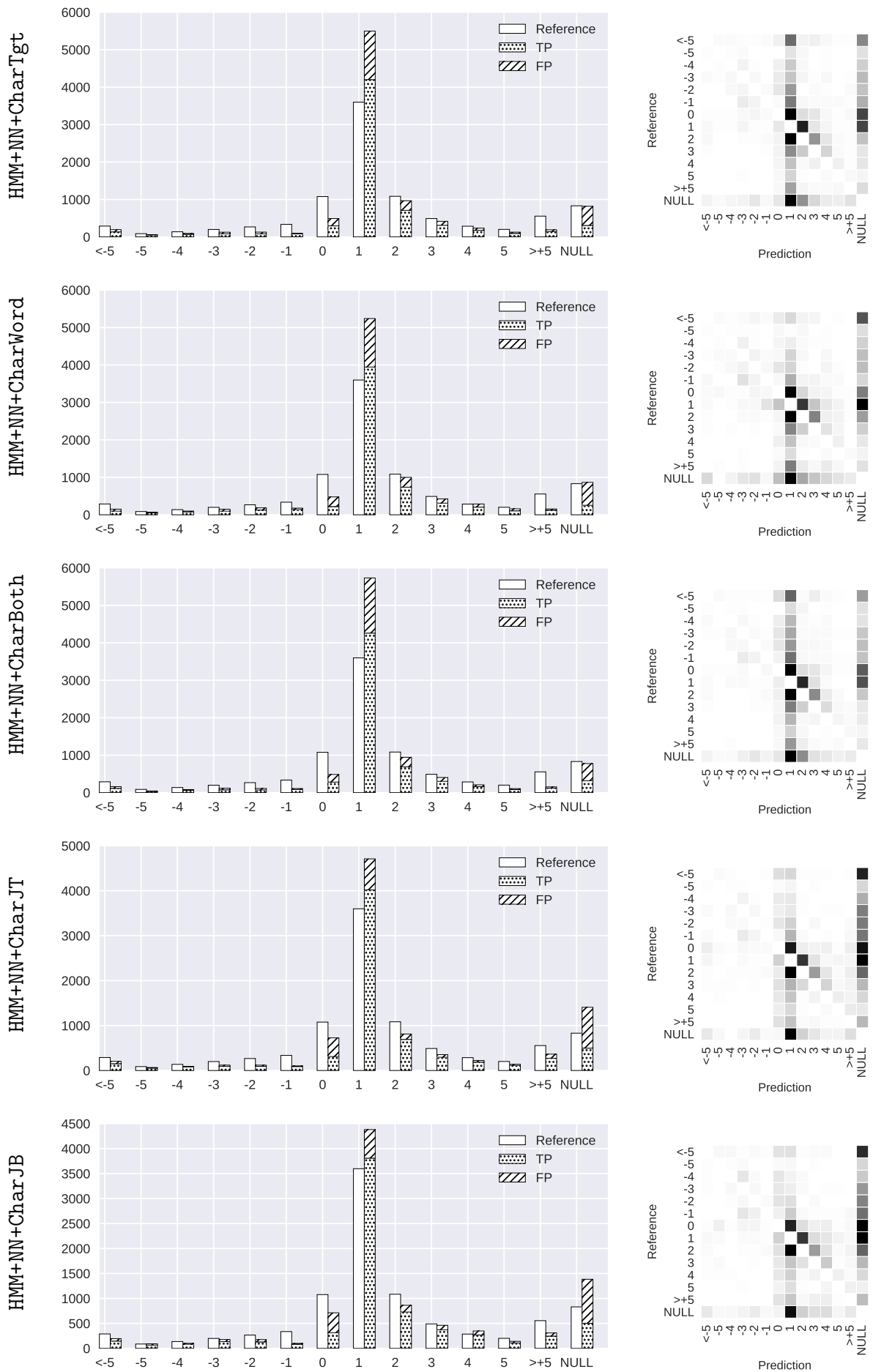
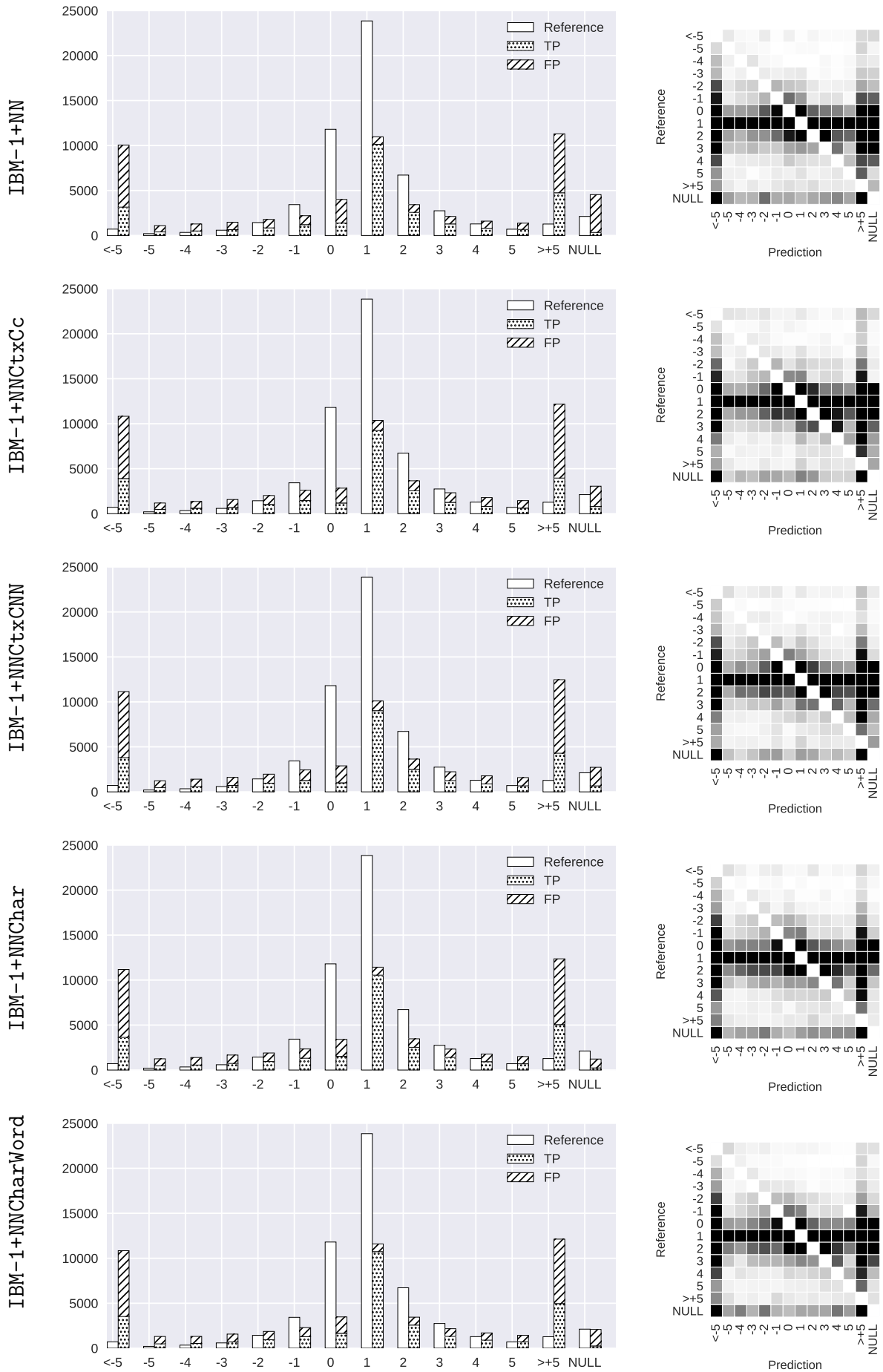
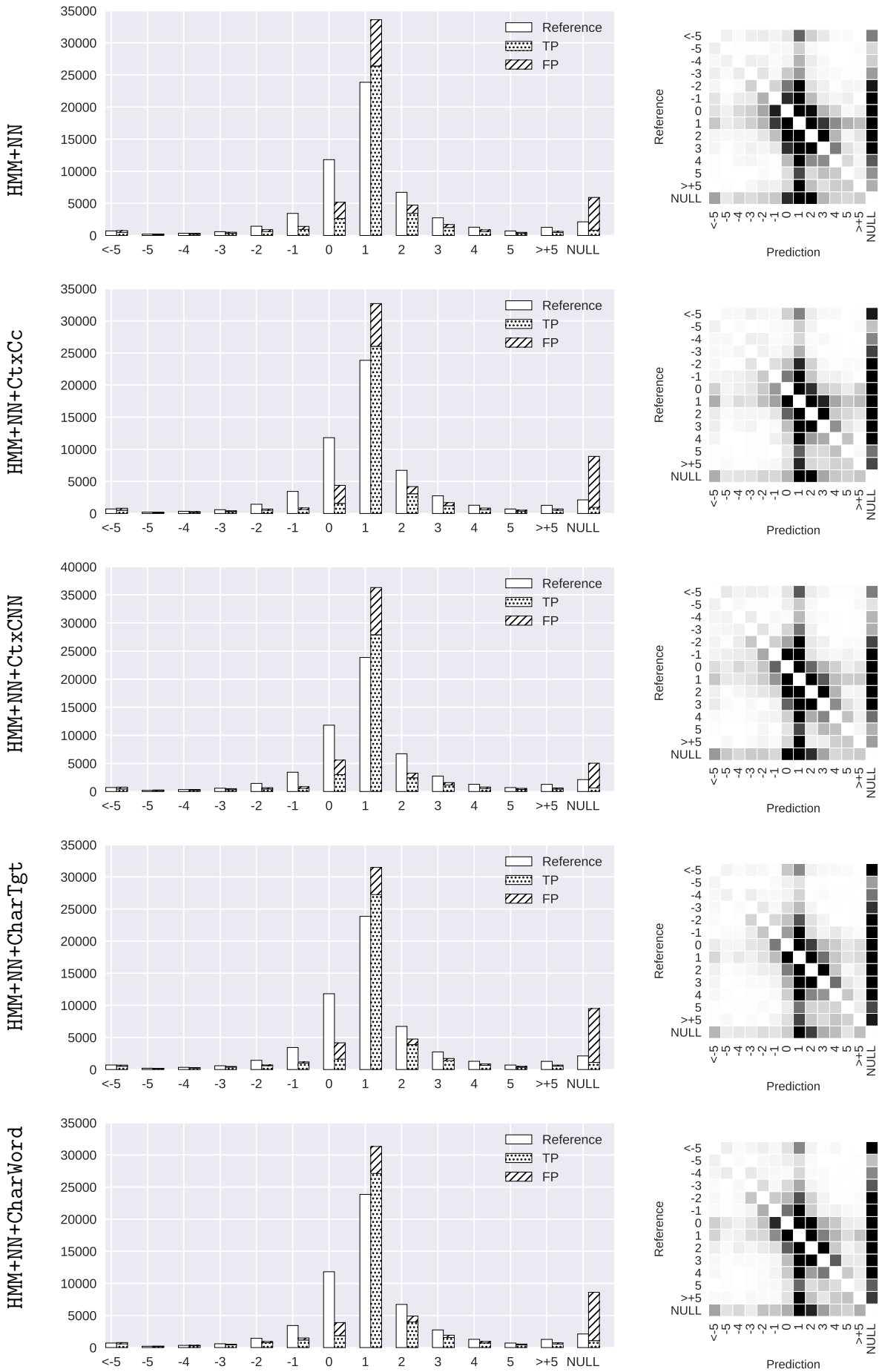


Figure B.28: Neural models for the direction German-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The

darker the cell, the greater the number of confusions.





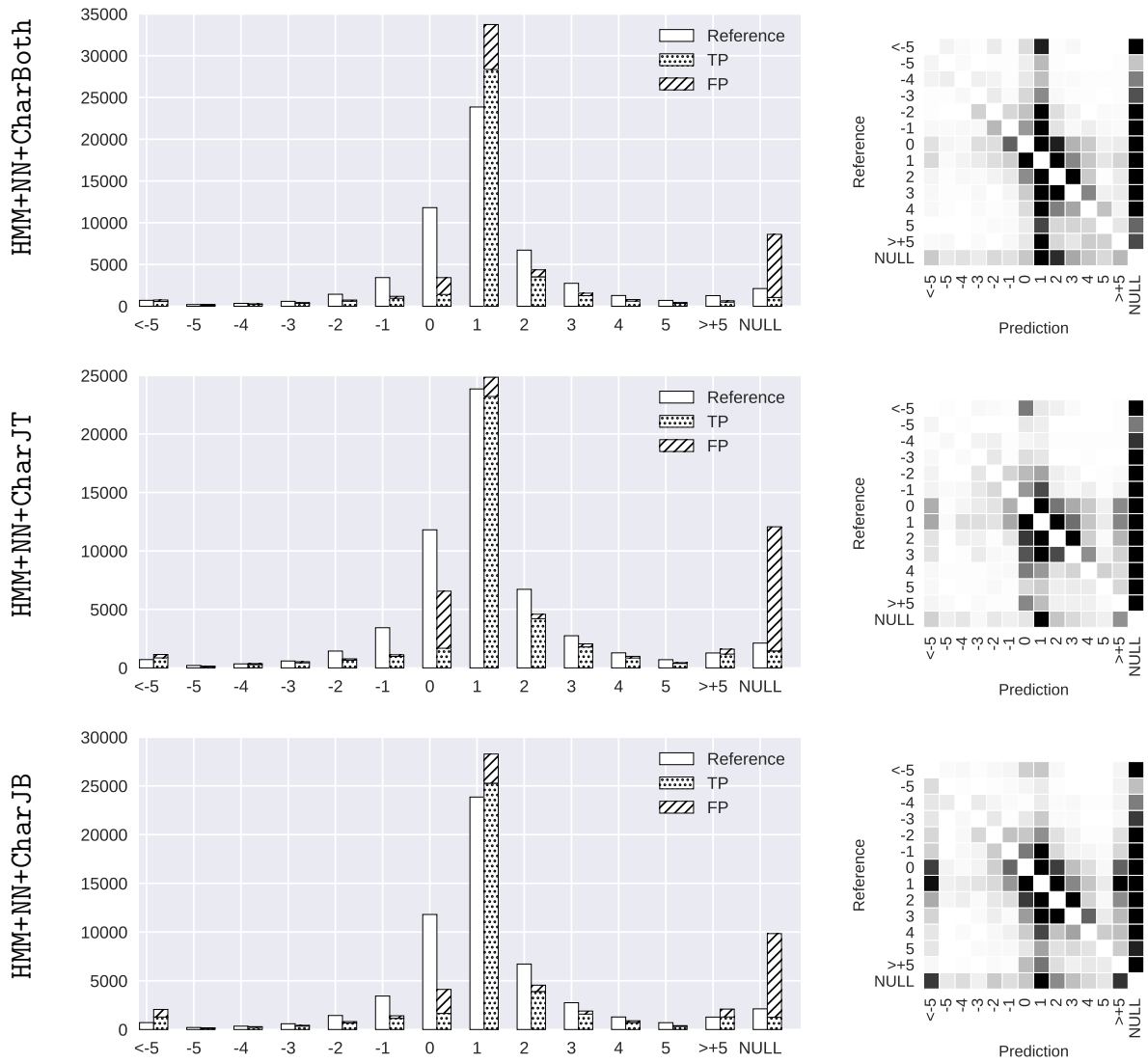
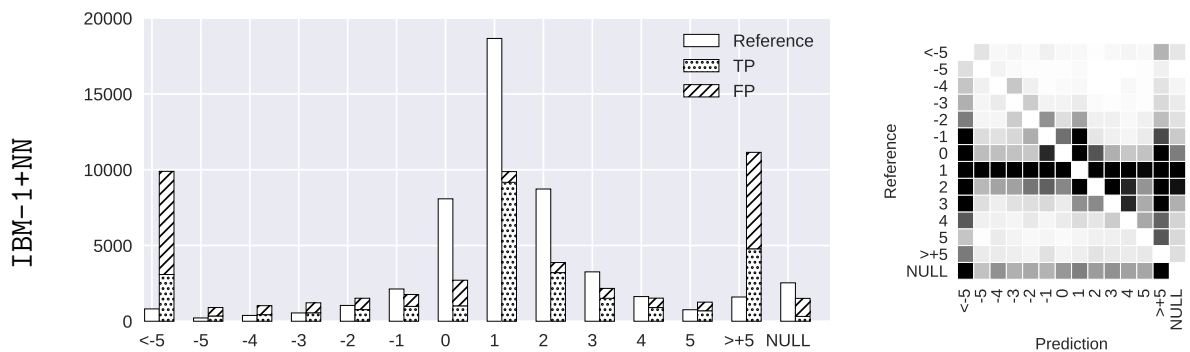
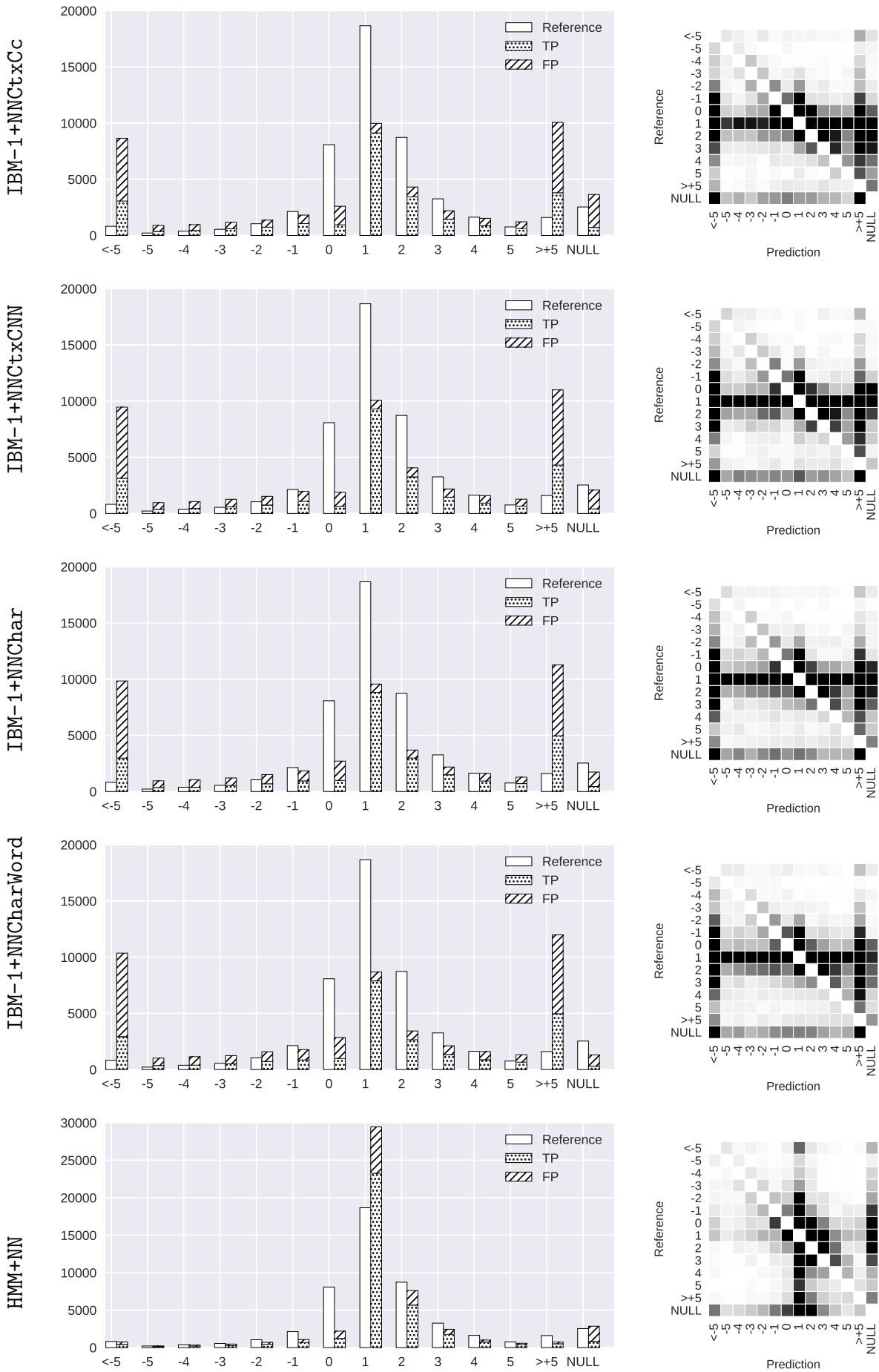
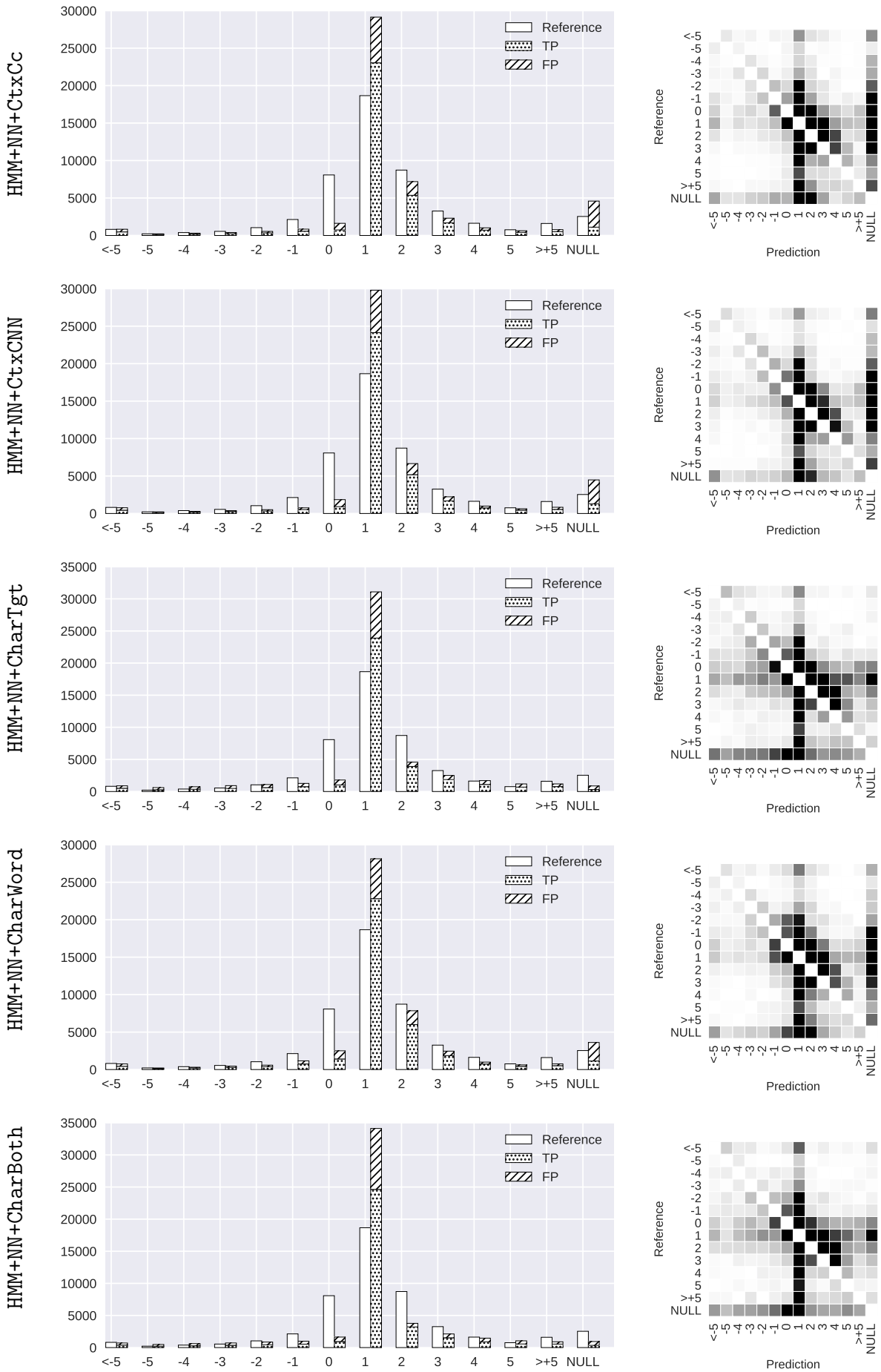


Figure B.29: Neural models for the direction English-Czech: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.







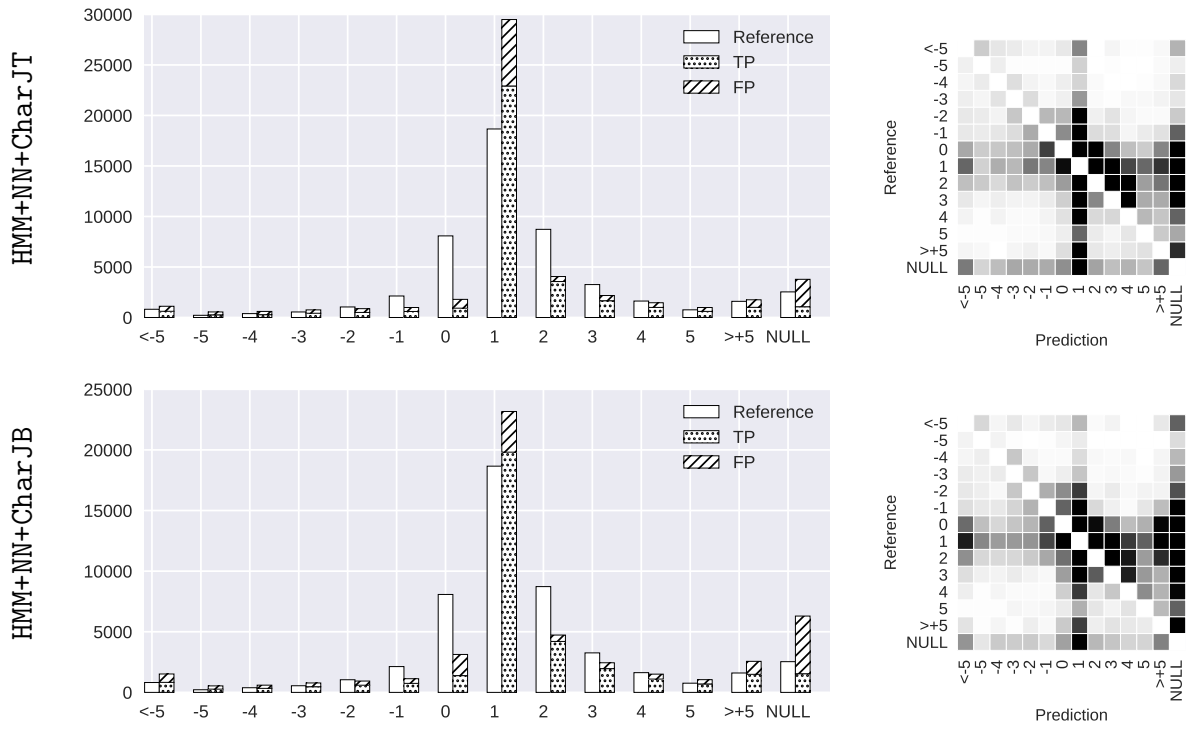
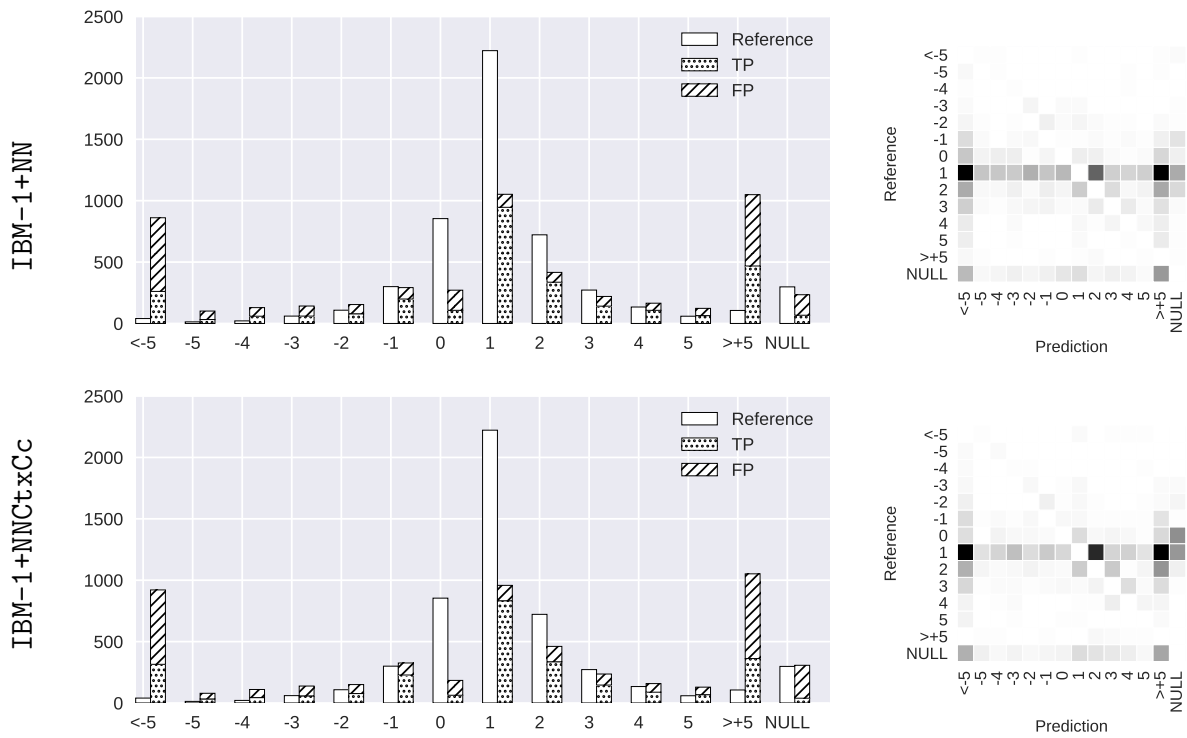
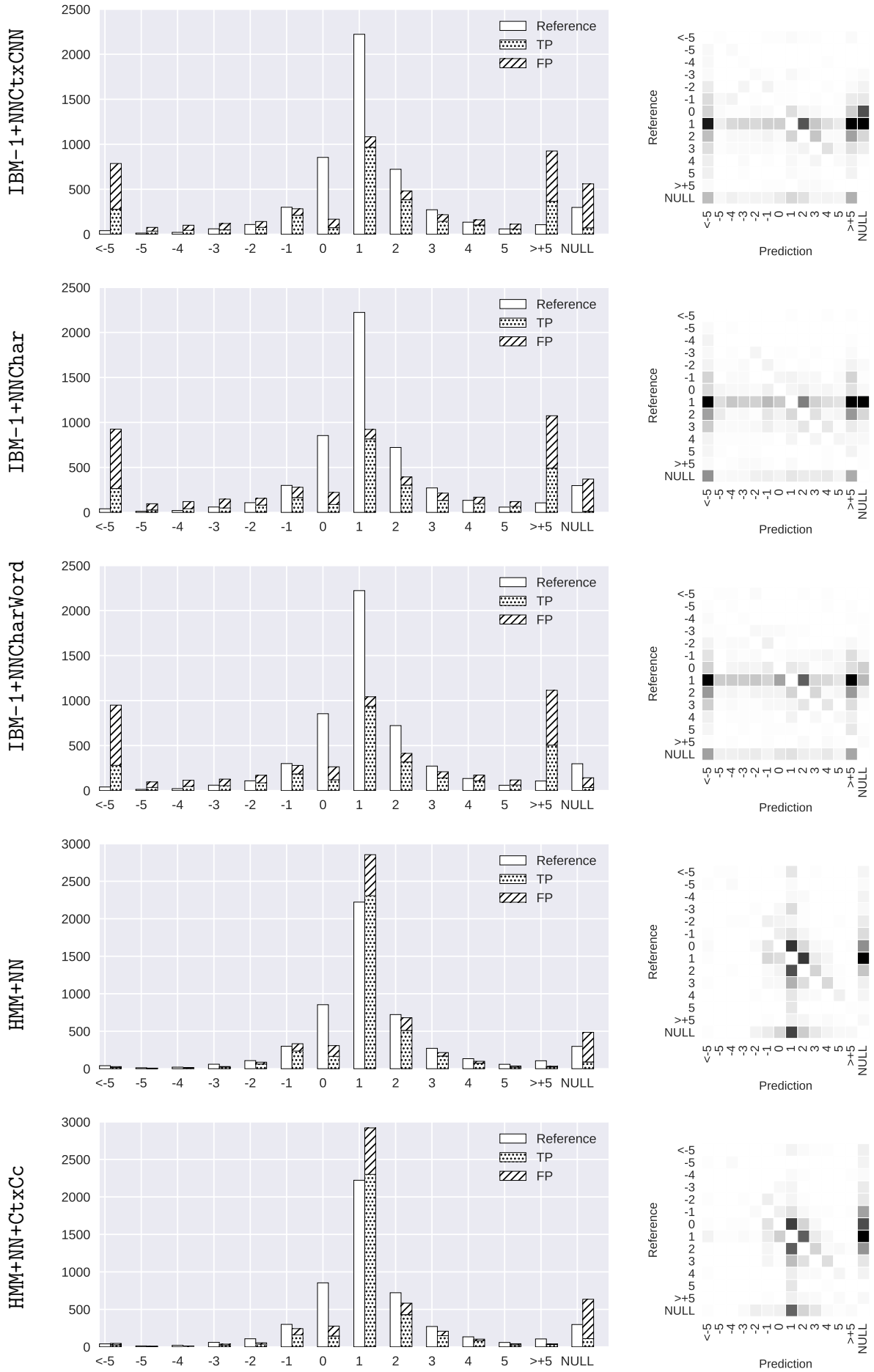
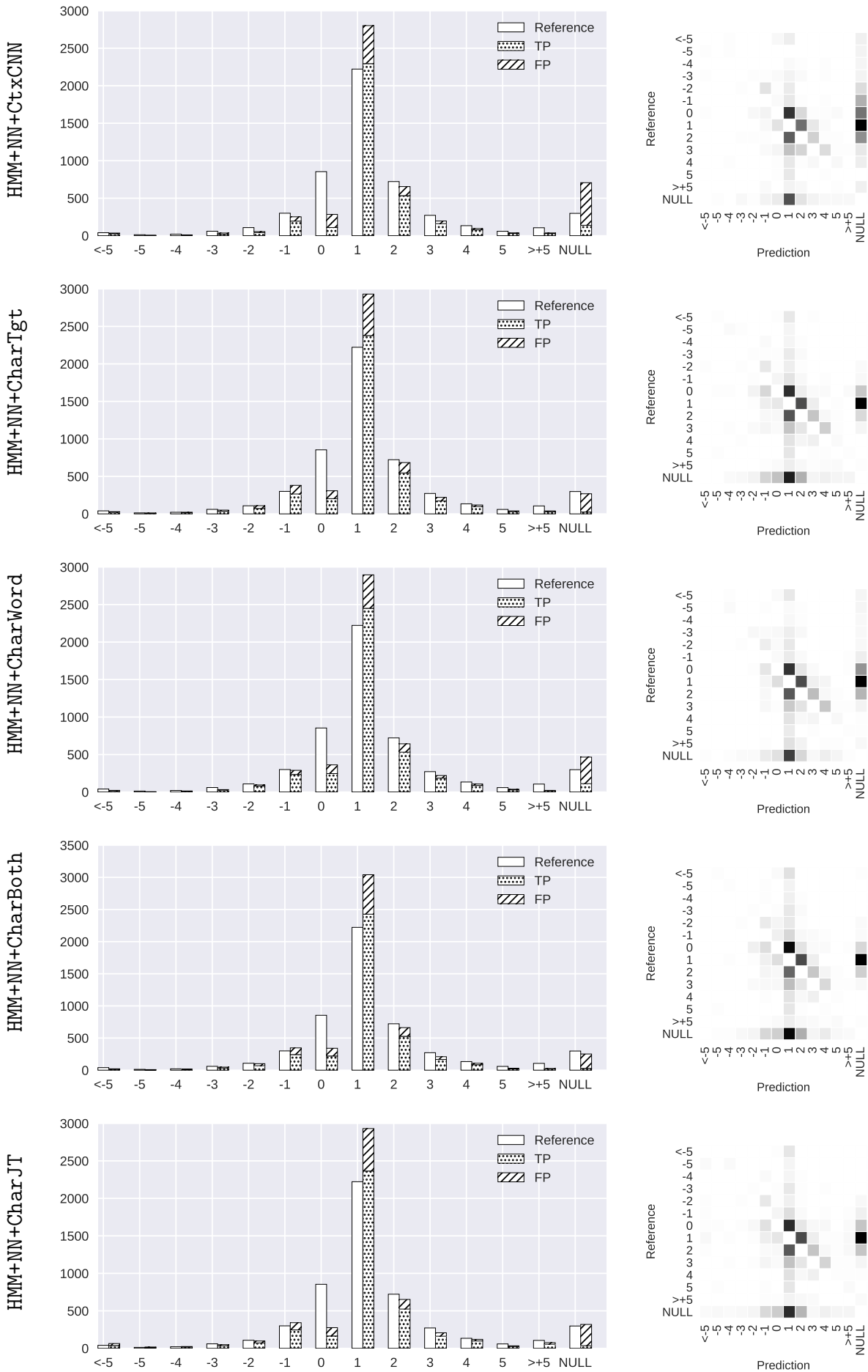


Figure B.30: Neural models for the direction Czech-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.







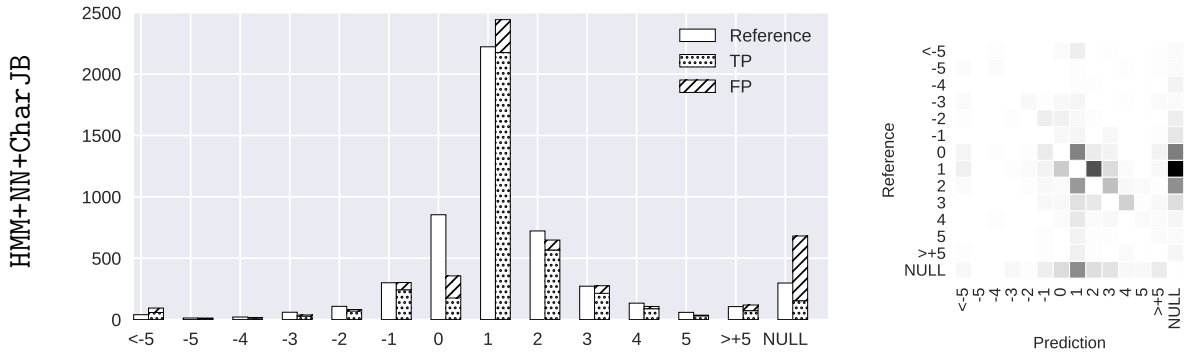
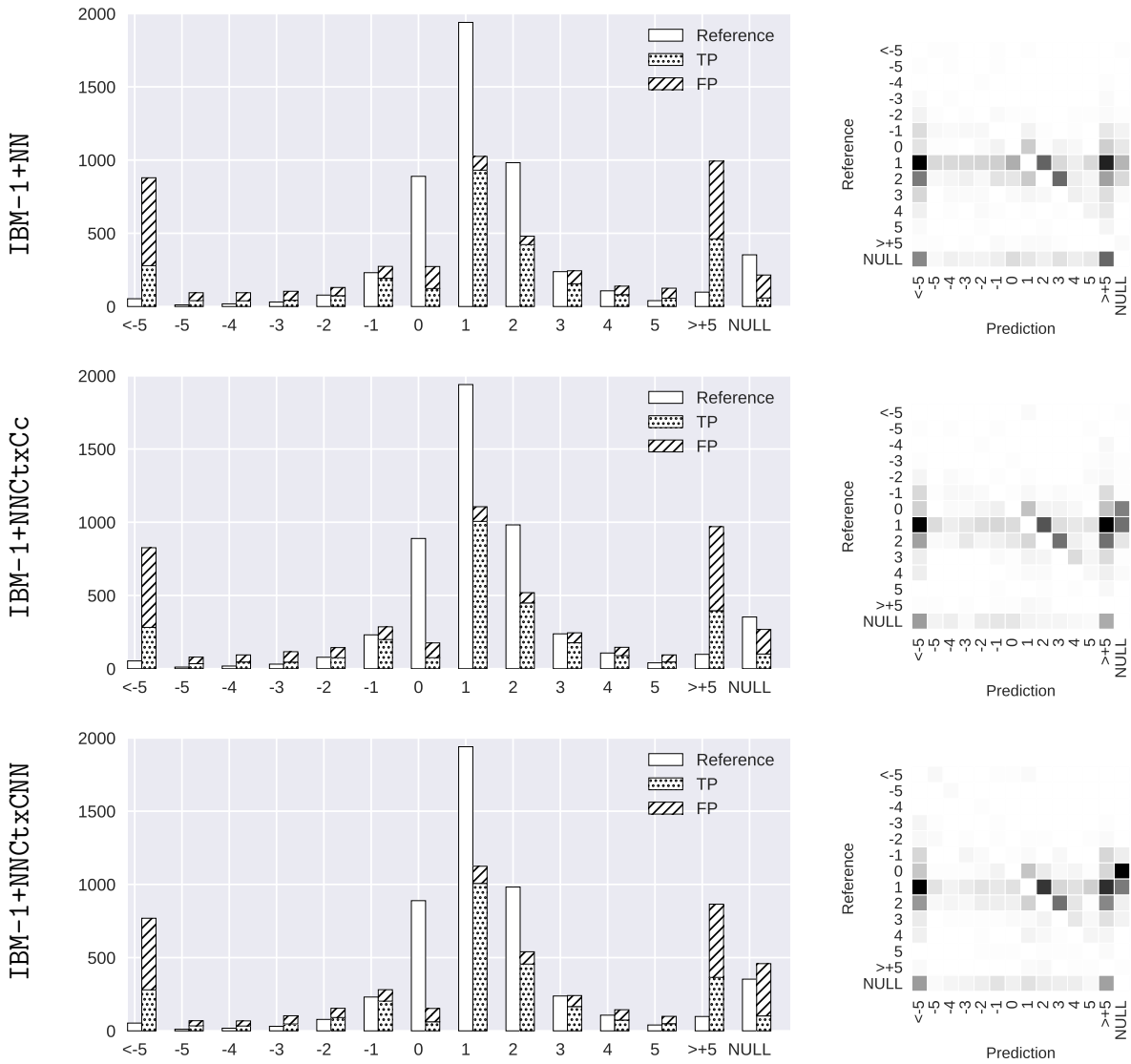
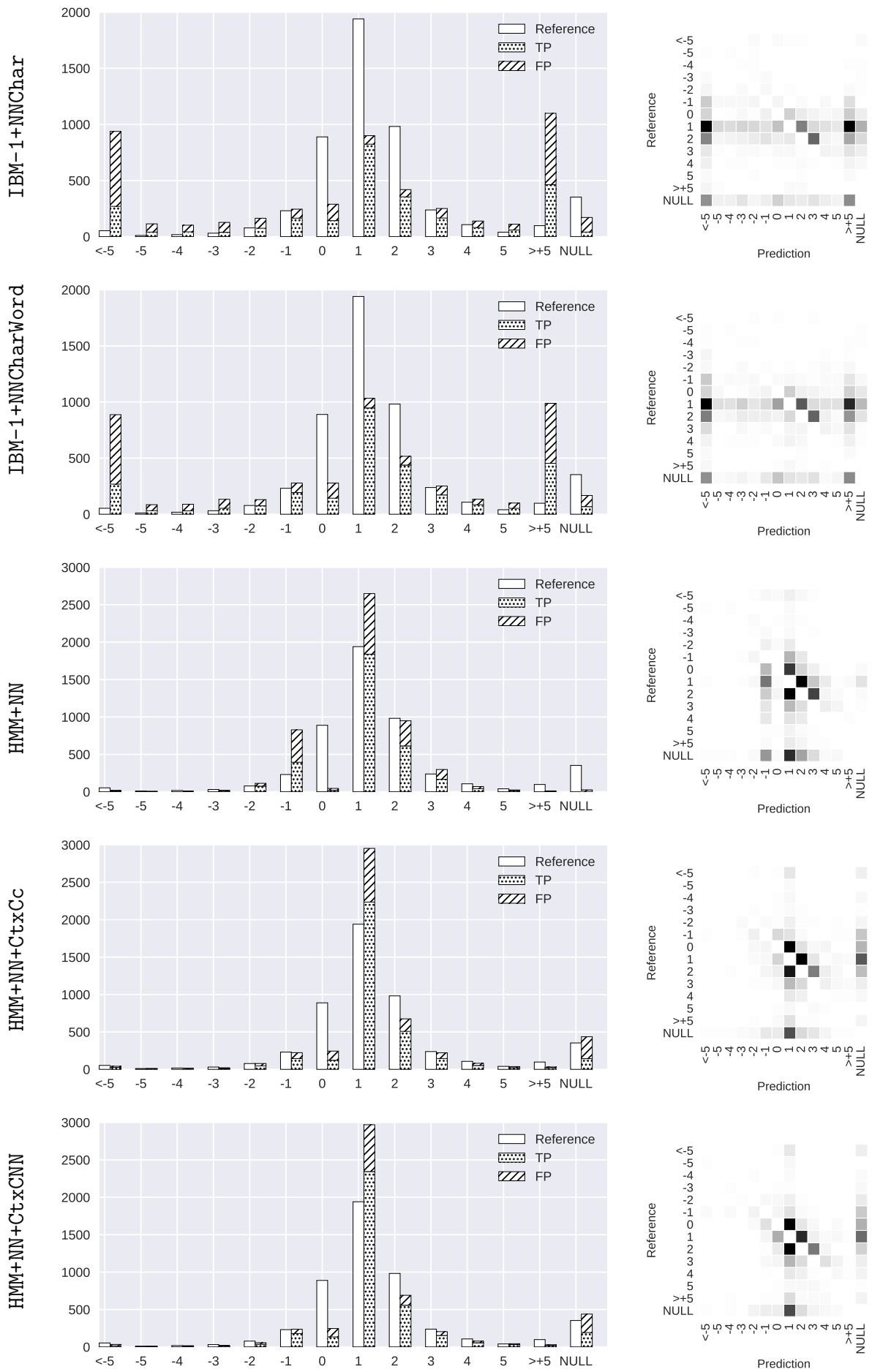


Figure B.31: Neural models for the direction English-Romanian: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.





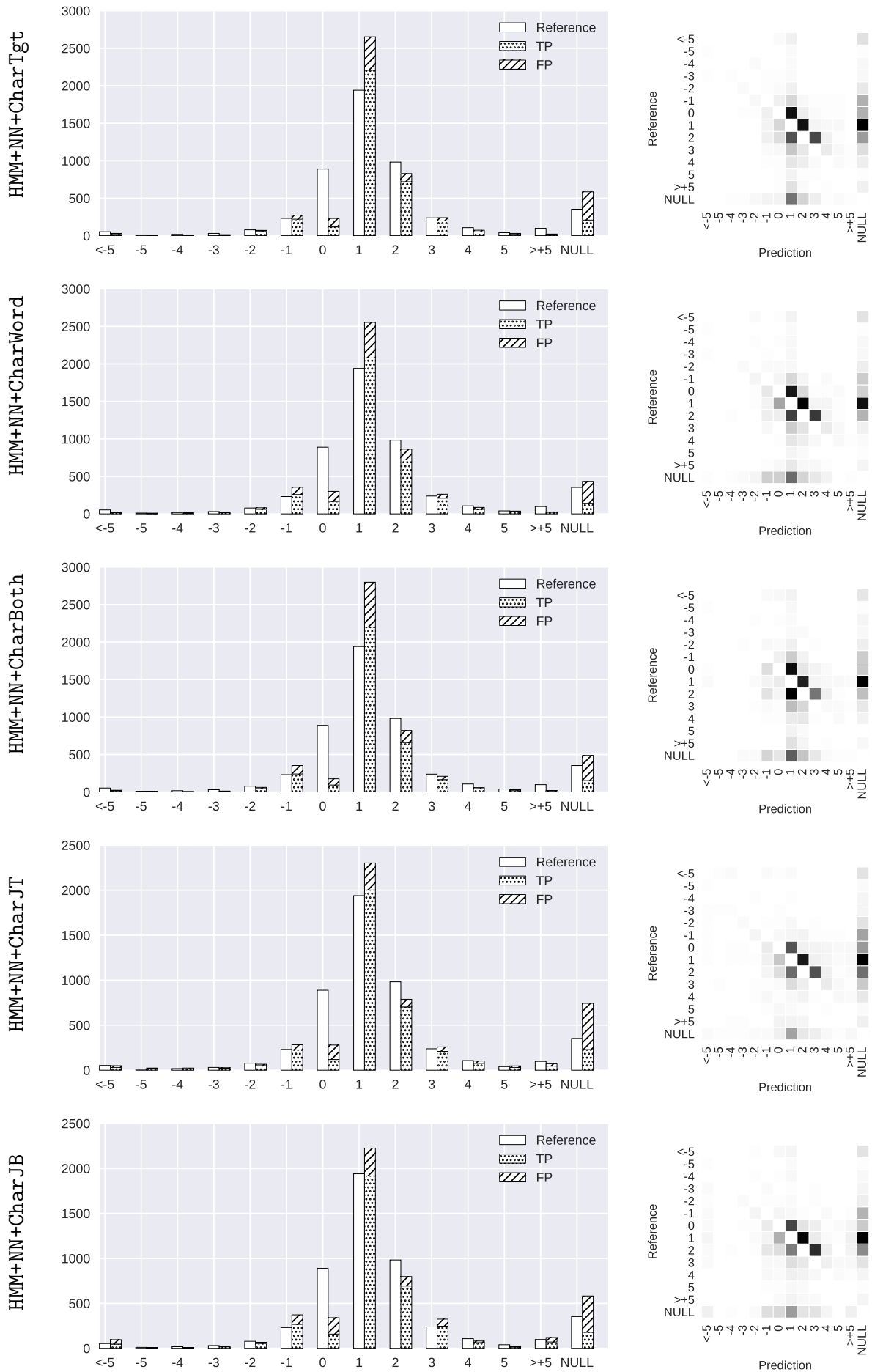
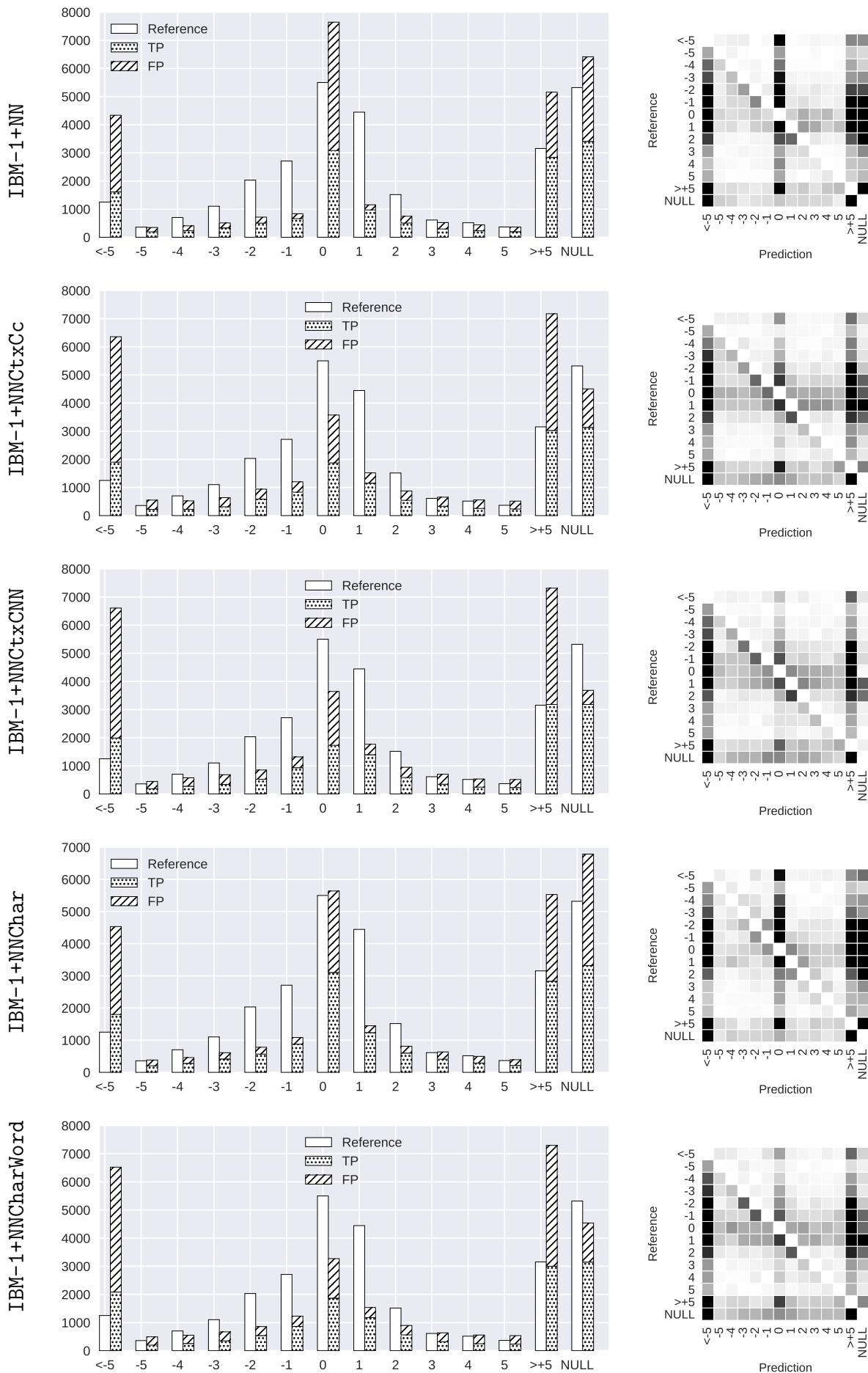
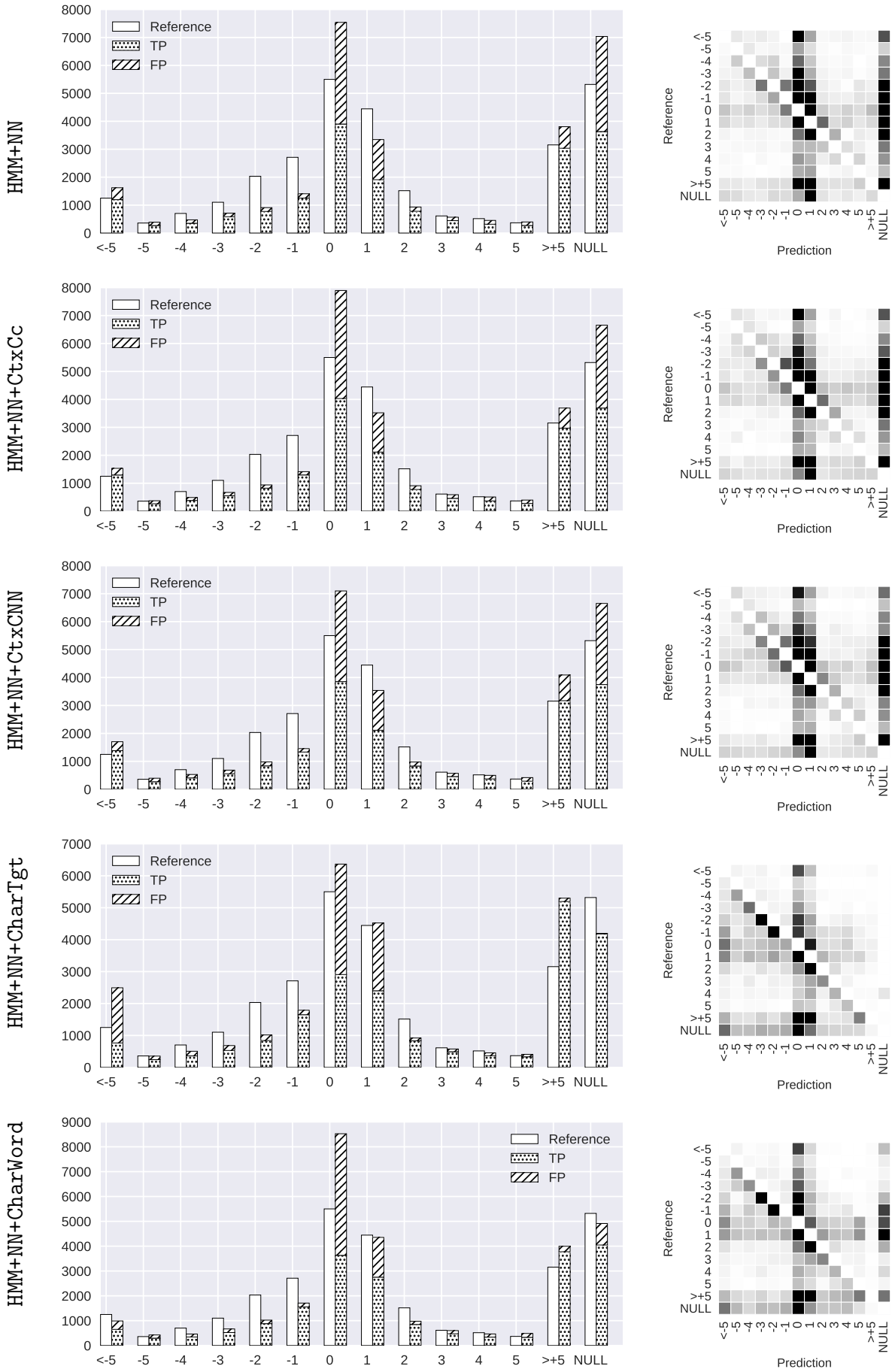


Figure B.32: Neural models for the direction Romanian-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph:

The darker the cell, the greater the number of confusions.





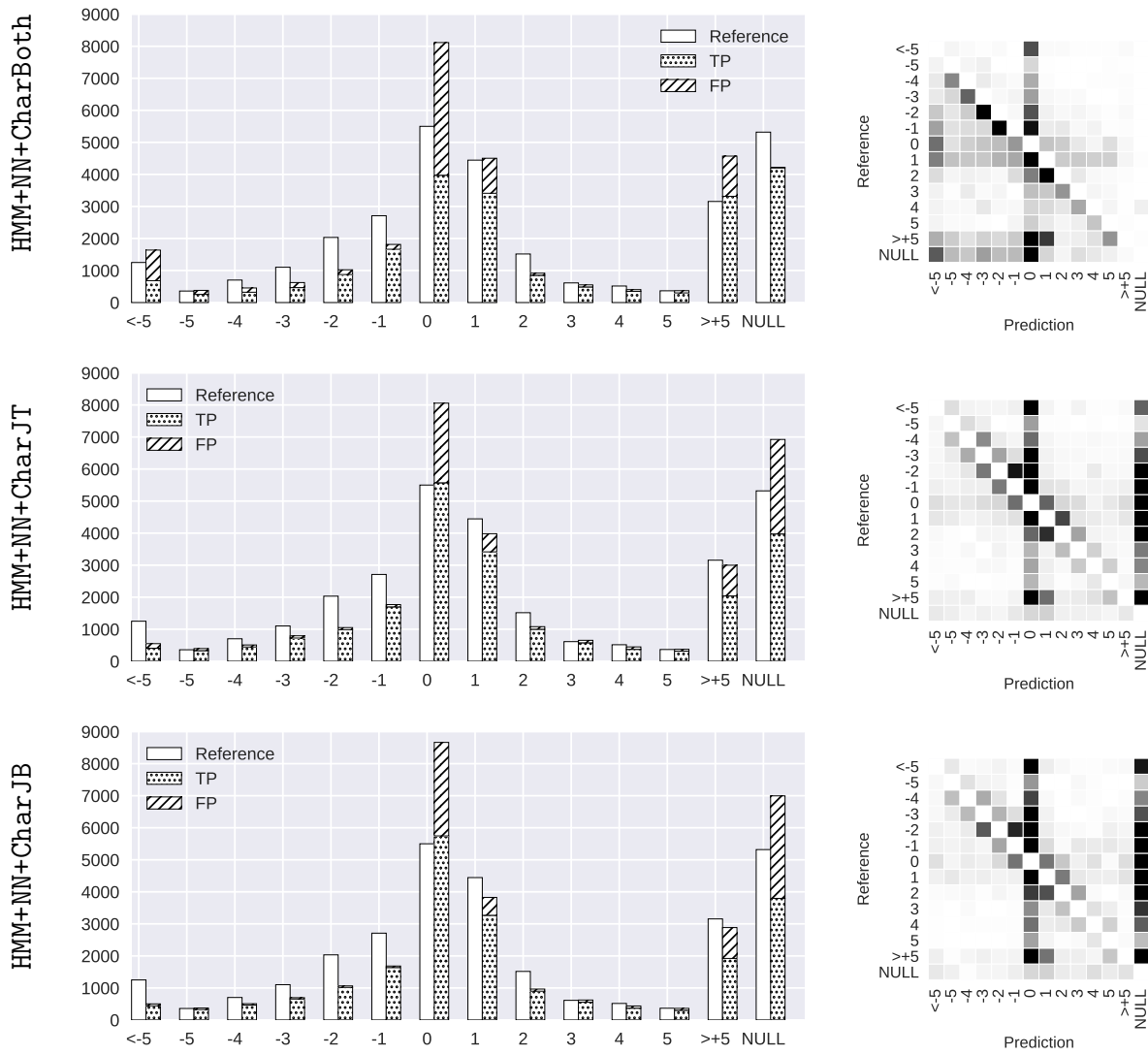
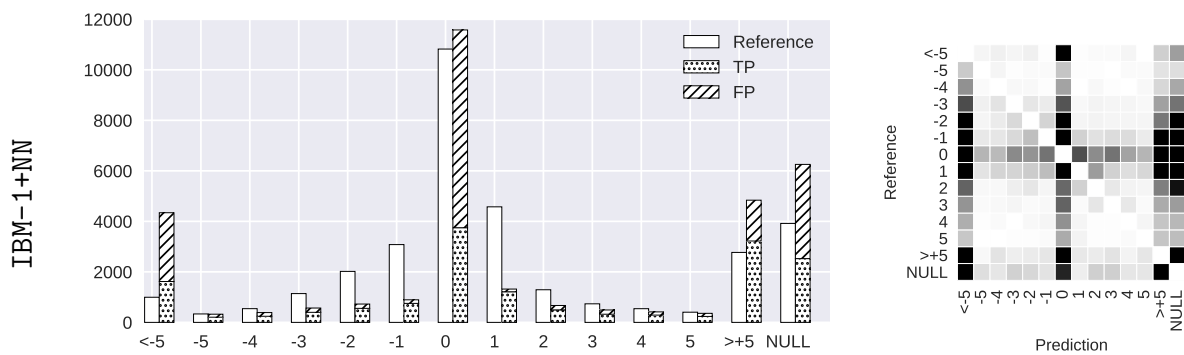
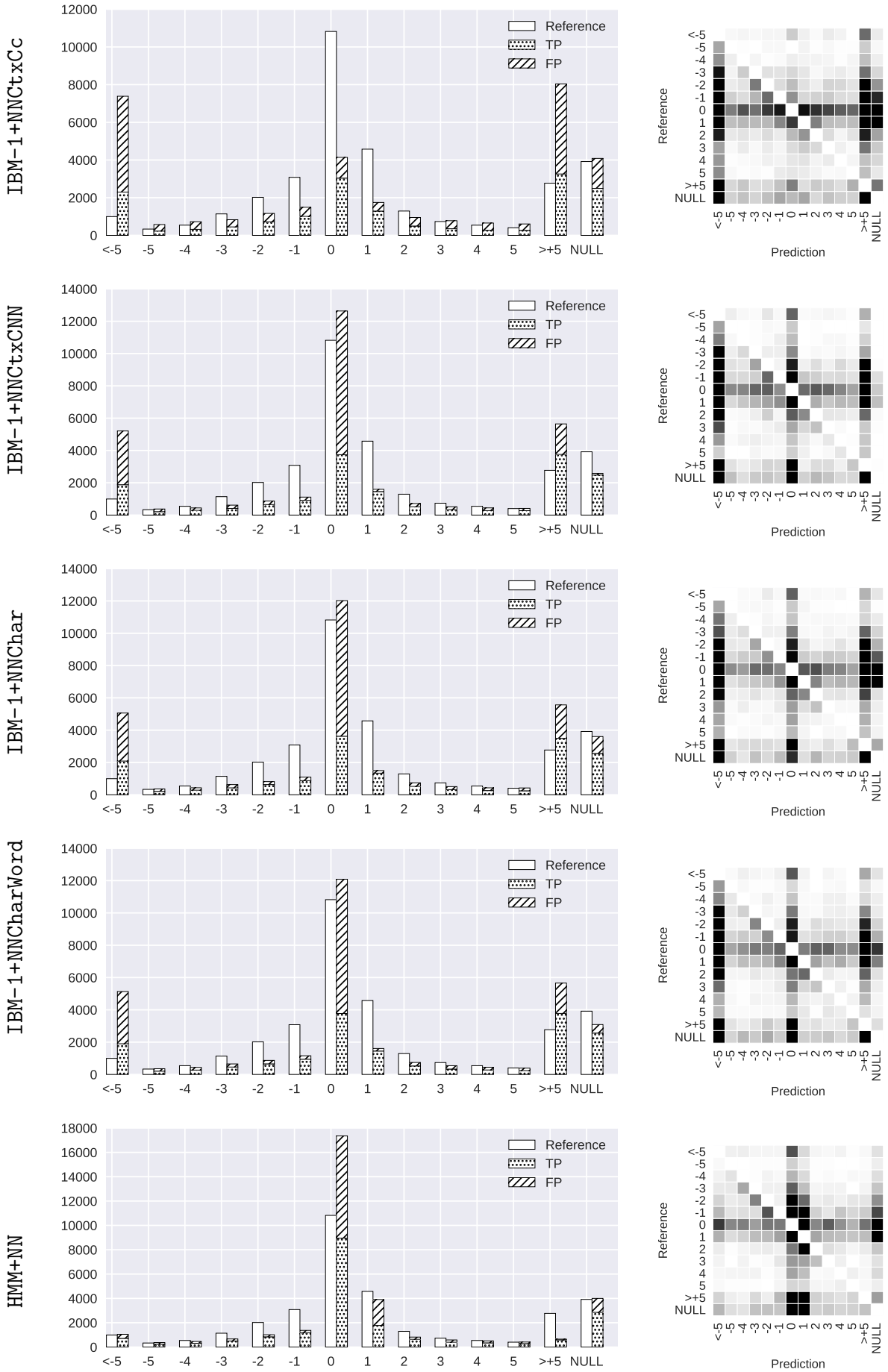
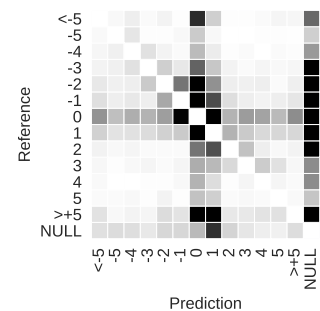
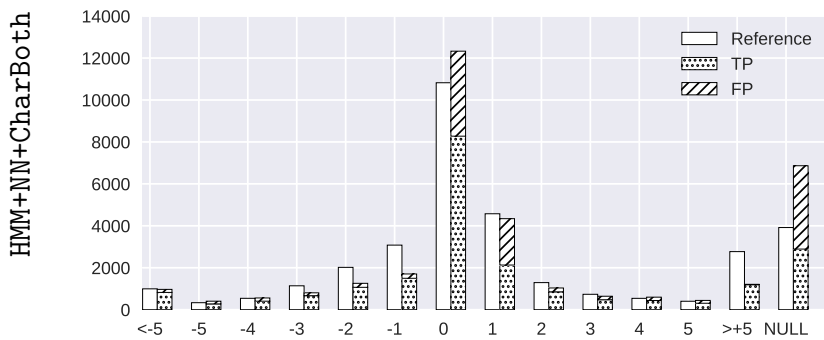
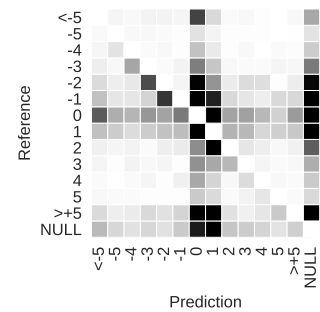
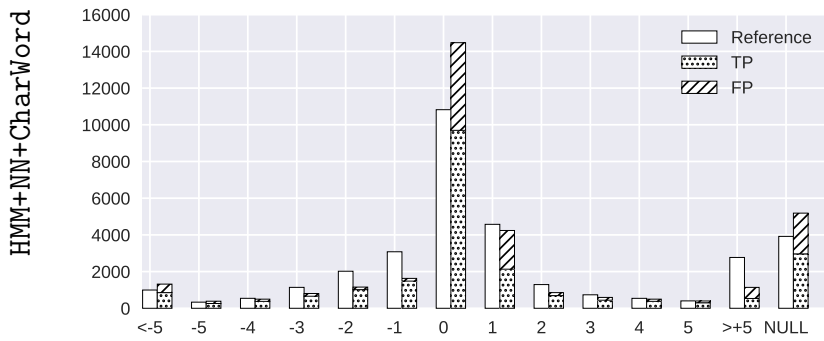
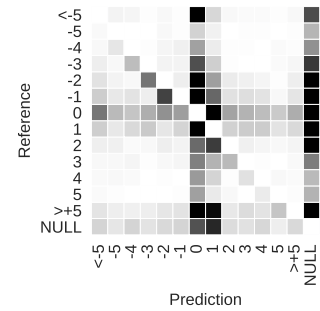
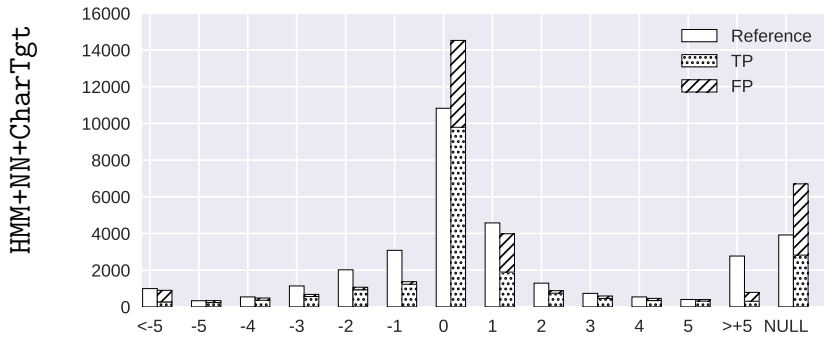
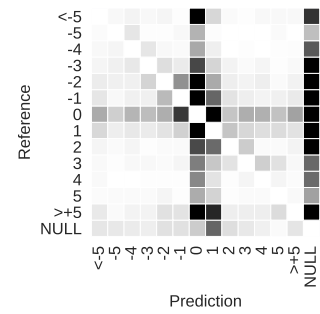
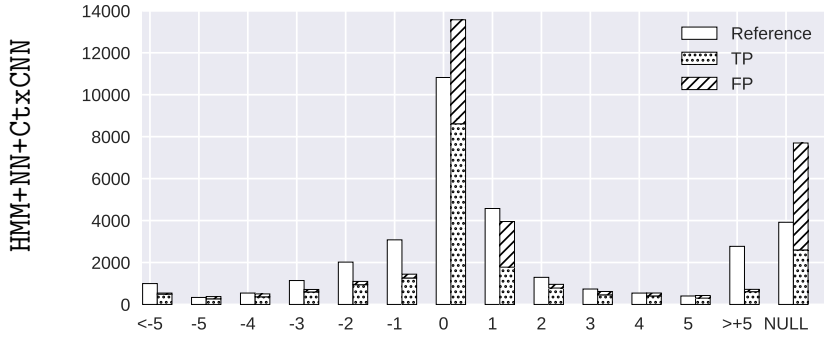
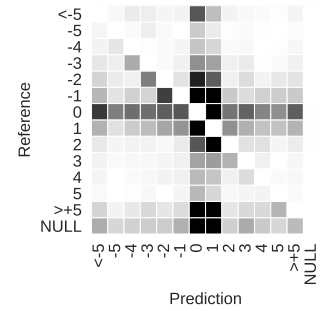
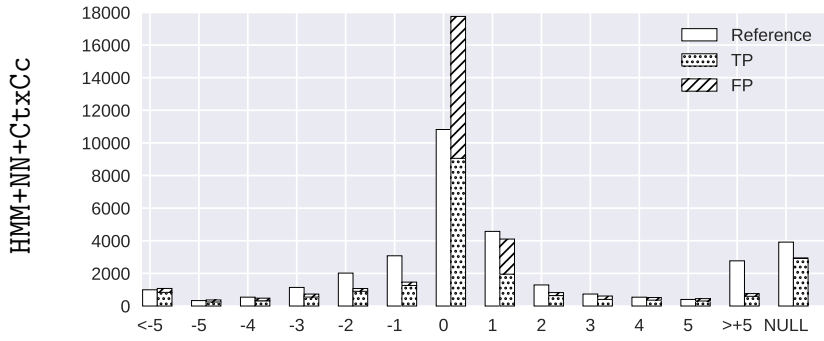


Figure B.33: Neural models for the direction English-Japanese: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.







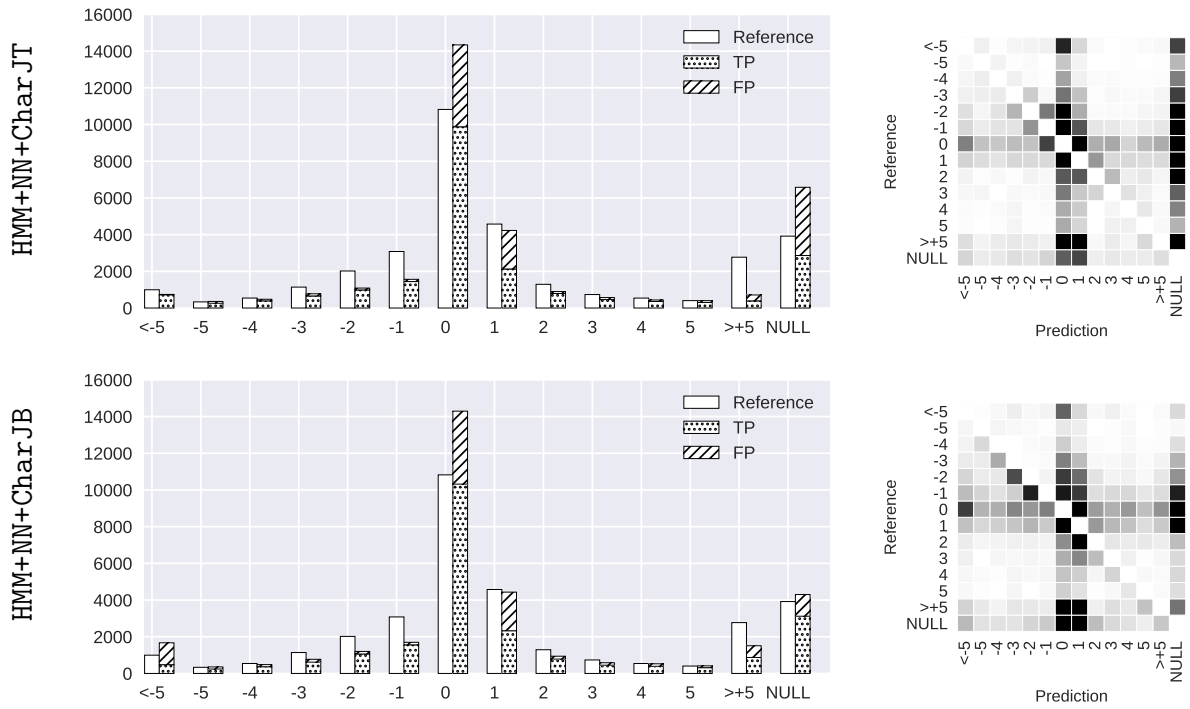
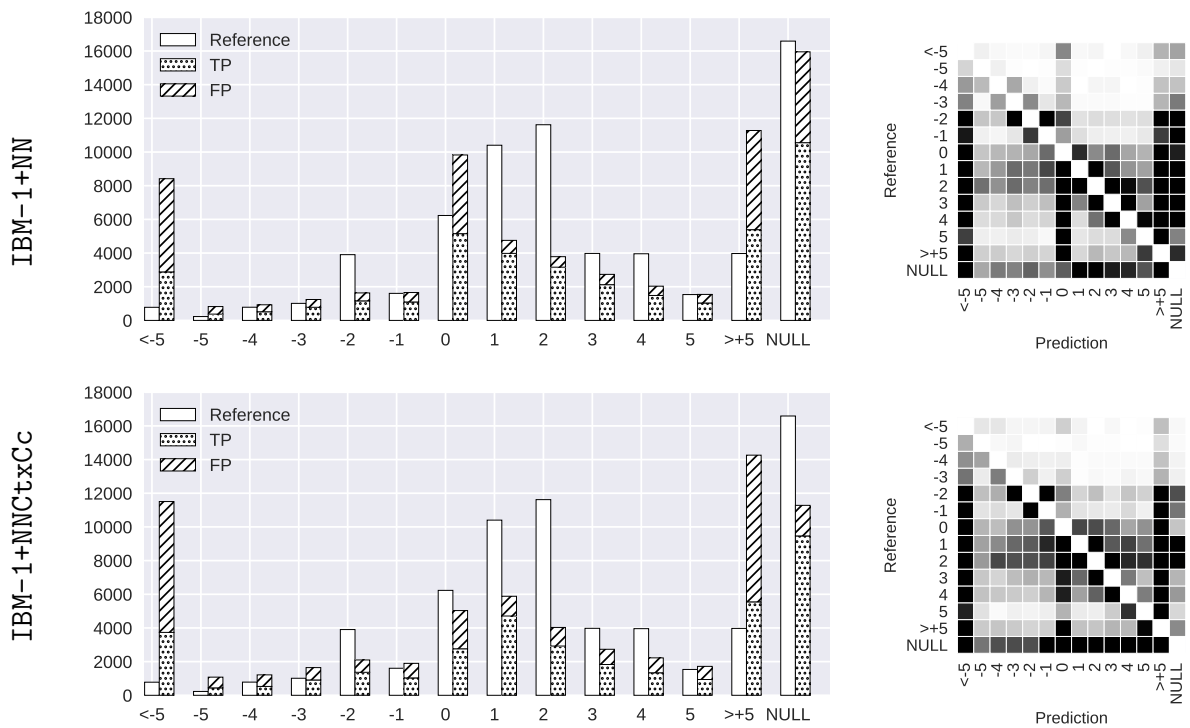
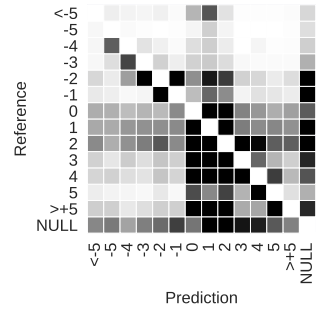
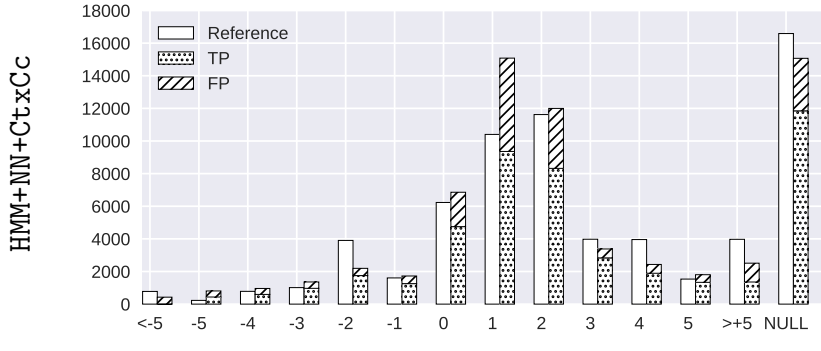
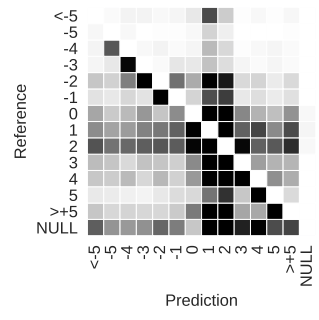
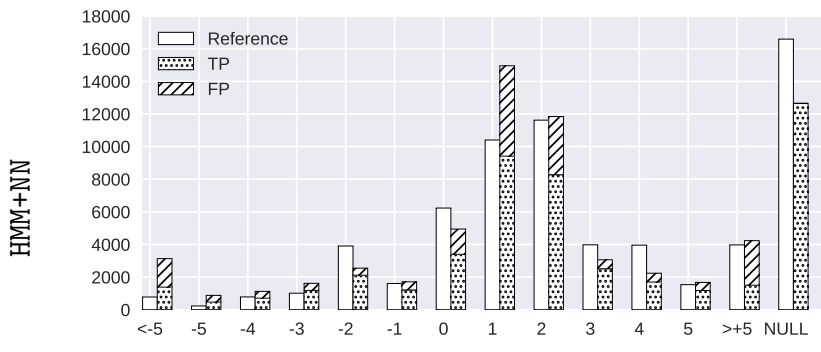
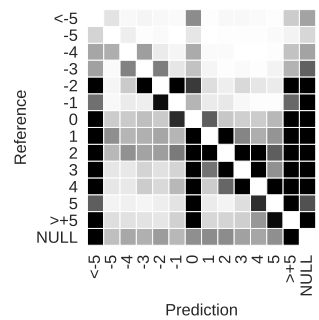
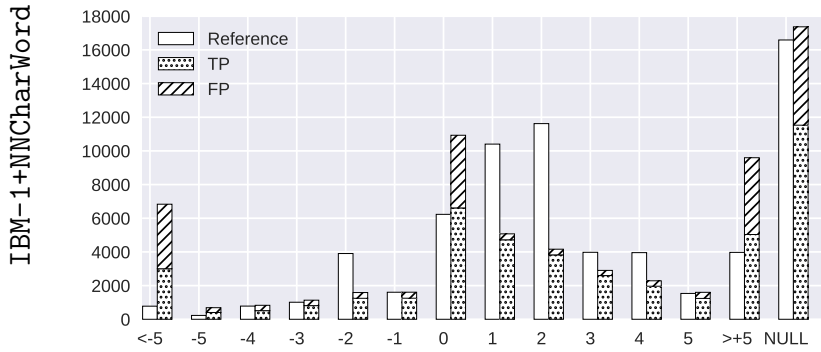
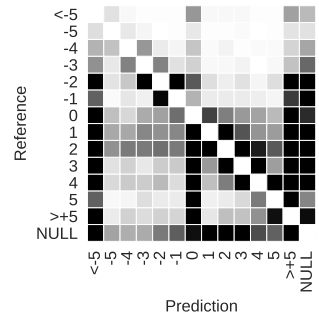
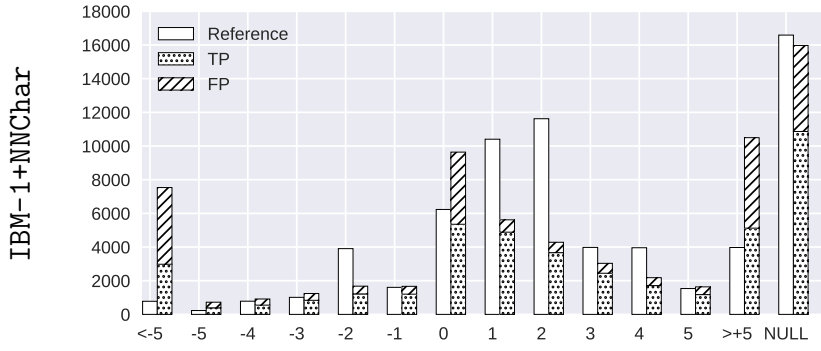
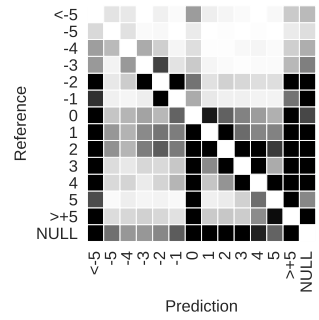
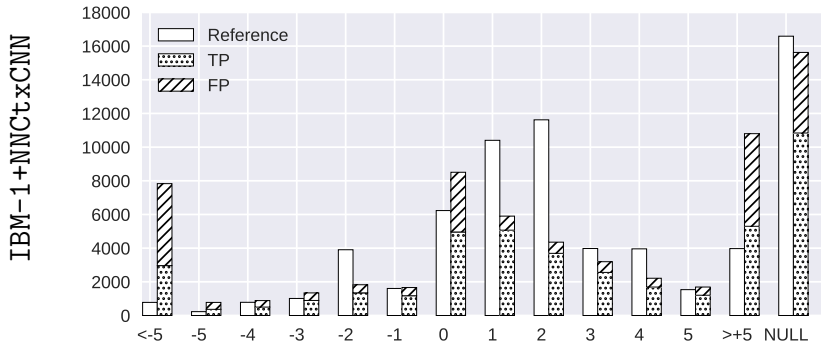
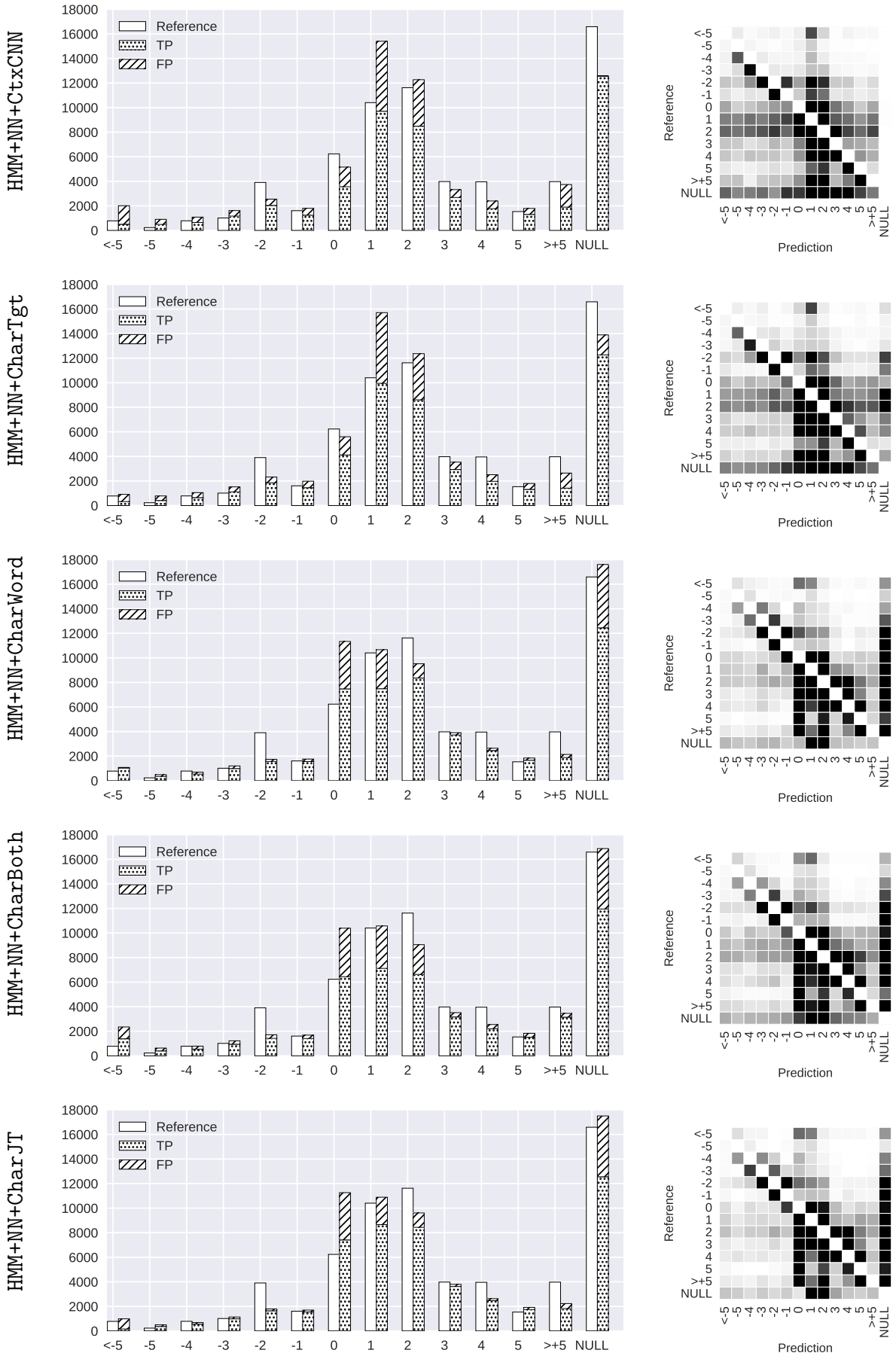


Figure B.34: Neural models for the direction Japanese-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.







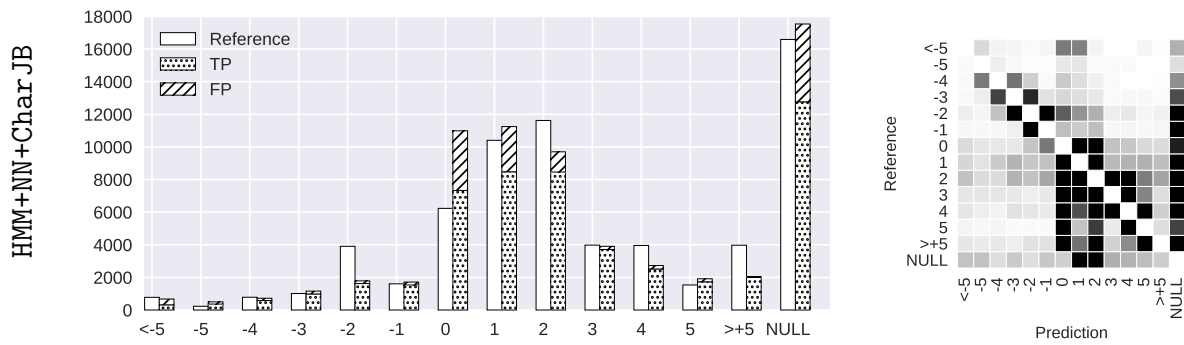
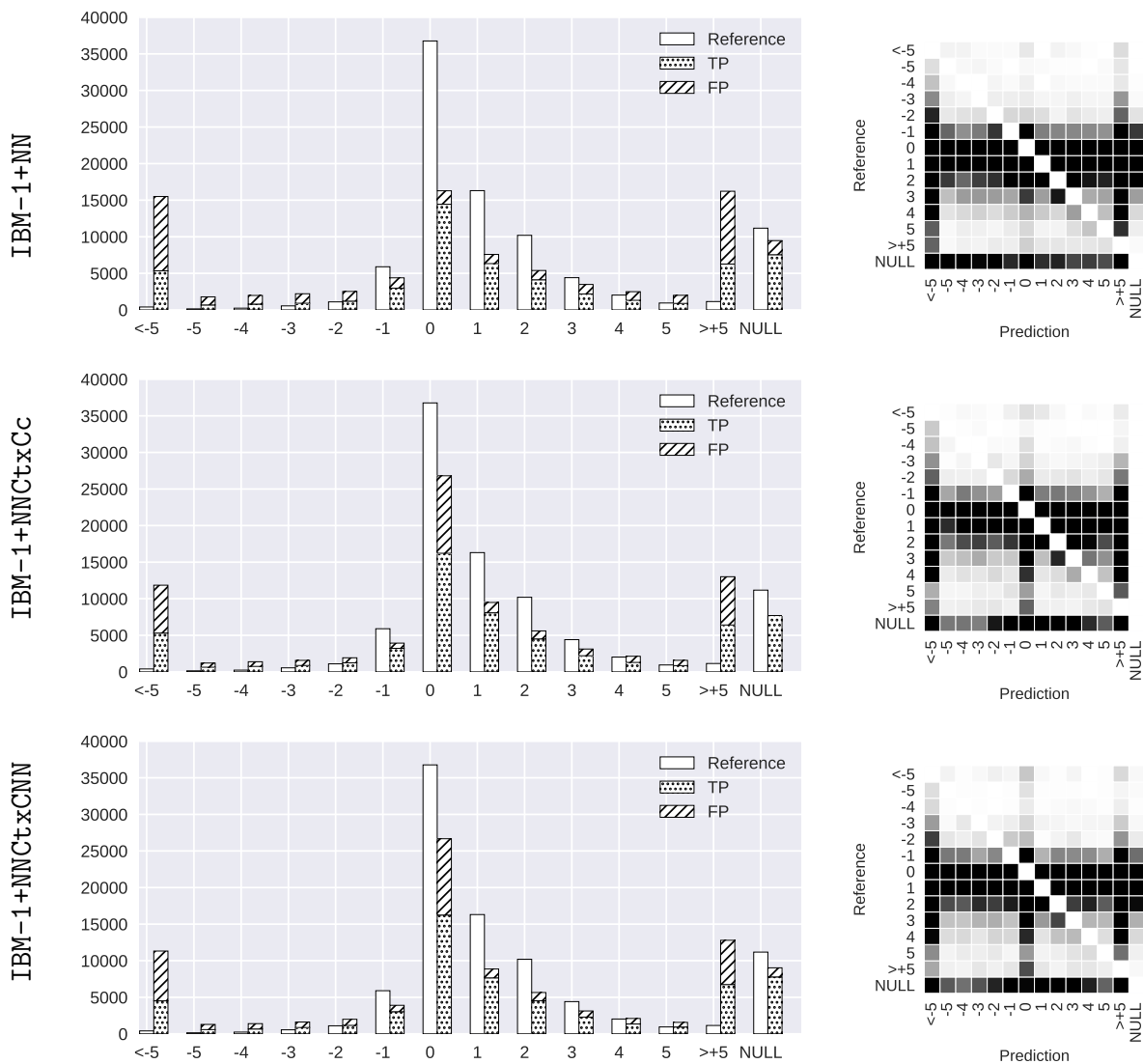
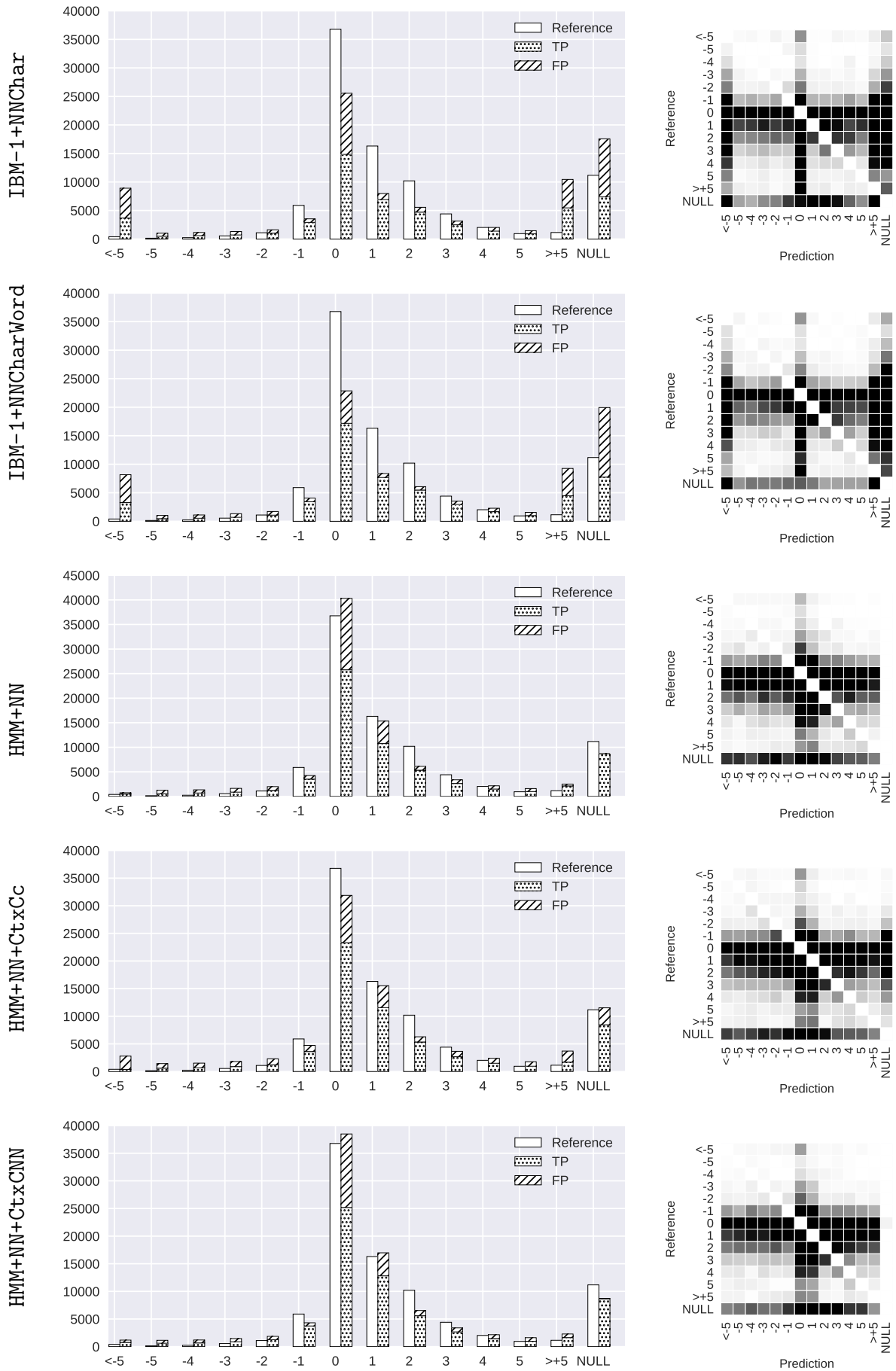


Figure B.35: Neural models for the direction English-Vietnamese: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.





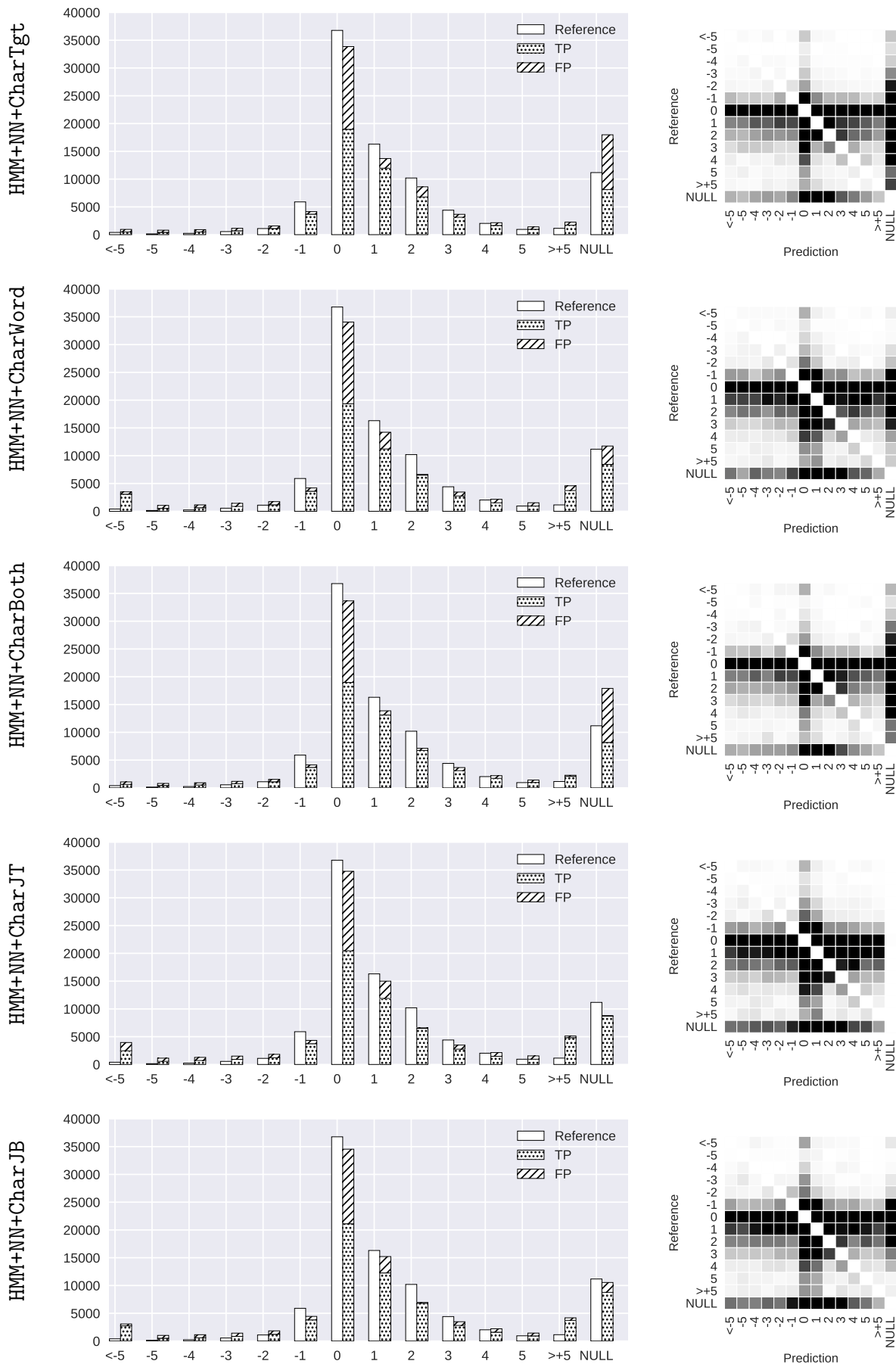


Figure B.36: Neural models for the direction Vietnamese-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. Confusion matrices on the right graph: The darker the cell, the greater the number of confusions.

B.6 Rare words

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	45	4.09	86.61	31.11	93.33	46.67	64	2.78	86.21	23.44	40.54	29.7
IBM-1+NN	7	0.64	93.31	42.86	20	27.27	38	1.65	89.71	28.95	29.73	29.33
IBM-1+NNCtxCc	5	0.45	94.98	80	26.67	40	39	1.7	88.74	23.08	24.32	23.68
IBM-1+NNCtxCNN	10	0.91	92.89	40	26.67	32	43	1.87	89.13	27.91	32.43	30
IBM-1+NNChar	7	0.64	94.98	71.43	33.33	45.45	45	1.96	87.57	20	24.32	21.95
IBM-1+NNCharWord	18	1.64	92.89	44.44	53.33	48.48	45	1.96	87.96	22.22	27.03	24.39
Fastalign	11	1	95.82	72.73	53.33	61.54	20	0.87	95.15	80	43.24	56.14
HMM Giza++	33	3	91.63	42.42	93.33	58.33	72	3.13	87.38	30.56	59.46	40.37
HMM+NN	11	1	96.65	81.82	60	69.23	34	1.48	92.82	50	45.95	47.89
HMM+NNCtxCc	10	0.91	96.23	80	53.33	64	29	1.26	93.01	51.72	40.54	45.45
HMM+NNCtxCNN	9	0.82	97.49	100	60	75	23	1	93.01	52.17	32.43	40
HMM+NNCharTgt	9	0.82	96.65	88.89	53.33	66.67	26	1.13	95.53	76.92	54.05	63.49
HMM+NNCharWord	10	0.91	96.23	80	53.33	64	31	1.35	94.56	64.52	54.05	58.82
HMM+NNCharBoth	10	0.91	96.23	80	53.33	64	23	1	93.79	60.87	37.84	46.67
HMM+NNCharJT	6	0.55	96.23	100	40	57.14	20	0.87	95.15	80	43.24	56.14
HMM+NNCharJB	8	0.73	97.07	100	53.33	69.57	18	0.78	94.76	77.78	37.84	50.91
IBM-4 Giza++	21	1.91	94.14	52.38	73.33	61.11	55	2.39	89.9	36.36	54.05	43.48

Table B.19: Models for English-French: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction French-English and in the direction English-French

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	26	4.33	78.02	23.08	100	37.5	88	4.63	85.37	35.23	77.5	48.44
IBM-1+NN	4	0.67	95.6	75	50	60	35	1.84	92.24	57.14	50	53.33
IBM-1+NNCtxCc	6	1	93.41	50	50	50	29	1.53	92.24	58.62	42.5	49.28
IBM-1+NNCtxCNN	6	1	93.41	50	50	50	42	2.21	91.13	50	52.5	51.22
IBM-1+NNChar	9	1.5	92.31	44.44	66.67	53.33	38	2	93.79	65.79	62.5	64.1
IBM-1+NNCharWord	6	1	95.6	66.67	66.67	66.67	41	2.16	94.01	65.85	67.5	66.67
Fastalign	8	1.33	95.6	62.5	83.33	71.43	27	1.42	95.34	85.19	57.5	68.66
HMM Giza++	22	3.67	80.22	22.73	83.33	35.71	79	4.16	88.69	43.04	85	57.14
HMM+NN	4	0.67	97.8	100	66.67	80	36	1.89	94.68	72.22	65	68.42
HMM+NNCtxCc	4	0.67	97.8	100	66.67	80	26	1.37	94.24	76.92	50	60.61
HMM+NNCtxCNN	5	0.83	98.9	100	83.33	90.91	28	1.47	95.12	82.14	57.5	67.65
HMM+NNCharTgt	5	0.83	96.7	80	66.67	72.73	31	1.63	96.67	90.32	70	78.87
HMM+NNCharWord	5	0.83	98.9	100	83.33	90.91	34	1.79	96.45	85.29	72.5	78.38
HMM+NNCharBoth	7	1.17	96.7	71.43	83.33	76.92	30	1.58	96.45	90	67.5	77.14
HMM+NNCharJT	5	0.83	96.7	80	66.67	72.73	31	1.63	96.67	90.32	70	78.87
HMM+NNCharJB	5	0.83	96.7	80	66.67	72.73	32	1.68	96.45	87.5	70	77.78
IBM-4 Giza++	15	2.5	87.91	33.33	83.33	47.62	72	3.79	90.24	47.22	85	60.71

Table B.20: Models for English-German: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction German-English and in the direction English-German

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	1961	4.25	85.54	15.96	56.09	24.85	3365	2.86	90.68	23.6	46.06	31.2
IBM-1+NN	582	1.26	93.19	21.31	22.22	21.75	1131	0.96	93.55	19.01	12.47	15.06
IBM-1+NNCtxCc	709	1.54	92.38	19.04	24.19	21.31	1458	1.24	92.62	13.99	11.83	12.82
IBM-1+NNCtxCNN	572	1.24	92.96	18.18	18.64	18.41	1362	1.16	93.01	16.81	13.28	14.84
IBM-1+NNChar	637	1.38	92.98	21.66	24.73	23.1	1817	1.55	93.54	30.6	32.25	31.4
IBM-1+NNCharWord	767	1.66	92.08	18.77	25.81	21.74	1596	1.36	93.99	33.21	30.74	31.93
Fastalign	700	1.52	95.94	51.86	65.05	57.71	1489	1.27	95.84	55.41	47.85	51.35
HMM Giza++	1623	3.52	89.42	24.52	71.33	36.5	2878	2.45	93.61	38.26	63.86	47.85
HMM+NN	521	1.13	96.63	61.23	57.17	59.13	1409	1.2	96.32	62.17	50.81	55.92
HMM+NNCtxCc	434	0.94	96.85	66.82	51.97	58.47	1142	0.97	96.16	62.35	41.3	49.69
HMM+NNCtxCNN	458	0.99	97.17	70.52	57.89	63.58	1408	1.2	96.23	60.94	49.77	54.79
HMM+NNCharTgt	461	1	97.14	69.85	57.71	63.2	1205	1.02	97.23	78.42	54.81	64.53
HMM+NNCharWord	512	1.11	97.11	67.58	62.01	64.67	1257	1.07	97.46	80.67	58.82	68.03
HMM+NNCharBoth	422	0.92	97.02	69.91	52.87	60.2	1176	1	97.33	80.61	54.99	65.38
HMM+NNCharJT	428	0.93	96.96	68.69	52.69	59.63	1044	0.89	97.13	80.94	49.01	61.05
HMM+NNCharJB	520	1.13	96.2	55.77	51.97	53.8	1272	1.08	97.17	75.94	56.03	64.49
IBM-4 Giza++	1468	3.18	90.83	28.13	74.01	40.77	2430	2.07	95	46.79	65.95	54.74

Table B.21: Models for English-Czech: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Czech-English and in the direction English-Czech

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	82	6.31	78.74	19.51	76.19	31.07	137	2.49	89.77	31.39	48.86	38.22
IBM-1+NN	17	1.31	92.81	41.18	33.33	36.84	42	0.76	93.38	47.62	22.73	30.77
IBM-1+NNCtxCc	10	0.77	92.51	30	14.29	19.35	57	1.04	91.39	24.56	15.91	19.31
IBM-1+NNCtxCNN	9	0.69	95.81	88.89	38.1	53.33	53	0.96	92.86	41.51	25	31.21
IBM-1+NNChar	12	0.92	92.51	33.33	19.05	24.24	77	1.4	90.8	25.97	22.73	24.24
IBM-1+NNCharWord	13	1	94.01	53.85	33.33	41.18	80	1.45	91.76	35	31.82	33.33
Fastalign	17	1.31	97.01	82.35	66.67	73.68	57	1.04	95.36	71.93	46.59	56.55
HMM Giza++	53	4.08	87.43	30.19	76.19	43.24	182	3.31	88.96	32.97	68.18	44.44
HMM+NN	15	1.15	94.01	53.33	38.1	44.44	71	1.29	95.22	66.2	53.41	59.12
HMM+NNCtxCc	11	0.85	95.81	81.82	42.86	56.25	59	1.07	94.92	66.1	44.32	53.06
HMM+NNCtxCNN	11	0.85	95.81	81.82	42.86	56.25	59	1.07	95.66	74.58	50	59.86
HMM+NNCharTgt	14	1.08	95.51	71.43	47.62	57.14	70	1.27	96.03	74.29	59.09	65.82
HMM+NNCharWord	14	1.08	96.11	78.57	52.38	62.86	67	1.22	96.84	83.58	63.64	72.26
HMM+NNCharBoth	12	0.92	93.71	50	28.57	36.36	70	1.27	95.88	72.86	57.95	64.56
HMM+NNCharJT	12	0.92	95.51	75	42.86	54.55	68	1.24	95.58	70.59	54.55	61.54
HMM+NNCharJB	12	0.92	96.11	83.33	47.62	60.61	71	1.29	95.51	69.01	55.68	61.64
IBM-4 Giza++	50	3.85	89.52	36	85.71	50.7	167	3.04	89.92	35.33	67.05	46.27

Table B.22: Models for English-Romanian: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Romanian-English and in the direction English-Romanian

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	642	3.75	92.43	31.62	65.48	42.65	412	4.12	89.35	16.99	51.47	25.55
IBM-1+NN	653	3.82	91.77	28.33	59.68	38.42	581	5.81	84.1	9.29	39.71	15.06
IBM-1+NNCtxCc	274	1.6	95.92	52.92	46.77	49.66	151	1.51	95.01	31.79	35.29	33.45
IBM-1+NNCtxCNN	751	4.39	90.72	26.1	63.23	36.95	351	3.51	89.95	14.53	37.5	20.94
IBM-1+NNChar	664	3.88	91.51	27.26	58.39	37.17	107	1.07	96.74	55.14	43.38	48.56
IBM-1+NNCharWord	871	5.09	88.94	22.04	61.94	32.51	257	2.57	92.45	20.23	38.24	26.46
Fastalign	231	1.35	96.8	67.1	50	57.3	100	1	96.45	50	36.76	42.37
HMM Giza++	639	3.74	92.58	32.39	66.77	43.62	442	4.42	89.03	17.87	58.09	27.34
HMM+NN	291	1.7	96.82	63.92	60	61.9	88	0.88	97.34	69.32	44.85	54.46
HMM+NNCtxCc	350	2.05	95.89	52	58.71	55.15	293	2.93	92.25	22.53	48.53	30.77
HMM+NNCtxCNN	171	1	97.38	85.38	47.1	60.71	151	1.51	95.38	36.42	40.44	38.33
HMM+NNCharTgt	280	1.64	96.34	58.21	52.58	55.25	160	1.6	96.87	55	64.71	59.46
HMM+NNCharWord	320	1.87	96.73	61.56	63.55	62.54	207	2.07	95.8	43.96	66.91	53.06
HMM+NNCharBoth	189	1.11	97.54	85.19	51.94	64.53	161	1.61	96.89	55.28	65.44	59.93
HMM+NNCharJT	516	3.02	94.64	42.64	70.97	53.27	89	0.89	98.15	86.52	56.62	68.44
HMM+NNCharJB	553	3.23	94.88	44.67	79.68	57.24	154	1.54	96.14	46.1	52.21	48.97
IBM-4 Giza++	667	3.9	92.74	34.03	73.23	46.47	425	4.25	89.48	18.59	58.09	28.16

Table B.23: Models for English-Japanese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Japanese-English and in the direction English-Japanese

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	3054	3.39	91.52	32.65	56.94	41.5	1147	2.76	90.04	23.63	64.68	34.61
IBM-1+NN	1512	1.68	96.23	66.53	57.45	61.66	390	0.94	96.97	63.85	59.43	61.56
IBM-1+NNCtxCc	2527	2.8	94.27	47.01	67.85	55.54	449	1.08	96.85	60.58	64.92	62.67
IBM-1+NNCtxCNN	2715	3.01	93.32	41.44	64.25	50.38	393	0.95	97.59	71.76	67.3	69.46
IBM-1+NNChar	2474	2.74	93.47	41.63	58.82	48.76	415	1	97.2	65.78	65.16	65.47
IBM-1+NNCharWord	1290	1.43	96.99	79.15	58.31	67.15	266	0.64	97.34	77.44	49.16	60.15
Fastalign	1280	1.42	96.32	70.78	51.74	59.78	488	1.18	97.11	62.5	72.79	67.25
HMM Giza++	2853	3.16	93.37	42.13	68.65	52.22	1090	2.63	91.1	27.25	70.88	39.36
HMM+NN	2715	3.01	94.22	46.96	72.82	57.1	361	0.87	97.04	65.93	56.8	61.03
HMM+NNCtxCc	1520	1.69	97.05	75.39	65.45	70.07	404	0.97	97.56	70.79	68.26	69.5
HMM+NNCtxCNN	2542	2.82	95.18	53.03	76.98	62.8	404	0.97	97.56	70.79	68.26	69.5
HMM+NNCharTgt	2007	2.23	95.95	60.14	68.93	64.24	427	1.03	97.57	69.79	71.12	70.45
HMM+NNCharWord	2238	2.48	95.56	56.21	71.84	63.07	267	0.64	97.98	89.51	57.04	69.68
HMM+NNCharBoth	2022	2.24	95.99	60.39	69.73	64.72	360	0.87	98.02	80	68.74	73.94
HMM+NNCharJT	2575	2.85	95.02	51.92	76.36	61.81	319	0.77	98.21	86.83	66.11	75.07
HMM+NNCharJB	2438	2.7	95.2	53.28	74.19	62.02	324	0.78	98.26	87.04	67.3	75.91
IBM-4 Giza++	2586	2.87	94.71	49.92	73.73	59.53	956	2.3	92.75	32.95	75.18	45.82

Table B.24: Models for English-Vietnamese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Vietnamese-English and in the direction English-Vietnamese

B.7 Unknown words

B.7.1 Unknown words in the case of a full vocabulary

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	680	4.33	79.06	22.65	52.38	31.62	298	4.66	78.04	21.14	62.38	31.58
IBM-1+NN	128	0.82	89.37	32.81	14.29	19.91	93	1.45	90.02	37.63	34.65	36.08
IBM-1+NNCtxCc	125	0.8	89.41	32.8	13.95	19.57	105	1.64	90.19	40	41.58	40.78
IBM-1+NNCtxCNN	174	1.11	89.81	41.38	24.49	30.77	97	1.52	89.7	36.08	34.65	35.35
IBM-1+NNChar	188	1.2	89.31	37.77	24.15	29.46	109	1.7	88.42	30.28	32.67	31.43
IBM-1+NNCharWord	204	1.3	90.38	47.06	32.65	38.55	137	2.14	88.9	36.5	49.5	42.02
Fastalign	269	1.71	91.86	56.51	51.7	54	82	1.28	93.81	64.63	52.48	57.92
HMM Giza++	432	2.75	88.05	40.05	58.84	47.66	226	3.53	83.67	27.43	61.39	37.92
HMM+NN	176	1.12	92.46	65.34	39.12	48.94	94	1.47	93	57.45	53.47	55.38
HMM+NNCtxCc	131	0.83	92.8	74.81	33.33	46.12	80	1.25	94.29	68.75	54.46	60.77
HMM+NNCtxCNN	151	0.96	93.49	78.81	40.48	53.48	79	1.23	94.05	67.09	52.48	58.89
HMM+NNCharTgt	153	0.97	93.43	77.78	40.48	53.24	70	1.09	94.45	72.86	50.5	59.65
HMM+NNCharWord	152	0.97	94.15	85.53	44.22	58.3	87	1.36	94.69	70.11	60.4	64.89
HMM+NNCharBoth	144	0.92	93.65	81.94	40.14	53.88	76	1.19	93.48	63.16	47.52	54.24
HMM+NNCharJT	127	0.81	93.3	81.89	35.37	49.41	76	1.19	93.16	60.53	45.54	51.98
HMM+NNCharJB	128	0.82	93.59	85.16	37.07	51.66	71	1.11	93.56	64.79	45.54	53.49
IBM-4 Giza++	388	2.47	88.81	42.01	55.44	47.8	194	3.03	86.4	32.47	62.38	42.71

Table B.25: Models for English-French: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction French-English and in the direction English-French.

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	112	7.47	73.74	13.39	68.18	22.39	393	6.78	79.13	25.45	77.52	38.31
IBM-1+NN	15	1	92.17	20	13.64	16.22	94	1.62	91.51	48.94	35.66	41.26
IBM-1+NNCtxCc	17	1.13	92.17	23.53	18.18	20.51	106	1.83	91.64	50	41.09	45.11
IBM-1+NNCtxCNN	18	1.2	91.41	16.67	13.64	15	98	1.69	92.55	57.14	43.41	49.34
IBM-1+NNChar	20	1.33	93.94	45	40.91	42.86	120	2.07	93.84	64.17	59.69	61.85
IBM-1+NNCharWord	21	1.4	92.68	33.33	31.82	32.56	115	1.98	94.43	68.7	61.24	64.75
Fastalign	15	1	94.19	46.67	31.82	37.84	48	0.83	92.81	68.75	25.58	37.29
HMM Giza++	88	5.87	81.82	21.59	86.36	34.55	323	5.57	84.83	33.75	84.5	48.23
HMM+NN	10	0.67	95.45	70	31.82	43.75	102	1.76	96.18	84.31	66.67	74.46
HMM+NNCtxCc	12	0.8	95.45	66.67	36.36	47.06	84	1.45	94.36	75	48.84	59.15
HMM+NNCtxCNN	14	0.93	95.45	64.29	40.91	50	90	1.55	95.92	86.67	60.47	71.23
HMM+NNCharTgt	14	0.93	96.97	85.71	54.55	66.67	89	1.53	96.63	93.26	64.34	76.15
HMM+NNCharWord	10	0.67	95.45	70	31.82	43.75	103	1.78	97.28	92.23	73.64	81.9
HMM+NNCharBoth	14	0.93	96.46	78.57	50	61.11	96	1.66	96.69	90.62	67.44	77.33
HMM+NNCharJT	13	0.87	96.72	84.62	50	62.86	92	1.59	96.57	91.3	65.12	76.02
HMM+NNCharJB	16	1.07	96.46	75	54.55	63.16	98	1.69	97.21	93.88	71.32	81.06
IBM-4 Giza++	76	5.07	84.85	25	86.36	38.78	270	4.66	88.66	41.48	86.82	56.14

Table B.26: Models for English-German: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction German-English and in the direction English-German.

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	6931	4.33	85.4	16.87	55.53	25.87	8487	3.33	89.86	20.9	48.91	29.29
IBM-1+NN	2073	1.3	92.8	21.13	20.81	20.97	2779	1.09	93.7	19.5	14.94	16.92
IBM-1+NNCtxCc	2364	1.48	92.19	18.78	21.09	19.87	3455	1.36	92.97	16.5	15.72	16.1
IBM-1+NNCtxCNN	2072	1.3	92.53	18.1	17.81	17.96	3433	1.35	93.06	17.39	16.46	16.91
IBM-1+NNChar	2210	1.38	92.9	23.94	25.13	24.52	4393	1.73	92.9	23.06	27.93	25.26
IBM-1+NNCharWord	2761	1.73	91.85	20.39	26.75	23.14	3690	1.45	93.68	26.8	27.27	27.03
Fastalign	2118	1.32	96.29	59.54	59.9	59.72	3056	1.2	96.22	57.04	48.06	52.16
HMM Giza++	5702	3.57	89.75	27.24	73.78	39.78	7488	2.94	92.59	32.41	66.91	43.67
HMM+NN	1925	1.2	96.24	59.84	54.73	57.17	3115	1.22	96.34	58.62	50.34	54.17
HMM+NNCtxCc	1584	0.99	96.44	64.84	48.79	55.68	2596	1.02	96.39	61.13	43.76	51
HMM+NNCtxCNN	1698	1.06	96.51	64.84	52.3	57.9	3233	1.27	96.32	58	51.7	54.66
HMM+NNCharTgt	1665	1.04	96.49	64.86	51.31	57.29	2489	0.98	97.18	74.97	51.45	61.02
HMM+NNCharWord	1853	1.16	96.59	64.6	56.86	60.49	2964	1.16	97.24	71.9	58.75	64.66
HMM+NNCharBoth	1635	1.02	96.24	61.59	47.84	53.85	2611	1.03	97.15	73.42	52.85	61.46
HMM+NNCharJT	1533	0.96	96.32	63.54	46.27	53.55	2130	0.84	96.97	74.98	44.03	55.48
HMM+NNCharJB	1931	1.21	95.91	55.88	51.26	53.47	3016	1.18	96.82	65.58	54.54	59.55
IBM-4 Giza++	5132	3.21	91.11	30.79	75.06	43.66	6058	2.38	94.39	40.82	68.18	51.07

Table B.27: Models for English-Czech: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Czech-English and in the direction English-Czech.

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	238	6.61	78.68	13.45	65.31	22.3	263	4.24	86.77	21.29	58.33	31.2
IBM-1+NN	32	0.89	95.12	46.88	30.61	37.04	68	1.1	93.79	35.29	25	29.27
IBM-1+NNCtxCc	35	0.97	94.65	40	28.57	33.33	58	0.94	92.93	18.97	11.46	14.29
IBM-1+NNCtxCNN	36	1	94.74	41.67	30.61	35.29	67	1.08	93.95	37.31	26.04	30.67
IBM-1+NNChar	37	1.03	93.12	18.92	14.29	16.28	76	1.23	94.11	40.79	32.29	36.05
IBM-1+NNCharWord	45	1.25	94.84	44.44	40.82	42.55	95	1.53	93.63	37.89	37.5	37.7
Fastalign	36	1	96.46	66.67	48.98	56.47	61	0.98	95.45	59.02	37.5	45.86
HMM Giza++	189	5.25	83.94	18.52	71.43	29.41	242	3.9	89.93	30.99	78.12	44.38
HMM+NN	42	1.17	96.65	66.67	57.14	61.54	70	1.13	96.57	72.86	53.12	61.45
HMM+NNCtxCc	38	1.06	96.85	71.05	55.1	62.07	68	1.1	95.93	64.71	45.83	53.66
HMM+NNCtxCNN	32	0.89	97.42	84.38	55.1	66.67	74	1.19	96.04	64.86	50	56.47
HMM+NNCharTgt	35	0.97	97.32	80	57.14	66.67	81	1.31	97.27	77.78	65.62	71.19
HMM+NNCharWord	46	1.28	97.42	73.91	69.39	71.58	84	1.35	97.54	79.76	69.79	74.44
HMM+NNCharBoth	33	0.92	96.75	72.73	48.98	58.54	80	1.29	97.32	78.75	65.62	71.59
HMM+NNCharJT	38	1.06	97.04	73.68	57.14	64.37	79	1.27	96.95	74.68	61.46	67.43
HMM+NNCharJB	43	1.19	97.32	74.42	65.31	69.57	81	1.31	96.73	71.6	60.42	65.54
IBM-4 Giza++	159	4.42	87.76	25.16	81.63	38.46	214	3.45	91.22	34.11	76.04	47.1

Table B.28: Models for English-Romanian: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Romanian-English and in the direction English-Romanian.

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	2649	4.73	89.29	25.78	57.44	35.59	1192	4.97	88.47	15.44	58.04	24.39
IBM-1+NN	2326	4.15	90.46	28.2	55.17	37.33	1377	5.74	85.33	8.79	38.17	14.29
IBM-1+NNCtxCc	926	1.65	95.7	60.58	47.18	53.05	398	1.66	94.78	24.87	31.23	27.69
IBM-1+NNCtxCNN	2552	4.56	89.94	27.78	59.63	37.9	864	3.6	90.37	13.19	35.96	19.31
IBM-1+NNChar	2439	4.36	89.91	26.61	54.58	35.78	219	0.91	97.45	64.84	44.79	52.99
IBM-1+NNCharWord	3067	5.48	88.09	24.55	63.33	35.39	582	2.42	93.6	22.85	41.96	29.59
Fastalign	684	1.22	95.6	62.72	36.08	45.81	167	0.7	96.97	55.09	29.02	38.02
HMM Giza++	2610	4.66	89.95	28.35	62.24	38.96	1121	4.67	89.49	17.75	62.78	27.68
HMM+NN	1063	1.9	95.54	57.48	51.39	54.26	174	0.72	97.99	83.91	46.06	59.47
HMM+NNCtxCc	785	1.4	96.98	81.27	53.66	64.64	638	2.66	92.94	20.06	40.38	26.81
HMM+NNCtxCNN	655	1.17	96.83	84.89	46.76	60.3	323	1.35	96.08	39.01	39.75	39.38
HMM+NNCharTgt	789	1.41	96.78	78.2	51.89	62.39	286	1.19	97.42	60.84	54.89	57.71
HMM+NNCharWord	1097	1.96	96.23	64.54	59.55	61.94	388	1.62	96.59	47.42	58.04	52.2
HMM+NNCharBoth	747	1.33	97.12	85.01	53.41	65.6	367	1.53	96.77	49.59	57.41	53.22
HMM+NNCharJT	1650	2.95	94.42	47.03	65.26	54.67	196	0.82	98.49	92.86	57.41	70.96
HMM+NNCharJB	1912	3.41	93.72	43.2	69.47	53.27	336	1.4	95.99	38.1	40.38	39.2
IBM-4 Giza++	2502	4.47	90.46	29.74	62.57	40.31	1081	4.5	90.06	19.15	65.3	29.61

Table B.29: Models for English-Japanese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Japanese-English and in the direction English-Japanese.

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	14589	3	92.22	23.11	56.59	32.82	7181	2.55	90.93	17.1	64.56	27.04
IBM-1+NN	6023	1.24	97.15	57.5	58.11	57.8	2444	0.87	97.1	45.58	58.57	51.27
IBM-1+NNCtxCc	11927	2.46	94.36	33.05	66.15	44.08	2738	0.97	97	44.74	64.41	52.8
IBM-1+NNCtxCNN	12060	2.48	94.16	31.74	64.24	42.49	2191	0.78	97.84	57.37	66.09	61.42
IBM-1+NNChar	8627	1.78	95.55	38.79	56.15	45.88	2054	0.73	97.89	58.76	63.46	61.02
IBM-1+NNCharWord	4399	0.91	97.95	76.34	56.35	64.84	1319	0.47	98.2	72.33	50.16	59.24
Fastalign	5808	1.2	96.62	49.62	48.36	48.98	3141	1.11	96.89	44.09	72.82	54.93
HMM Giza++	13469	2.77	93.79	31.23	70.58	43.3	6888	2.44	91.85	20.6	74.61	32.29
HMM+NN	13112	2.7	94.1	32.84	72.26	45.16	2141	0.76	97.87	58.15	65.46	61.59
HMM+NNCtxCc	5778	1.19	97.81	68.02	65.95	66.97	2187	0.78	97.94	59.12	67.98	63.24
HMM+NNCtxCNN	9426	1.94	96.17	45.6	72.13	55.87	2549	0.9	97.49	51.31	68.77	58.77
HMM+NNCharTgt	7931	1.63	96.58	49.3	65.62	56.3	2278	0.81	97.92	58.34	69.87	63.59
HMM+NNCharWord	9806	2.02	95.8	42.37	69.73	52.71	1361	0.48	98.67	84.06	60.15	70.12
HMM+NNCharBoth	8308	1.71	96.45	47.93	66.82	55.82	1849	0.66	98.53	72.47	70.45	71.45
HMM+NNCharJT	9403	1.94	96.31	46.85	73.92	57.35	1626	0.58	98.67	78.6	67.19	72.45
HMM+NNCharJB	9402	1.94	96.32	46.99	74.14	57.52	1720	0.61	98.77	79.19	71.61	75.21
IBM-4 Giza++	11283	2.32	95.07	37.69	71.35	49.32	5977	2.12	93.41	25.63	80.55	38.89

Table B.30: Models for English-Vietnamese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Vietnamese-English and in the direction English-Vietnamese.

B.7.2 Unknown words in the case of 50K word vocabulary

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	61	0.35	89.67	32.79	5.67	9.66	116	1.12	89.49	35.34	22.91	27.8
IBM-1+NN	145	0.82	88.9	33.1	13.6	19.28	153	1.47	89.54	39.22	33.52	36.14
IBM-1+NNCtxCc	136	0.77	89.26	36.76	14.16	20.45	173	1.66	89.54	40.46	39.11	39.77
IBM-1+NNCtxCNN	195	1.11	89.4	42.05	23.23	29.93	171	1.64	89.44	39.77	37.99	38.86
IBM-1+NNChar	217	1.23	88.9	38.71	23.8	29.47	188	1.81	88.1	33.51	35.2	34.33
IBM-1+NNCharWord	245	1.39	89.89	47.35	32.86	38.8	205	1.97	88.45	36.59	41.9	39.06
Fastalign	106	0.6	91.3	67.92	20.4	31.37	126	1.21	91.76	54.76	38.55	45.25
HMM Giza++	132	0.75	93.07	88.64	33.14	48.25	88	0.85	92.55	65.91	32.4	43.45
HMM+NN	195	1.11	92.16	67.69	37.39	48.18	152	1.46	92.25	57.24	48.6	52.57
HMM+NNCtxCc	149	0.85	92.38	75.84	32.01	45.02	128	1.23	93.04	64.84	46.37	54.07
HMM+NNCtxCNN	171	0.97	93.1	80.12	38.81	52.29	123	1.18	92.79	63.41	43.58	51.66
HMM+NNCharTgt	170	0.97	92.9	78.24	37.68	50.86	120	1.15	94.23	75.83	50.84	60.87
HMM+NNCharWord	171	0.97	93.65	85.96	41.64	56.11	144	1.38	94.03	70.14	56.42	62.54
HMM+NNCharBoth	163	0.93	93.15	82.21	37.96	51.94	121	1.16	93.19	66.94	45.25	54
HMM+NNCharJT	139	0.79	92.82	83.45	32.86	47.15	112	1.08	93.14	67.86	42.46	52.23
HMM+NNCharJB	142	0.81	93.12	86.62	34.84	49.7	106	1.02	93.34	70.75	41.9	52.63
IBM-4 Giza++	82	0.47	91.63	80.49	18.7	30.34	85	0.82	92.3	63.53	30.17	40.91

Table B.31: Models for English-French: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction French-English and in the direction English-French. Note that the training data for for all models including the baselines only has a vocabulary containing the most frequent 50K words.

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	27	1.04	92.73	33.33	27.27	30	82	0.44	91.3	15.85	3.65	5.94
IBM-1+NN	26	1	92.91	34.62	27.27	30.51	273	1.46	91.95	45.42	34.83	39.43
IBM-1+NNCtxCc	28	1.08	92.91	35.71	30.3	32.79	300	1.6	92.27	48.33	40.73	44.21
IBM-1+NNCtxCNN	31	1.19	92.39	32.26	30.3	31.25	317	1.7	92.5	50.16	44.66	47.25
IBM-1+NNChar	35	1.35	93.43	42.86	45.45	44.12	354	1.89	94	60.17	59.83	60
IBM-1+NNCharWord	34	1.31	93.25	41.18	42.42	41.79	346	1.85	94.21	61.85	60.11	60.97
Fastalign	20	0.77	94.98	60	36.36	45.28	203	1.09	94.36	71.92	41.01	52.24
HMM Giza++	19	0.73	93.77	42.11	24.24	30.77	81	0.43	93.39	76.54	17.42	28.38
HMM+NN	18	0.69	95.67	72.22	39.39	50.98	309	1.65	95.58	73.79	64.04	68.57
HMM+NNCtxCc	22	0.85	96.37	77.27	51.52	61.82	260	1.39	94.84	71.54	52.25	60.39
HMM+NNCtxCNN	24	0.92	96.37	75	54.55	63.16	277	1.48	95.96	79.78	62.08	69.83
HMM+NNCharTgt	22	0.85	97.4	90.91	60.61	72.73	271	1.45	96.81	87.82	66.85	75.92
HMM+NNCharWord	19	0.73	96.19	78.95	45.45	57.69	303	1.62	97.02	85.48	72.75	78.6
HMM+NNCharBoth	24	0.92	97.06	83.33	60.61	70.18	271	1.45	96.81	87.82	66.85	75.92
HMM+NNCharJT	22	0.85	96.71	81.82	54.55	65.45	264	1.41	96.87	89.39	66.29	76.13
HMM+NNCharJB	26	1	96.71	76.92	60.61	67.8	285	1.52	97.02	87.72	70.22	78
IBM-4 Giza++	15	0.58	94.81	60	27.27	37.5	49	0.26	93.05	77.55	10.67	18.77

Table B.32: Models for English-German: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction German-English and in the direction English-German. Note that the training data for for all models including the baselines only has a vocabulary containing the most frequent 50K words.

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	2059	1.09	92.62	13.4	11.26	12.24	3630	0.7	93.84	11.49	5.57	7.5
IBM-1+NN	2433	1.29	92.83	21.33	21.18	21.25	5318	1.03	93.59	19.76	14.04	16.42
IBM-1+NNCtxCc	2925	1.56	92	18.56	22.15	20.2	6595	1.28	92.84	16.03	14.12	15.01
IBM-1+NNCtxCNN	2530	1.35	92.44	18.3	18.89	18.59	6596	1.28	92.95	17.5	15.41	16.39
IBM-1+NNChar	2608	1.39	92.84	23.39	24.89	24.12	8501	1.64	93.3	28.18	32	29.97
IBM-1+NNCharWord	3336	1.77	91.74	20.41	27.78	23.54	7202	1.39	93.92	31.42	30.23	30.81
Fastalign	2258	1.2	95.17	46.99	43.29	45.06	4335	0.84	95.28	45.49	26.34	33.36
HMM Giza++	1642	0.87	95.85	56.82	38.07	45.59	2174	0.42	95.65	54.92	15.95	24.72
HMM+NN	2298	1.22	96.28	59.97	56.22	58.03	6313	1.22	96.24	59.51	50.18	54.45
HMM+NNCtxCc	1867	0.99	96.5	65.4	49.82	56.55	5141	0.99	96.15	60.34	41.43	49.13
HMM+NNCtxCNN	2012	1.07	96.61	65.71	53.94	59.24	6456	1.25	96.24	59.37	51.2	54.98
HMM+NNCharTgt	1952	1.04	96.6	66.03	52.59	58.55	5173	1	97.23	77.67	53.67	63.48
HMM+NNCharWord	2214	1.18	96.69	65.27	58.96	61.95	5882	1.14	97.37	76.27	59.92	67.11
HMM+NNCharBoth	1890	1	96.36	63.23	48.76	55.06	5311	1.03	97.26	77.44	54.94	64.28
HMM+NNCharJT	1813	0.96	96.4	64.42	47.65	54.78	4476	0.87	97.09	79.33	47.43	59.37
HMM+NNCharJB	2290	1.22	95.98	56.46	52.75	54.55	5904	1.14	97.06	71.88	56.68	63.39
IBM-4 Giza++	1251	0.67	95.46	50.6	25.83	34.2	1239	0.24	95.71	62.79	10.39	17.83

Table B.33: Models for English-Czech: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Czech-English and in the direction English-Czech. Note that the training data for for all models including the baselines only has a vocabulary containing the most frequent 50K words.

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	40	0.93	94.98	45	30.51	36.36	33	0.2	94.82	48.48	6.5	11.47
IBM-1+NN	40	0.93	94.98	45	30.51	36.36	163	0.98	93.89	36.2	23.98	28.85
IBM-1+NNCtxCc	45	1.05	94.42	37.78	28.81	32.69	158	0.95	92.78	18.99	12.2	14.85
IBM-1+NNCtxCNN	41	0.95	95.06	46.34	32.2	38	154	0.93	94.04	37.66	23.58	29
IBM-1+NNChar	42	0.98	93.22	19.05	13.56	15.84	227	1.37	93.05	31.28	28.86	30.02
IBM-1+NNCharWord	54	1.26	94.98	46.3	42.37	44.25	243	1.46	93.47	36.63	36.18	36.4
Fastalign	39	0.91	96.17	64.1	42.37	51.02	117	0.7	95.49	63.25	30.08	40.77
HMM Giza++	27	0.63	96.49	77.78	35.59	48.84	39	0.23	95.24	74.36	11.79	20.35
HMM+NN	51	1.19	96.33	62.75	54.24	58.18	186	1.12	96.31	68.82	52.03	59.26
HMM+NNCtxCc	44	1.02	96.73	70.45	52.54	60.19	173	1.04	96.03	66.47	46.75	54.89
HMM+NNCtxCNN	37	0.86	97.45	86.49	54.24	66.67	175	1.05	96.29	69.71	49.59	57.96
HMM+NNCharTgt	44	1.02	97.21	77.27	57.63	66.02	205	1.23	97.13	76.59	63.82	69.62
HMM+NNCharWord	55	1.28	97.29	72.73	67.8	70.18	198	1.19	97.69	84.34	67.89	75.23
HMM+NNCharBoth	42	0.98	96.41	66.67	47.46	55.45	205	1.23	97.04	75.61	63.01	68.74
HMM+NNCharJT	47	1.09	96.97	72.34	57.63	64.15	204	1.23	96.73	72.06	59.76	65.33
HMM+NNCharJB	51	1.19	97.29	74.51	64.41	69.09	205	1.23	96.66	71.22	59.35	64.75
IBM-4 Giza++	23	0.53	96.65	86.96	33.9	48.78	28	0.17	95.26	85.71	9.76	17.52

Table B.34: Models for English-Romanian: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Romanian-English and in the direction English-Romanian. Note that the training data for for all models including the baselines only has a vocabulary containing the most frequent 50K words.

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	452	0.48	94.97	34.51	8.41	13.52	468	0.82	95.69	27.78	18.44	22.17
IBM-1+NN	3533	3.74	91.55	28.84	54.9	37.82	2494	4.38	87.44	10.83	38.3	16.88
IBM-1+NNCtxCc	1496	1.58	95.99	58.89	47.47	52.57	793	1.39	95.23	30.77	34.61	32.58
IBM-1+NNCtxCNN	4150	4.4	90.5	26.94	60.24	37.23	1586	2.79	91.91	18.22	40.99	25.23
IBM-1+NNChar	3879	4.11	90.72	26.45	55.28	35.78	504	0.89	97.24	61.9	44.26	51.61
IBM-1+NNCharWord	4534	4.8	89.6	24.97	60.99	35.43	1152	2.02	93.94	24.91	40.71	30.91
Fastalign	836	0.89	95.25	48.21	21.71	29.94	548	0.96	96.08	38.5	29.93	33.68
HMM Giza++	443	0.47	95.19	44.02	10.51	16.96	374	0.66	96.27	38.77	20.57	26.88
HMM+NN	1783	1.89	95.87	56.14	53.93	55.02	413	0.73	97.78	78.45	45.96	57.96
HMM+NNCtxCc	1785	1.89	96.05	58.1	55.87	56.96	1264	2.22	93.67	24.84	44.54	31.89
HMM+NNCtxCNN	1032	1.09	97.09	84.01	46.71	60.04	711	1.25	96.18	42.62	42.98	42.8
HMM+NNCharTgt	1687	1.79	95.93	57.2	51.99	54.47	653	1.15	97.65	65.85	60.99	63.33
HMM+NNCharWord	1775	1.88	96.57	64	61.21	62.57	870	1.53	96.61	49.31	60.85	54.48
HMM+NNCharBoth	1167	1.24	97.31	83.8	52.69	64.7	773	1.36	97.08	55.63	60.99	58.19
HMM+NNCharJT	2499	2.65	95.11	48.34	65.09	55.48	447	0.79	98.35	89.71	56.88	69.62
HMM+NNCharJB	2933	3.11	94.59	45.04	71.17	55.17	715	1.26	96.45	46.71	47.38	47.04
IBM-4 Giza++	313	0.33	95.37	53.04	8.94	15.31	341	0.6	96.28	37.83	18.3	24.67

Table B.35: Models for English-Japanese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Japanese-English and in the direction English-Japanese. Note that the training data for for all models including the baselines only has a vocabulary containing the most frequent 50K words.

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	3140	0.68	95.76	23.38	13.32	16.97	3506	1.27	94.71	22.3	42.66	29.29
IBM-1+NN	5649	1.22	97.21	56.93	58.38	57.64	2392	0.87	97.1	44.98	58.7	50.93
IBM-1+NNCtxCc	11391	2.46	94.31	31.88	65.93	42.98	2679	0.97	97.01	44.38	64.87	52.7
IBM-1+NNCtxCNN	11470	2.47	94.12	30.61	63.73	41.36	2125	0.77	97.86	57.18	66.28	61.39
IBM-1+NNChar	8075	1.74	95.62	38.14	55.91	45.35	1999	0.73	97.91	58.53	63.83	61.06
IBM-1+NNCharWord	4049	0.87	97.98	75.87	55.76	64.28	1280	0.47	98.22	72.11	50.35	59.3
Fastalign	5249	1.13	96.01	38.08	36.29	37.16	3366	1.22	96.06	35.44	65.08	45.89
HMM Giza++	2660	0.57	96.6	45.34	21.89	29.53	3217	1.17	96.28	37.21	65.3	47.41
HMM+NN	12416	2.68	94.11	32.02	72.15	44.35	2083	0.76	97.89	57.9	65.79	61.59
HMM+NNCtxCc	5368	1.16	97.84	67.27	65.55	66.4	2125	0.77	97.99	59.29	68.74	63.67
HMM+NNCtxCNN	8920	1.92	96.15	44.28	71.7	54.75	2488	0.9	97.49	50.84	69.01	58.55
HMM+NNCharTgt	7448	1.61	96.6	48.35	65.37	55.58	2220	0.81	97.94	58.2	70.49	63.76
HMM+NNCharWord	9281	2	95.79	41.26	69.5	51.78	1317	0.48	98.68	83.9	60.28	70.16
HMM+NNCharBoth	7828	1.69	96.48	47.1	66.93	55.29	1794	0.65	98.55	72.3	70.76	71.52
HMM+NNCharJT	8815	1.9	96.34	46.11	73.79	56.76	1577	0.57	98.68	78.19	67.27	72.32
HMM+NNCharJB	8872	1.91	96.33	46.02	74.12	56.78	1668	0.61	98.78	78.9	71.79	75.18
IBM-4 Giza++	2530	0.55	96.61	45.38	20.84	28.56	3218	1.17	96.26	37.01	64.98	47.16

Table B.36: Models for English-Vietnamese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Vietnamese-English and in the direction English-Vietnamese. Note that the training data for for all models including the baselines only has a vocabulary containing the most frequent 50K words.

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
Fastalign	20	0.77	94.98	60	36.36	45.28	203	1.09	94.36	71.92	41.01	52.24
HMM Giza++	19	0.73	93.77	42.11	24.24	30.77	81	0.43	93.39	76.54	17.42	28.38
IBM-4 Giza++	15	0.58	94.81	60	27.27	37.5	49	0.26	93.05	77.55	10.67	18.77
HMM+NN	10	0.67	95.45	70	31.82	43.75	102	1.76	96.18	84.31	66.67	74.46
HMM+NN+3M	20	0.77	96.02	75	45.45	56.6	309	1.65	95.8	75.4	65.45	70.08
HMM+NN+6M	20	0.77	96.02	75	45.45	56.6	309	1.65	96.01	77.02	66.85	71.58
Fastalign	15	1	94.19	46.67	31.82	37.84	48	0.83	92.81	68.75	25.58	37.29
HMM Giza++	88	5.87	81.82	21.59	86.36	34.55	323	5.57	84.83	33.75	84.5	48.23
IBM-4 Giza++	76	5.07	84.85	25	86.36	38.78	270	4.66	88.66	41.48	86.82	56.14
HMM+NNCharTgt+3M	16	1.07	97.47	87.5	63.64	73.68	93	1.6	96.89	93.55	67.44	78.38
HMM+NNCharTgt+6M	16	1.07	97.47	87.5	63.64	73.68	95	1.64	97.02	93.68	68.99	79.46
HMM+NNCharJB+3M	15	1	96.72	80	54.55	64.86	96	1.66	97.47	96.88	72.09	82.67
HMM+NNCharJB+6M	15	1	97.22	86.67	59.09	70.27	103	1.78	97.67	95.15	75.97	84.48

Table B.37: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction German-English and in the direction English-German. Note that the training data for for all models including the baselines only has a vocabulary containing the most frequent 50K words.

B.8 Part-of-Speech

B.8.1 Alignment links for content and function words

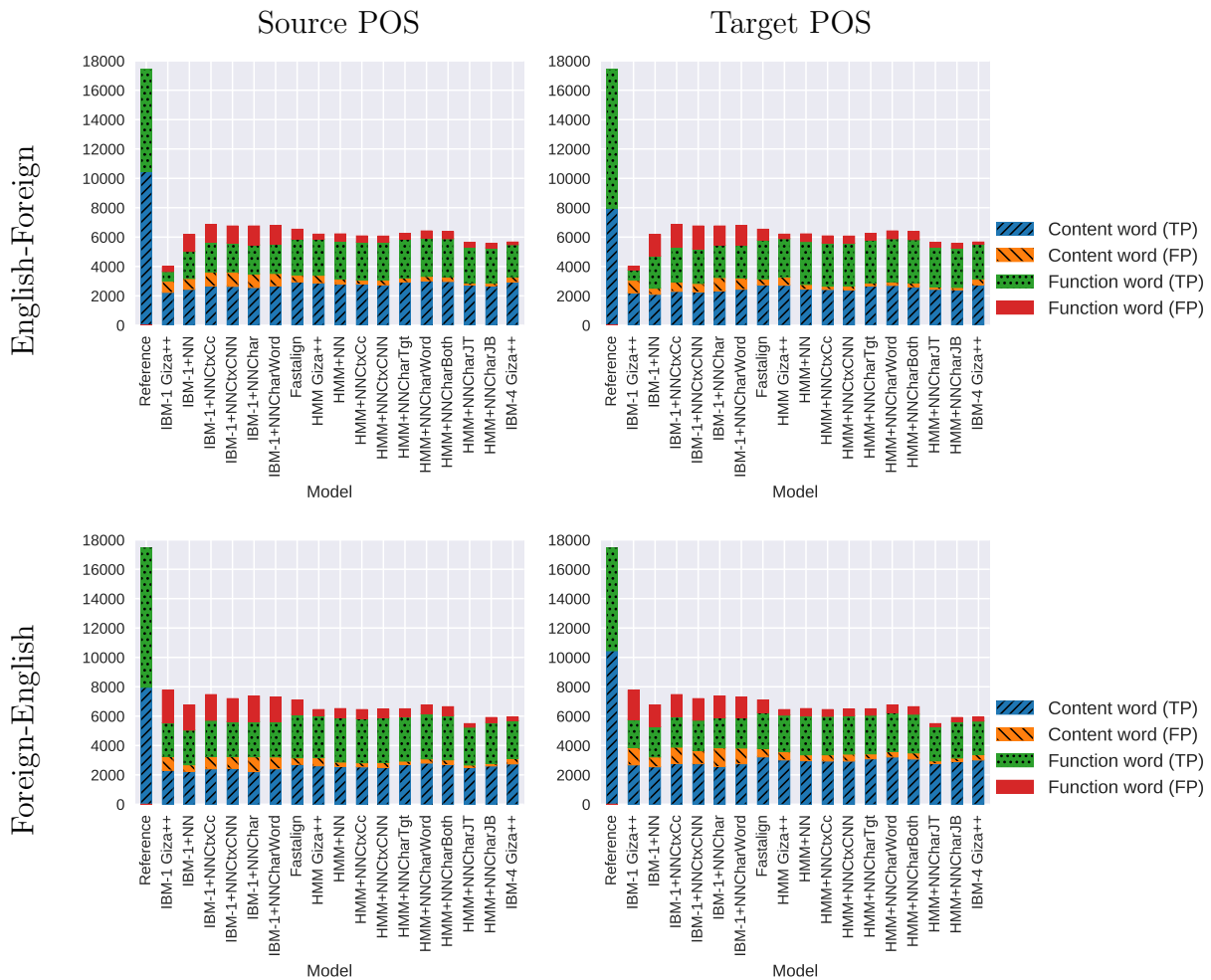


Figure B.37: Neural models for English-French. The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).

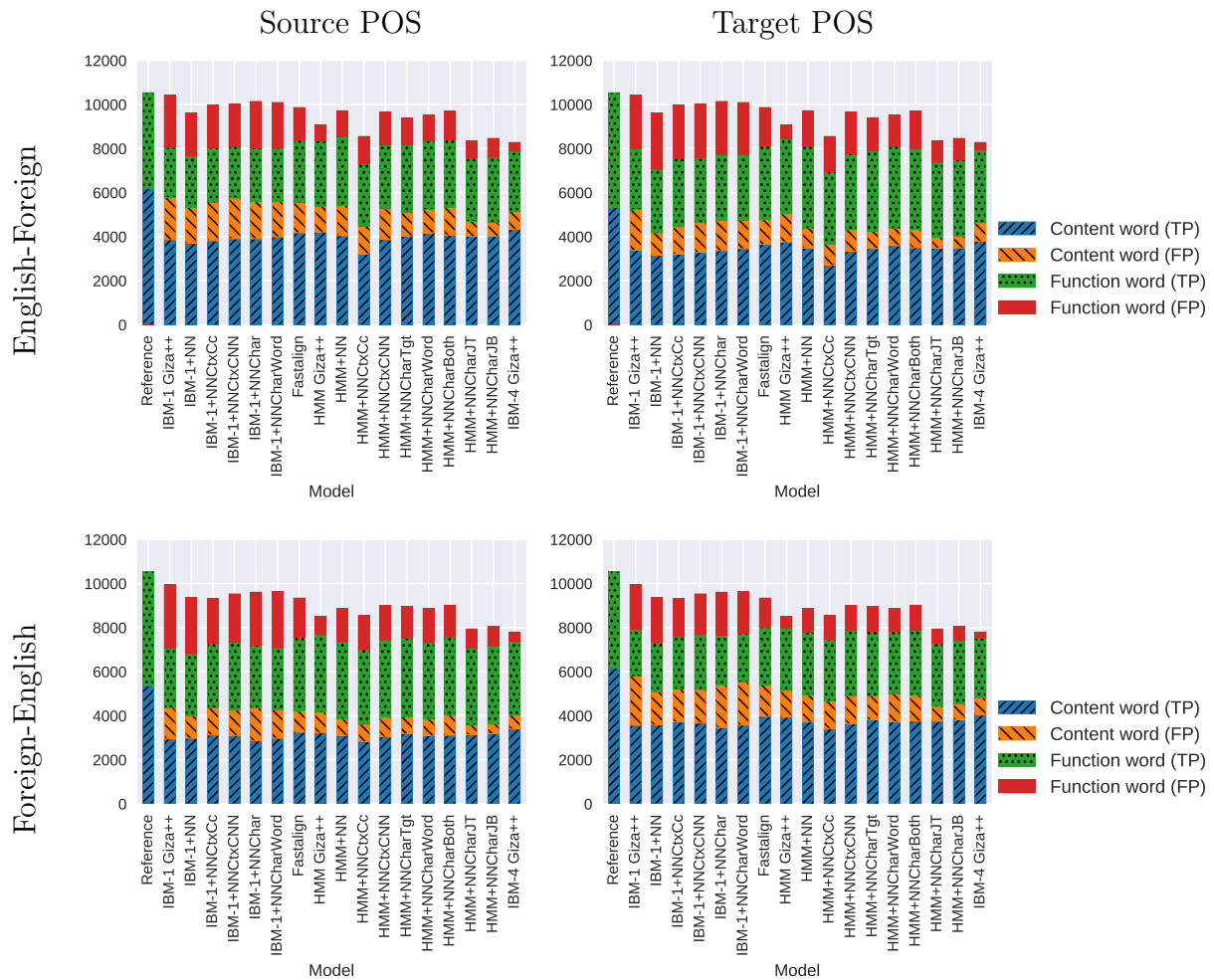


Figure B.38: Neural models for English-German The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).

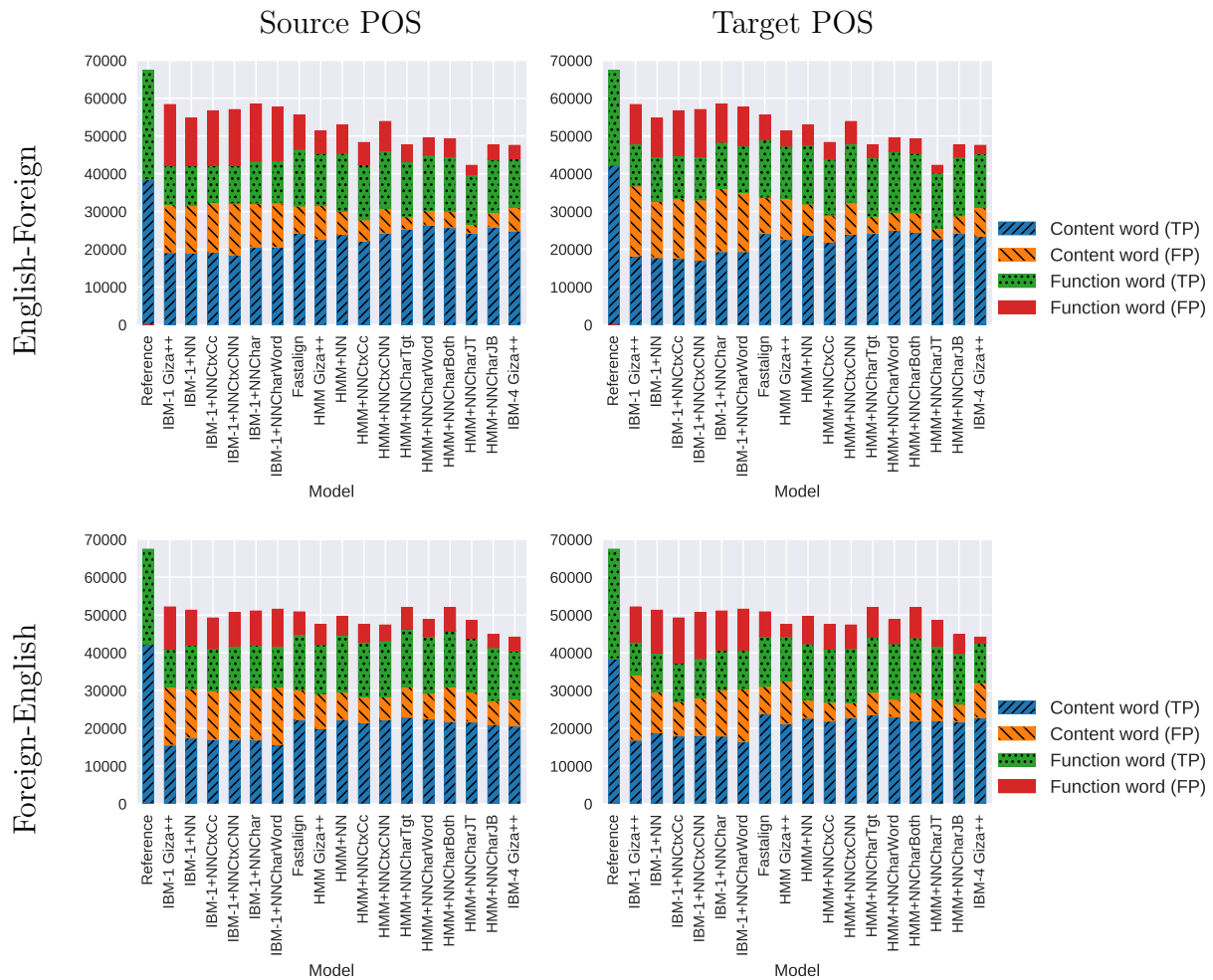


Figure B.39: Neural models for English-Czech. The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).

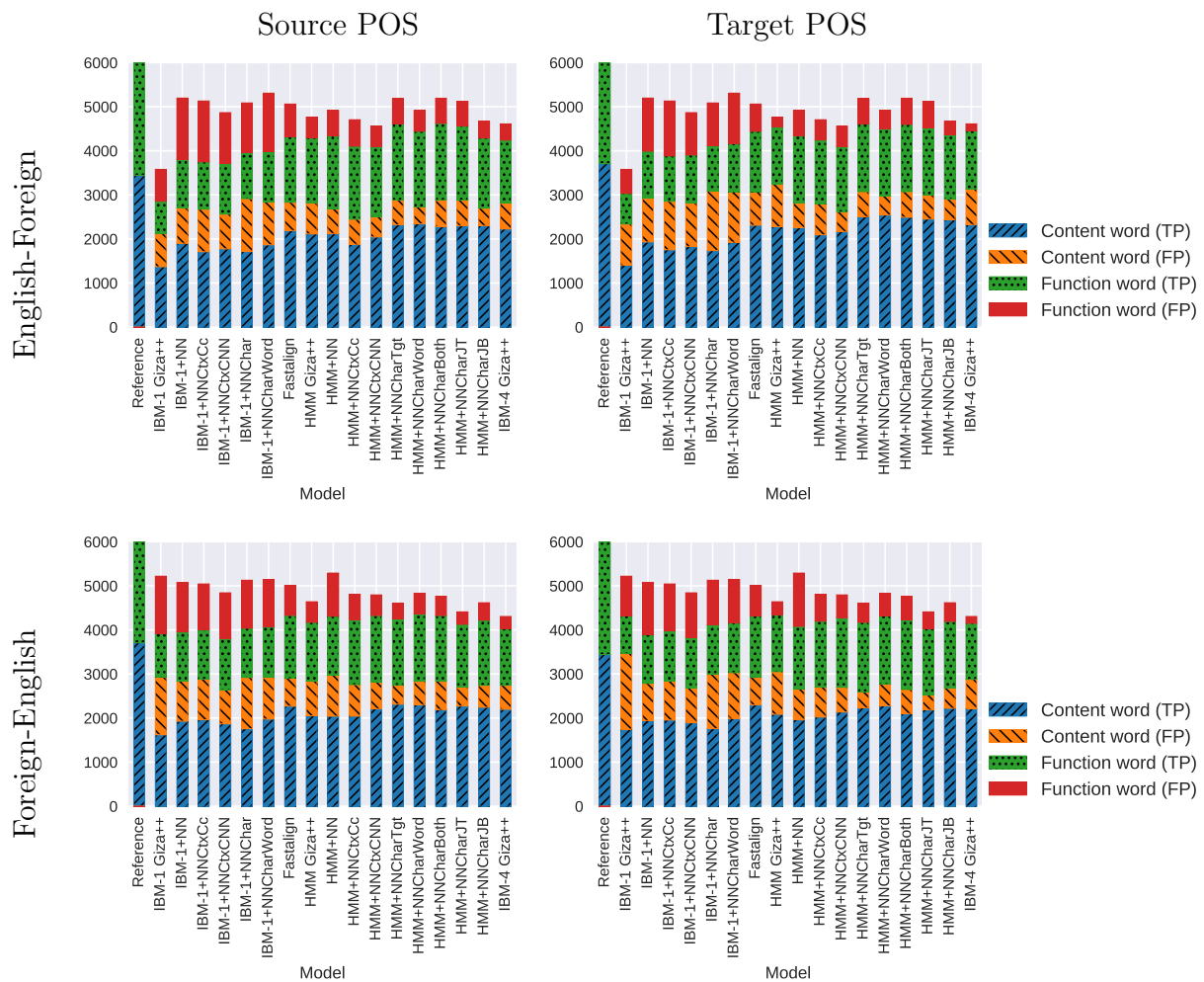


Figure B.40: Neural models for English-Romanian The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).

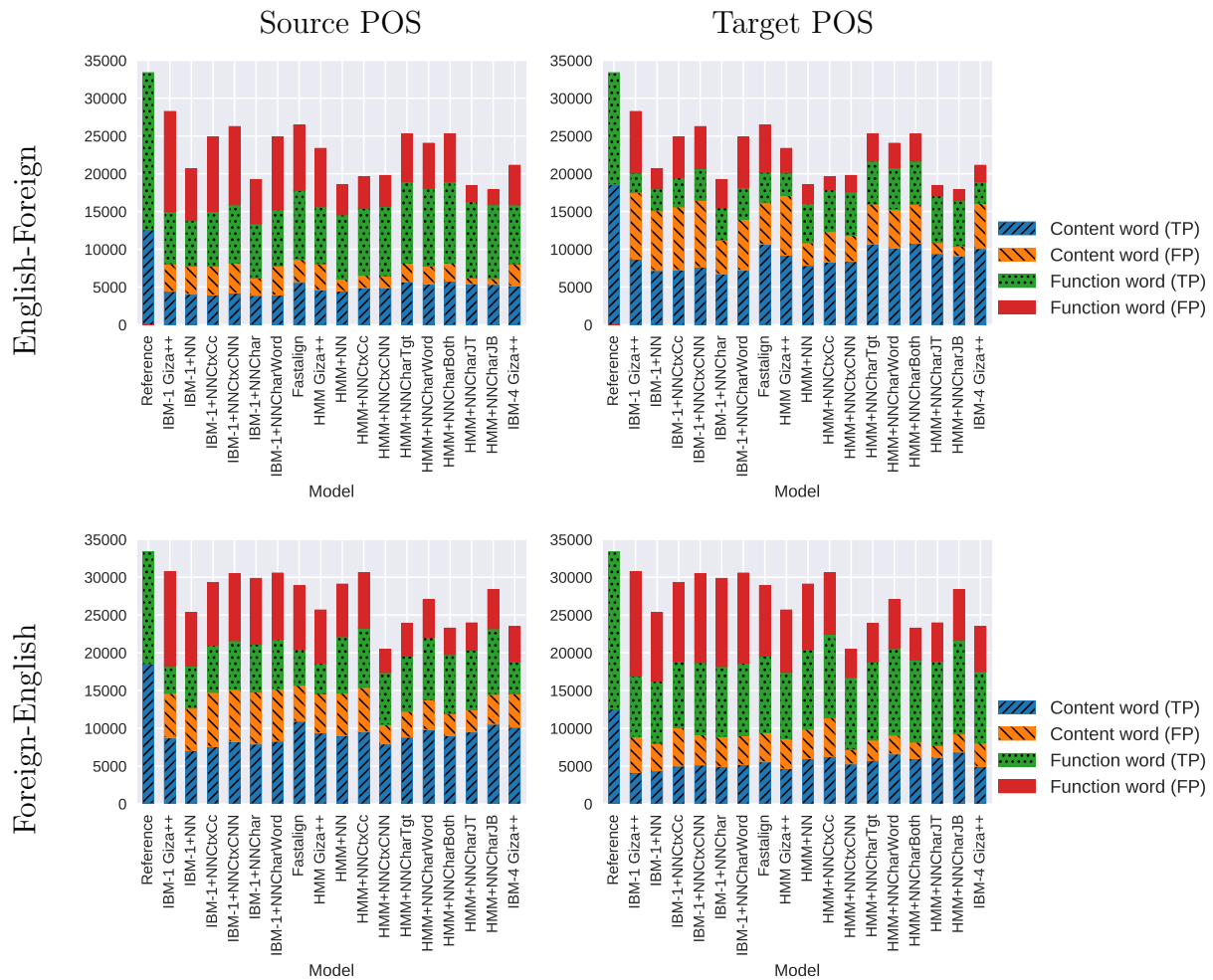


Figure B.41: Neural models for English-Japanese: The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).

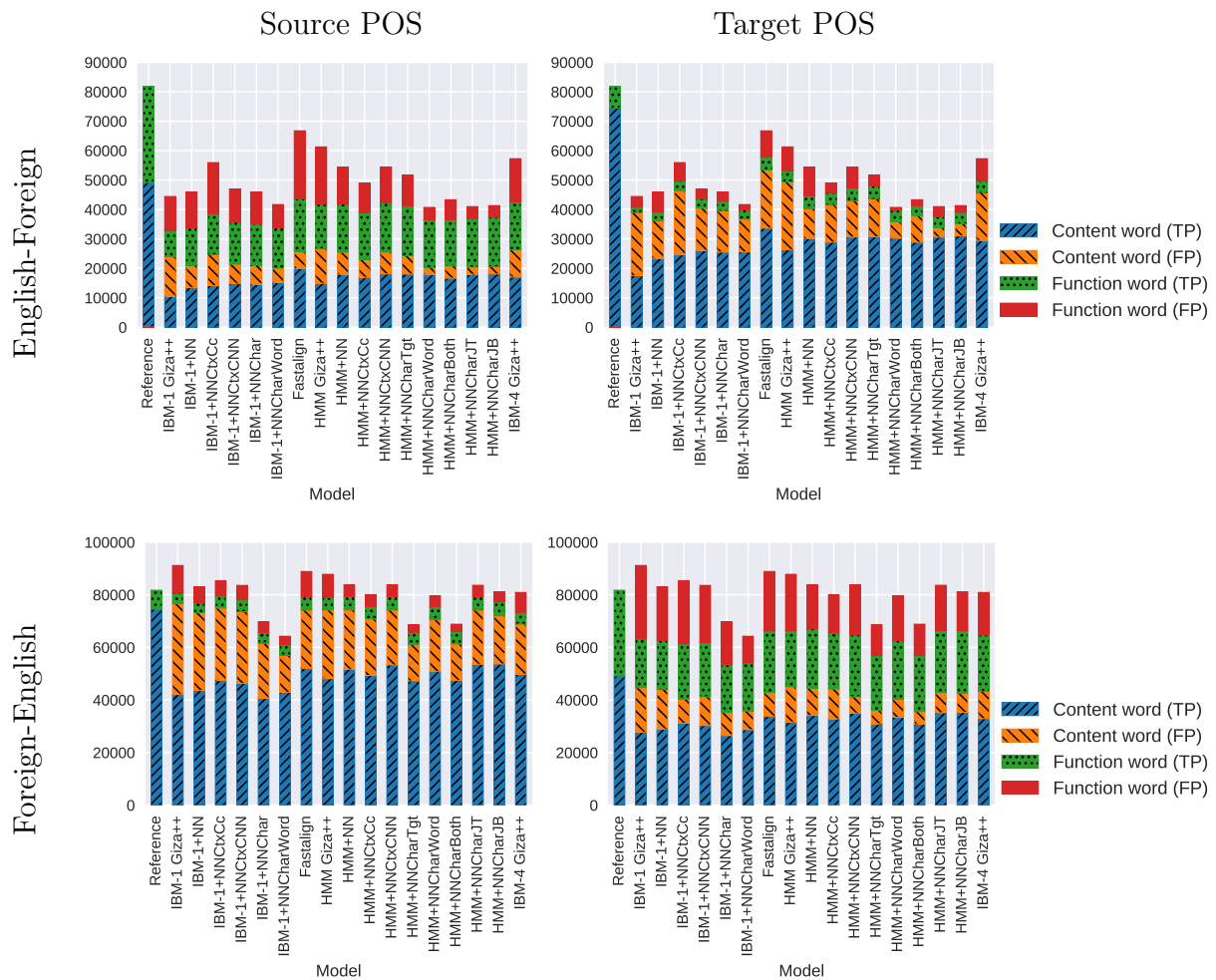


Figure B.42: Neural models for English-Vietnamese: The number of target words that align with a content/function source word (left graph). The number of source words that align with a content/function target words (right graph).

B.8.2 Unaligned words for content and function words



Figure B.43: Neural models for English-French: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).

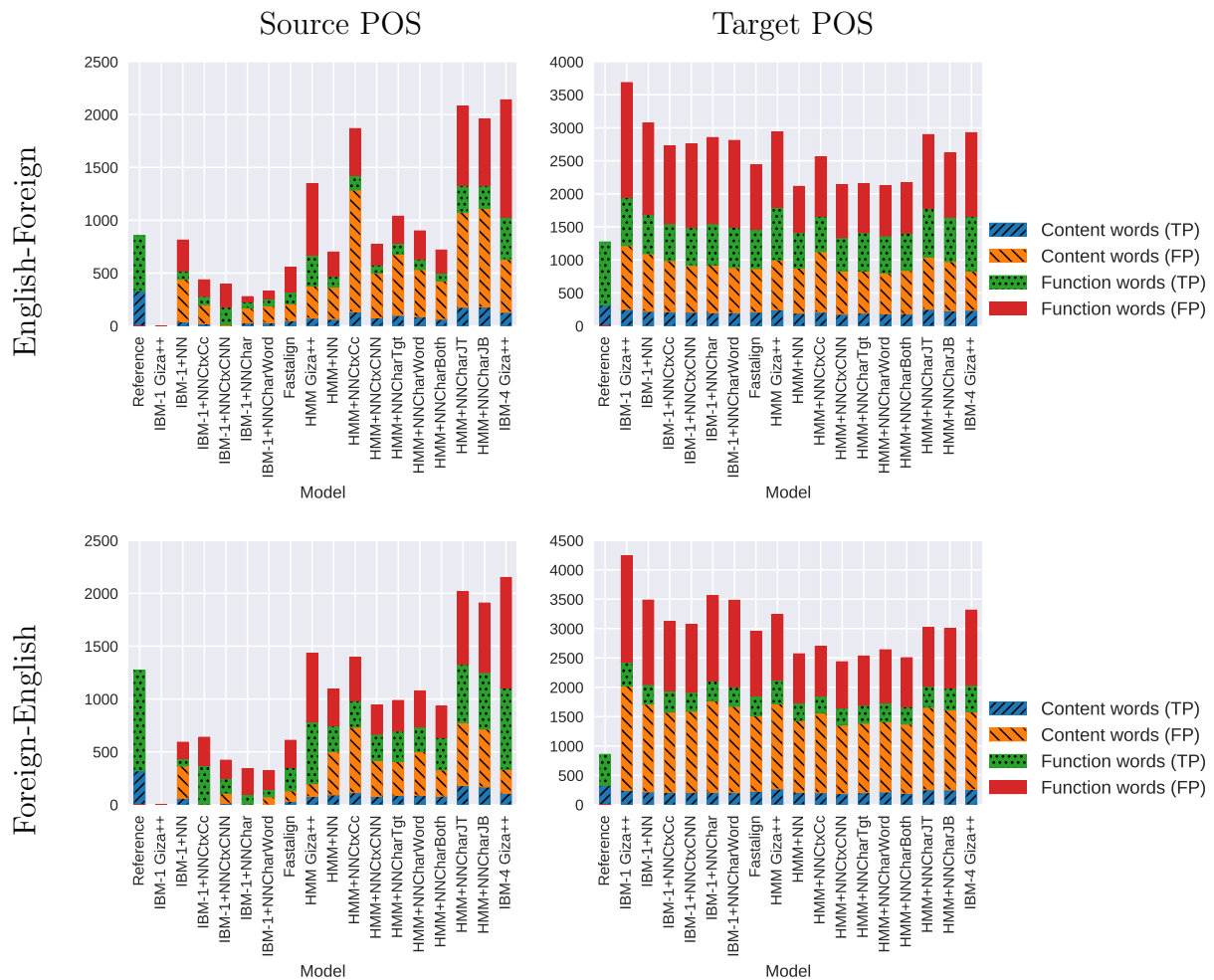


Figure B.44: Neural models for English-German: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).

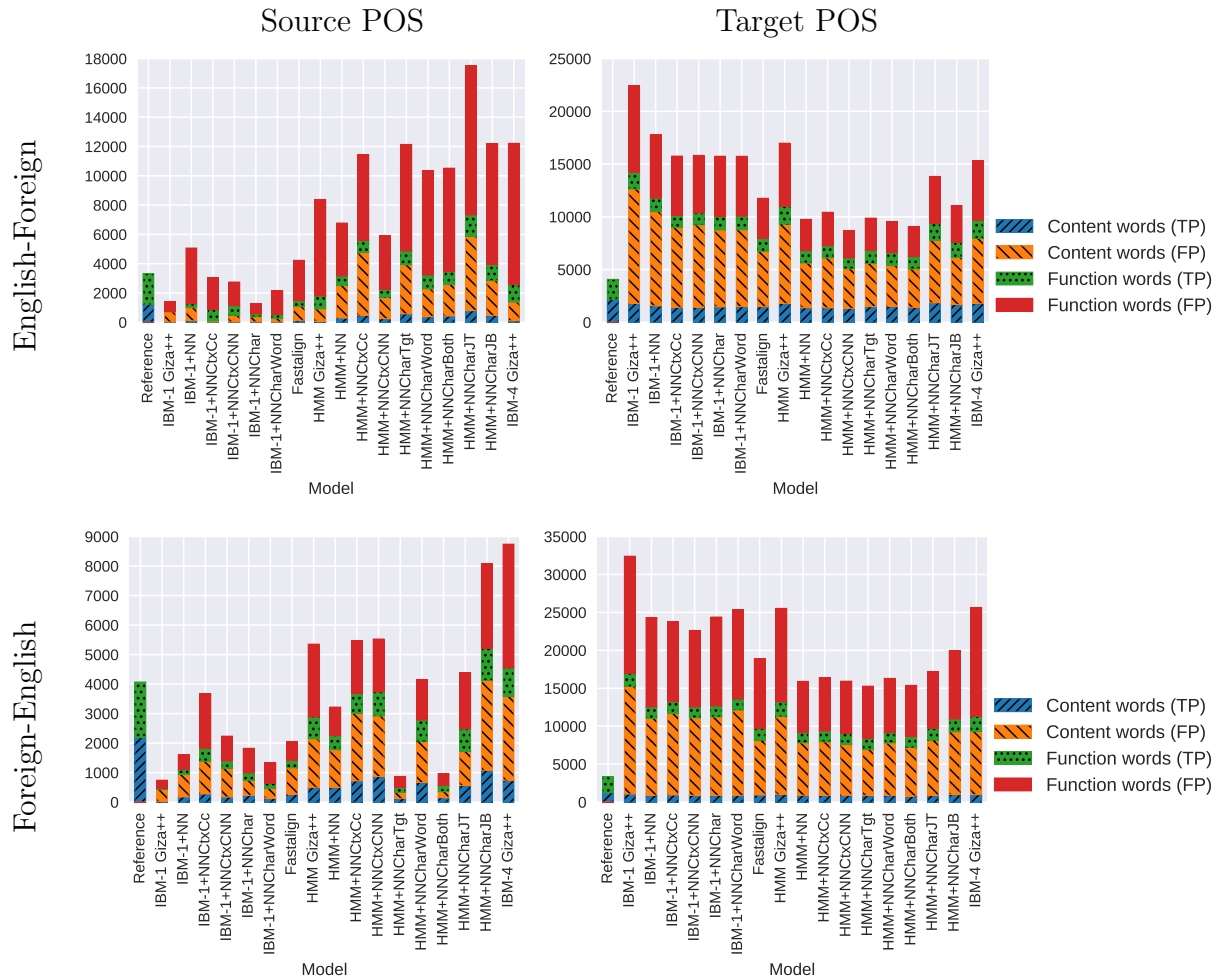


Figure B.45: Neural models for English-Czech: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).

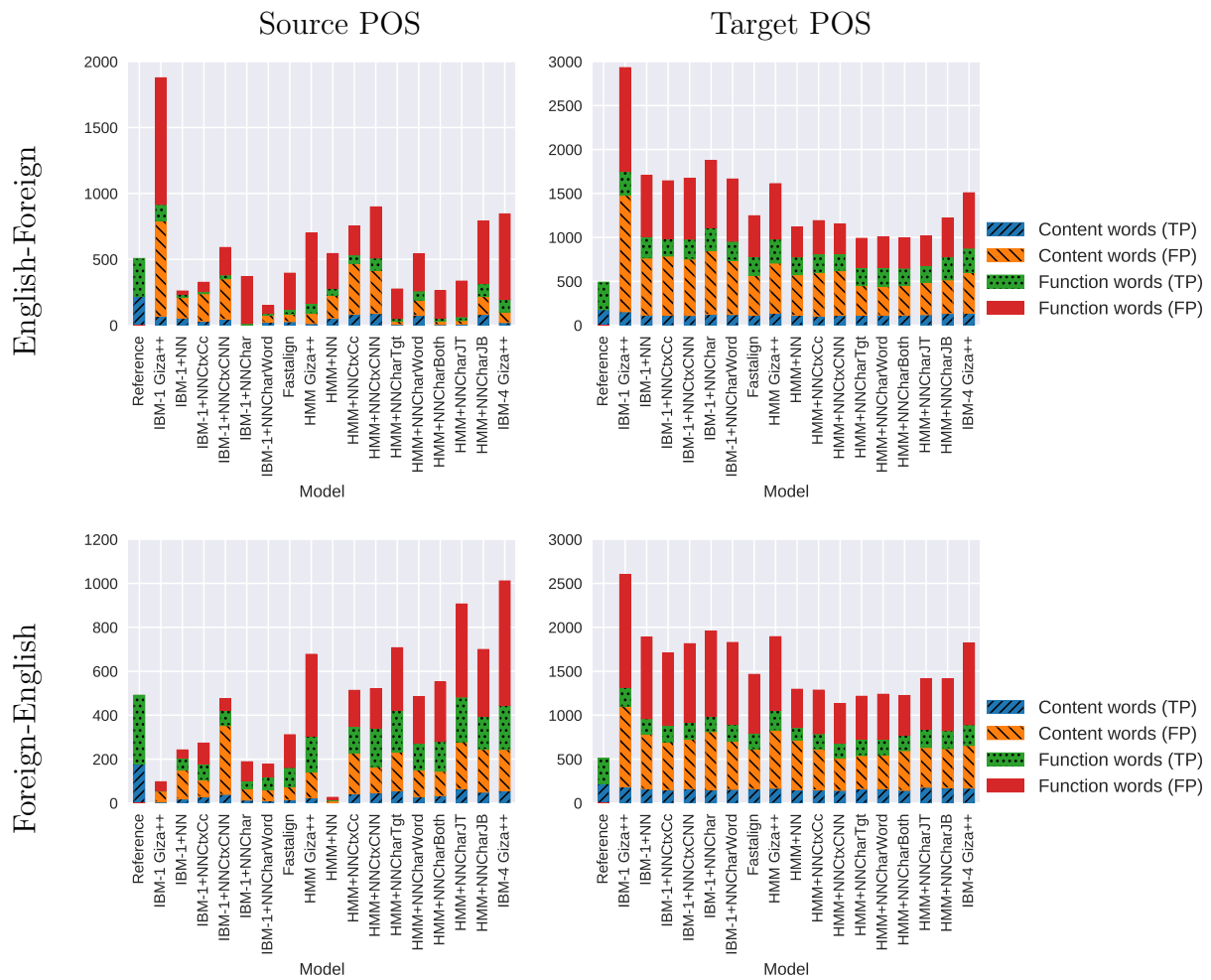


Figure B.46: Neural models for English-Romanian: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).

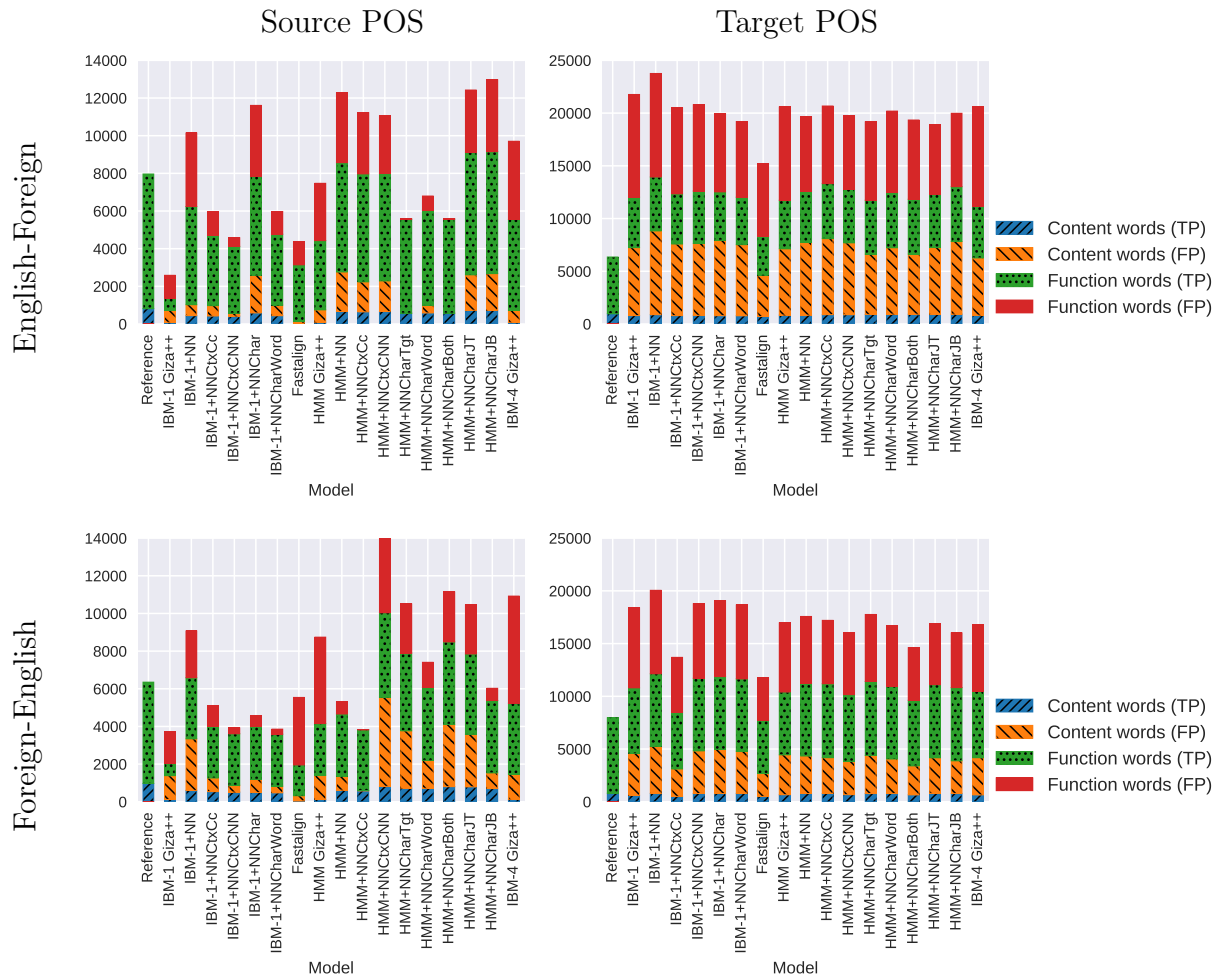


Figure B.47: Neural models for English-Japanese: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).

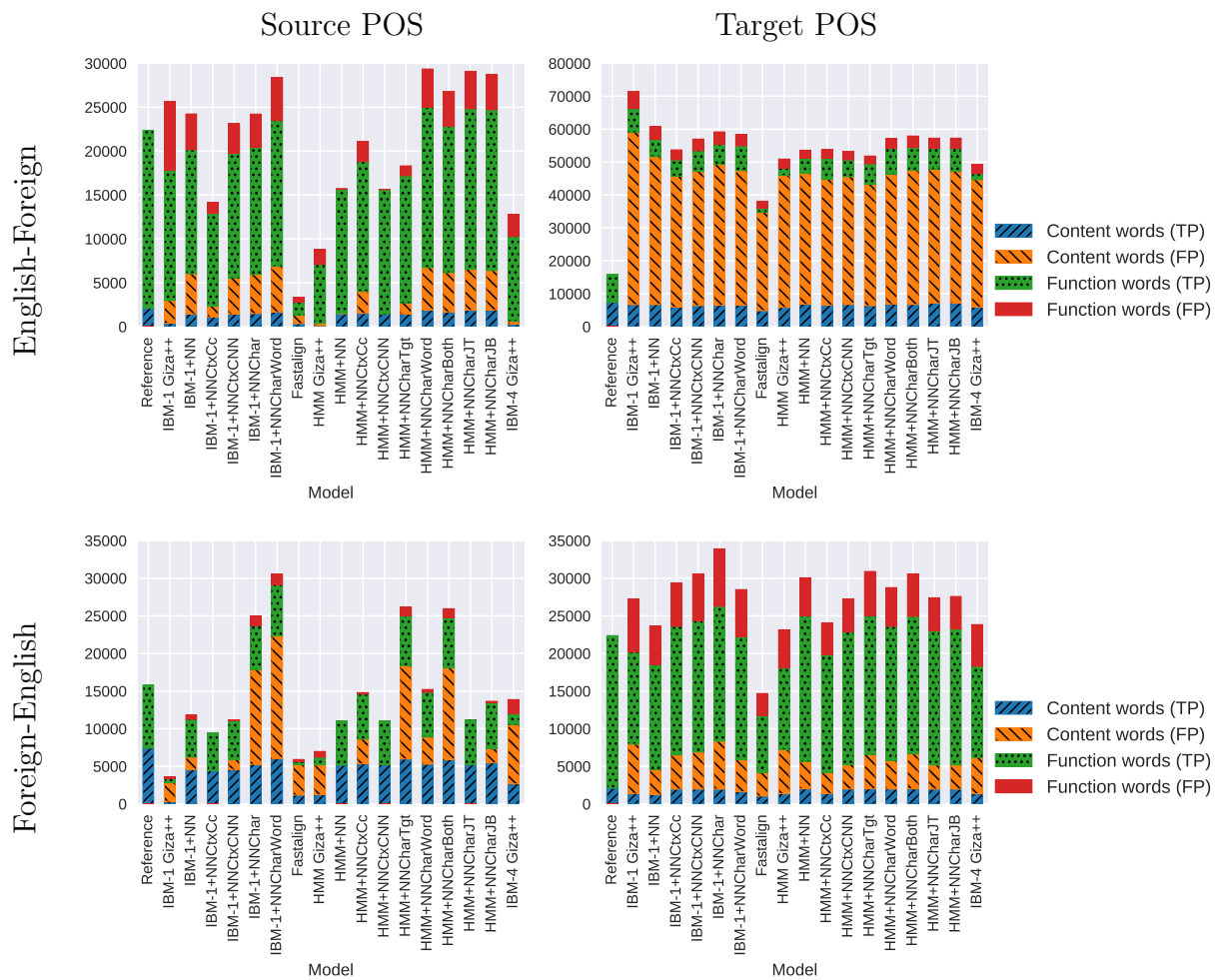


Figure B.48: Neural models for English-Vietnamese: The number of unaligned content/function source words (left graph). The number of unaligned content/function target words (right graph).

B.9 Symmetrization

B.9.1 Intersection

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
IBM-1 Giza++	2441	0.31	0.61	36.9	23.04	93.81	13.13	89.3	0.43	0.55
IBM-1+NN	4023	0.59	0.65	21.46	32.78	87.42	20.17	89.91	0.52	0.36
IBM-1+NNCtxCc	4099	0.55	0.6	19.6	33.8	88.8	20.87	90.03	0.43	0.36
IBM-1+NNCtxCNN	4061	0.56	0.6	19.37	33.82	89.53	20.85	90.05	0.57	0.36
IBM-1+NNChar	4106	0.56	0.61	22.86	32.3	84.73	19.95	89.8	0.26	0.09
IBM-1+NNCharWord	4275	0.59	0.63	20.65	33.94	86.2	21.13	89.97	0.43	0.36
Fastalign	4879	0.69	0.75	11.09	40.48	92.58	25.9	90.71	0.7	0.73
HMM Giza++	4683	0.73	0.76	7.59	41.16	97.2	26.1	90.9	0.65	0.82
HMM+NN	4771	0.73	0.77	7.42	41.53	96.67	26.45	90.92	0.57	0.64
HMM+NNCtxCc	4707	0.73	0.77	7.49	41.13	96.75	26.12	90.88	0.7	0.64
HMM+NNCtxCNN	4654	0.72	0.77	8.78	40.6	96.37	25.72	90.82	0.52	0.64
HMM+NNCharTgt	5049	0.78	0.81	6	43.54	96.95	28.07	91.12	0.43	0.64
HMM+NNCharWord	5071	0.75	0.79	6.55	43.58	96.73	28.13	91.12	0.52	0.45
HMM+NNCharBoth	4982	0.75	0.78	7.04	42.91	96.55	27.58	91.05	0.57	0.55
HMM+NNCharJT	4562	0.83	0.81	6.69	40.81	98.4	25.74	90.89	0.3	0.45
HMM+NNCharJB	4698	0.8	0.85	6.27	41.62	98.06	26.42	90.96	0.39	0.64
IBM-4 Giza++	4588	0.77	0.81	7.76	40.88	98.13	25.82	90.89	0.65	0.64

Table B.38: Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-French

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
IBM-1 Giza++	5531	0.56	0.53	37.11	59.5	86.4	45.37	97.29	0.68	1
IBM-1+NN	5908	0.63	0.62	36.51	60.22	83.78	47	97.28	0.47	0.5
IBM-1+NNCtxCc	5823	0.63	0.58	33.75	62.83	88.24	48.78	97.47	0.58	0.5
IBM-1+NNCtxCNN	5819	0.61	0.58	33.73	62.88	88.35	48.81	97.47	0.42	0.5
IBM-1+NNChar	5851	0.61	0.58	36.16	60.55	84.77	47.09	97.31	0.21	0.5
IBM-1+NNCharWord	6026	0.63	0.6	35.96	60.8	83.54	47.79	97.3	0.47	0.67
Fastalign	6746	0.72	0.68	28.41	68.31	87.49	56.03	97.72	0.84	1
HMM Giza++	6265	0.74	0.69	25.46	70.91	95.07	56.55	97.97	0.84	0.83
HMM+NN	6696	0.76	0.69	27.39	69.43	89.32	56.78	97.81	0.84	0.67
HMM+NNCtxCc	5770	0.67	0.68	32.76	63.91	90.29	49.46	97.55	0.74	0.67
HMM+NNCtxCNN	6464	0.72	0.67	30.09	66.82	87.86	53.92	97.65	0.74	0.83
HMM+NNCharTgt	6809	0.76	0.73	26.2	70.55	89.84	58.07	97.87	0.89	0.67
HMM+NNCharWord	6528	0.74	0.69	27.08	69.54	90.87	56.32	97.84	0.58	0.67
HMM+NNCharBoth	6779	0.75	0.7	27.25	69.59	88.86	57.19	97.81	0.95	0.83
HMM+NNCharJT	6270	0.79	0.75	25.9	70.43	94.37	56.18	97.93	0.84	0.67
HMM+NNCharJB	6441	0.8	0.76	25.58	70.84	93.34	57.08	97.94	0.79	0.67
IBM-4 Giza++	6413	0.82	0.78	23.82	72.55	95.85	58.36	98.06	0.89	0.83

Table B.39: Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-German

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
IBM-1 Giza++	23298	0.45	0.4	40.22	45.12	87.85	30.36	97	0.39	0.55
IBM-1+NN	28673	0.56	0.52	34.94	50.34	84.36	35.88	97.13	0.18	0.27
IBM-1+NNCtxCc	25658	0.52	0.45	36.42	48.54	88.05	33.51	97.12	0.13	0.26
IBM-1+NNCtxCNN	25909	0.51	0.45	36.45	48.6	87.54	33.64	97.11	0.14	0.19
IBM-1+NNChar	27439	0.54	0.47	35.66	49.55	85.65	34.86	97.12	0.12	0.15
IBM-1+NNCharWord	26145	0.51	0.45	38.87	46.93	83.98	32.57	97.01	0.16	0.27
Fastalign	36091	0.71	0.65	20.68	63.06	90.43	48.41	97.7	0.63	0.77
HMM Giza++	28415	0.6	0.55	25.65	57.2	96.46	40.65	97.53	0.51	0.68
HMM+NN	36067	0.73	0.68	20.64	63.4	90.95	48.65	97.72	0.56	0.61
HMM+NNCtxCc	33167	0.7	0.69	22.91	60.89	92.33	45.42	97.63	0.48	0.58
HMM+NNCtxCNN	35350	0.75	0.66	21.48	62.72	91.17	47.8	97.69	0.55	0.6
HMM+NNCharTgt	35571	0.68	0.75	18.44	64.83	93.85	49.52	97.82	0.5	0.53
HMM+NNCharWord	35679	0.73	0.72	18.29	65.01	93.93	49.71	97.83	0.44	0.56
HMM+NNCharBoth	34473	0.66	0.7	20.69	62.92	92.99	47.54	97.72	0.5	0.54
HMM+NNCharJT	31610	0.65	0.75	20.51	61.98	97.08	45.52	97.73	0.4	0.4
HMM+NNCharJB	33178	0.74	0.7	19.61	63.23	95.86	47.17	97.77	0.34	0.48
IBM-4 Giza++	30648	0.69	0.65	21.43	60.84	97.33	44.24	97.69	0.58	0.73

Table B.40: Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Czech

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
IBM-1 Giza++	1811	0.35	0.51	59.32	40.7	87.63	26.5	97.12	0.49	0.69
IBM-1+NN	3074	0.61	0.59	42.95	57.07	84.12	43.19	97.58	0.22	0.54
IBM-1+NNCtxCc	2662	0.53	0.52	46.72	53.29	86.59	38.49	97.48	0.09	0.31
IBM-1+NNCtxCNN	2630	0.54	0.54	46.46	53.56	87.76	38.54	97.51	0.11	0.31
IBM-1+NNChar	2680	0.52	0.53	48.03	51.98	84.07	37.63	97.41	0.18	0.08
IBM-1+NNCharWord	3006	0.58	0.57	43.2	56.82	85	42.67	97.58	0.16	0.31
Fastalign	3659	0.73	0.72	32.85	67.17	88.55	54.11	98.03	0.67	0.92
HMM Giza++	3119	0.67	0.66	35.7	64.32	93.91	48.91	97.98	0.6	0.92
HMM+NN	3247	0.61	0.66	36.11	63.91	90.88	49.28	97.92	0.29	0.23
HMM+NNCtxCc	3226	0.67	0.69	37.07	62.95	89.89	48.43	97.87	0.44	0.38
HMM+NNCtxCNN	3483	0.73	0.76	33.02	67	91.1	52.99	98.05	0.33	0.38
HMM+NNCharTgt	3751	0.81	0.72	28.52	71.51	92.83	58.15	98.27	0.58	0.62
HMM+NNCharWord	3761	0.78	0.77	28.73	71.29	92.4	58.03	98.26	0.53	0.54
HMM+NNCharBoth	3646	0.77	0.7	30.5	69.52	91.85	55.93	98.17	0.6	0.46
HMM+NNCharJT	3605	0.82	0.7	29.24	70.78	94.17	56.7	98.25	0.53	0.69
HMM+NNCharJB	3587	0.78	0.77	29.53	70.5	94.09	56.36	98.24	0.29	0.46
IBM-4 Giza++	3315	0.77	0.72	32.19	67.83	95.17	52.69	98.14	0.69	0.85

Table B.41: Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Romanian

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
IBM-1 Giza++	9540	0.31	0.34	64.7	35.3	79.4	22.7	98.01	0.58	0.57
IBM-1+NN	3921	0.15	0.19	79.82	20.18	96	11.28	97.87	0.18	0.11
IBM-1+NNCtxCc	4903	0.17	0.2	75.45	24.55	95.84	14.08	97.94	0.19	0.11
IBM-1+NNCtxCNN	5497	0.18	0.21	72.77	27.23	96.27	15.86	97.98	0.21	0.11
IBM-1+NNChar	4789	0.16	0.25	75.63	24.37	97.1	13.93	97.94	0.28	0.33
IBM-1+NNCharWord	5188	0.17	0.21	73.87	26.13	97.11	15.09	97.96	0.14	0.08
Fastalign	15622	0.54	0.59	51.23	48.77	76.49	35.8	98.21	0.53	0.67
HMM Giza++	9372	0.37	0.4	61.12	38.88	88.67	24.9	98.13	0.53	0.57
HMM+NN	6579	0.23	0.35	67.51	32.49	98.65	19.44	98.07	0.23	0.33
HMM+NNCtxCc	7369	0.24	0.38	64.29	35.71	98.72	21.8	98.13	0.29	0.15
HMM+NNCtxCNN	6437	0.32	0.33	67.8	32.2	99.58	19.2	98.07	0.12	0.08
HMM+NNCharTgt	8057	0.34	0.32	61.45	38.55	99.13	23.93	98.18	0.36	0.4
HMM+NNCharWord	8587	0.32	0.36	59.42	40.58	99.15	25.51	98.22	0.38	0.48
HMM+NNCharBoth	8598	0.37	0.34	59.3	40.7	99.34	25.59	98.22	0.34	0.37
HMM+NNCharJT	8103	0.34	0.44	61.02	38.98	99.77	24.22	98.19	0.34	0.46
HMM+NNCharJB	8652	0.3	0.48	58.96	41.04	99.69	25.84	98.23	0.31	0.18
IBM-4 Giza++	10137	0.43	0.48	56.29	43.71	93.82	28.5	98.25	0.56	0.63

Table B.42: Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Japanese

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
IBM-1 Giza++	20603	0.23	0.46	67.18	32.82	81.52	20.55	97.26	0.62	0.39
IBM-1+NN	15580	0.19	0.34	68.68	31.33	97.83	18.65	97.33	0.34	0.32
IBM-1+NNCtxCc	20884	0.24	0.37	61.86	38.14	93.71	23.94	97.47	0.57	0.48
IBM-1+NNCtxCNN	20155	0.24	0.43	61.95	38.06	96.2	23.72	97.48	0.52	0.43
IBM-1+NNChar	17882	0.26	0.39	65.23	34.77	96.85	21.19	97.41	0.5	0.4
IBM-1+NNCharWord	16797	0.26	0.4	66.1	33.9	99.43	20.43	97.4	0.28	0.35
Fastalign	48975	0.55	0.73	46.21	53.79	71.78	43.01	97.59	0.78	0.71
HMM Giza++	35598	0.41	0.58	56.09	43.92	72.38	31.52	97.38	0.67	0.47
HMM+NN	24457	0.29	0.45	54.43	45.57	98.94	29.6	97.7	0.42	0.52
HMM+NNCtxCc	22360	0.28	0.46	57.48	42.52	98.98	27.08	97.61	0.41	0.45
HMM+NNCtxCNN	26282	0.31	0.48	52.46	47.54	97.7	31.42	97.74	0.61	0.59
HMM+NNCharTgt	24144	0.35	0.47	55.2	44.81	98.24	29.02	97.67	0.55	0.52
HMM+NNCharWord	23551	0.3	0.58	55.42	44.59	99.66	28.72	97.67	0.45	0.5
HMM+NNCharBoth	22361	0.32	0.52	57.51	42.5	98.92	27.06	97.61	0.56	0.53
HMM+NNCharJT	25612	0.31	0.63	52.53	47.48	99.5	31.18	97.75	0.56	0.56
HMM+NNCharJB	25998	0.32	0.63	52.04	47.96	99.38	31.61	97.76	0.56	0.58
IBM-4 Giza++	38925	0.48	0.68	51.18	48.82	75.67	36.04	97.54	0.72	0.55

Table B.43: Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Vietnamese

B.9.2 Grow-diag-final (GDF)

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
IBM-1 Giza++	40.1	26.7	33.9	36.49	25.19	33.83	82.75	21.26
IBM-1+NN	27.96	36.42	27.21	38.08	17.86	39.48	82.89	25.91
IBM-1+NNCtxCc	27.42	38.66	28.86	39.13	17.48	42.07	81.07	28.41
IBM-1+NNCtxCNN	27.86	38.2	27.99	39.07	17.58	41.65	82.14	27.9
IBM-1+NNChar	28.76	37.5	31.4	37.21	19.62	40.07	79.39	26.8
IBM-1+NNCharWord	27.03	38.69	28.33	38.97	17.88	41.83	80.81	28.21
Fastalign	15.19	44.98	16.23	46.32	14.25	49.05	81.84	35.02
HMM Giza++	11.99	45.18	11.97	45.98	8.54	47.5	90.78	32.17
HMM+NN	11.84	45.57	11.15	46.86	9.8	48.07	88.85	32.95
HMM+NNCtxCc	10.39	45.83	11.57	46.6	10.28	48.16	88.18	33.12
HMM+NNCtxCNN	11.64	45.16	12.71	46.42	11.25	47.97	87.31	33.07
HMM+NNCharTgt	9.17	47.22	9.56	47.87	8.41	48.99	90.12	33.64
HMM+NNCharWord	10.45	47.33	10.27	48.56	9.33	49.64	88.61	34.48
HMM+NNCharBoth	10.9	46.74	11.17	47.5	10.51	48.51	87.16	33.6
HMM+NNCharJT	8.41	44.71	7.7	44.45	6.26	45.39	94.55	29.87
HMM+NNCharJB	8.47	44.38	7.74	46.26	6.81	45.83	93.32	30.38
IBM-4 Giza++	10	44.43	9.64	45.43	7.03	46.32	93.55	30.78

Table B.44: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-French

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
IBM-1 Giza++	39.03	58.76	42.66	55.39	32.31	65.06	72.19	59.2
IBM-1+NN	37.64	60.07	39.22	58.53	30.82	66.36	76.51	58.59
IBM-1+NNCtxCc	36.4	61.36	36.31	61.36	27.88	69.18	79.7	61.11
IBM-1+NNCtxCNN	37.16	60.67	36.02	61.68	27.81	69.37	79.3	61.64
IBM-1+NNChar	36.22	61.55	40.88	56.99	30.29	66.89	76.44	59.47
IBM-1+NNCharWord	35.31	62.41	40.48	57.44	29.88	67.38	75.96	60.54
Fastalign	28.98	68.75	31.28	66.47	28.21	69.84	70.12	69.57
HMM Giza++	23.92	73.3	26.33	71.04	22.87	74.48	81.74	68.39
HMM+NN	26.78	70.95	29.44	68.21	26.92	71	73.93	68.29
HMM+NNCtxCc	34.36	63.19	32.31	65.34	31.72	65.94	73.33	59.9
HMM+NNCtxCNN	30.64	67.18	30.34	67.41	29.38	68.56	71.13	66.18
HMM+NNCharTgt	26.04	71.57	28.11	69.48	26.26	71.54	75.18	68.23
HMM+NNCharWord	24.98	72.64	29.77	67.76	26.35	71.43	74.89	68.28
HMM+NNCharBoth	27.14	70.6	29.31	68.34	27.86	70.04	72.51	67.73
HMM+NNCharJT	23.79	73.15	25.21	71.85	21.76	75.27	86.26	66.76
HMM+NNCharJB	23.69	73.38	24.9	72.16	22.42	74.71	84.83	66.74
IBM-4 Giza++	21.46	75.48	23.31	73.63	19.59	77.29	89.26	68.16

Table B.45: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-German

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
IBM-1 Giza++	45.09	46.75	48.47	42.88	35.47	52.67	71.16	41.81
IBM-1+NN	42.29	48.64	40.97	49.08	30.89	56.25	75.65	44.77
IBM-1+NNCtxCc	44.63	46.86	42.33	48.22	32.33	54.94	75.91	43.04
IBM-1+NNCtxCNN	45.47	46.13	41.92	48.4	31.6	55.28	76.18	43.37
IBM-1+NNChar	40.85	50.35	42.35	47.99	29.85	57.16	77.02	45.44
IBM-1+NNCharWord	40.28	50.82	46.2	44.94	31.44	55.84	75.91	44.16
Fastalign	25.75	64.09	25.3	62.86	23.3	66.93	72.95	61.82
HMM Giza++	27.86	61.22	30.38	57.28	25.25	62.96	75.67	53.91
HMM+NN	23.5	65.45	24.06	64.03	22.18	67.27	75.3	60.78
HMM+NNCtxCc	24.31	63.74	25.02	62.66	22.52	65.83	77.35	57.3
HMM+NNCtxCNN	24.18	65.38	22.92	64.63	22.14	67.56	75.24	61.3
HMM+NNCharTgt	16.74	69.36	24.61	63.94	18.87	69.16	79.35	61.28
HMM+NNCharWord	16.04	70.34	22.8	64.94	17.79	69.96	80.68	61.75
HMM+NNCharBoth	17.38	69.09	28.26	61.04	21.62	67.17	76	60.18
HMM+NNCharJT	15.94	68.31	25.82	62.11	18.29	68.04	83.66	57.33
HMM+NNCharJB	16.38	69.35	23.88	62.72	16.18	69.5	86.22	58.21
IBM-4 Giza++	20.92	65.7	26.5	59.81	19.13	66.67	84.22	55.17

Table B.46: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Czech

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
IBM-1 Giza++	56.02	43.99	53.52	46.49	51.13	48.88	73.82	36.54
IBM-1+NN	46.4	53.62	44.9	55.11	38.38	61.64	75.8	51.94
IBM-1+NNCtxCc	49.93	50.09	43.95	56.07	38.52	61.5	77.04	51.17
IBM-1+NNCtxCNN	46.15	53.87	43.93	56.09	38.13	61.88	79.03	50.85
IBM-1+NNChar	50.16	49.85	48.28	51.73	41.42	58.6	75.25	47.98
IBM-1+NNCharWord	46.54	53.47	43.94	56.08	37.06	62.95	76.62	53.42
Fastalign	33.36	66.65	32.91	67.1	30.42	69.6	72.65	66.8
HMM Giza++	33.36	66.65	36.38	63.64	32.4	67.62	76.63	60.5
HMM+NN	30.69	69.33	40.12	59.89	34.36	65.66	65.89	65.43
HMM+NNCtxCc	33.84	66.18	34.83	65.19	31.33	68.69	74.21	63.93
HMM+NNCtxCNN	30.87	69.15	31.03	68.99	28.86	71.16	77.34	65.9
HMM+NNCharTgt	27.54	72.47	28.01	72.01	25.97	74.05	79.04	69.66
HMM+NNCharWord	25.51	74.51	29.19	70.83	25.91	74.11	78.97	69.81
HMM+NNCharBoth	28.14	71.88	31.53	68.49	28.37	71.65	75.18	68.44
HMM+NNCharJT	28.27	71.75	29.02	71	26.27	73.75	80.42	68.1
HMM+NNCharJB	26.85	73.17	29.76	70.26	26.13	73.89	83.22	66.43
IBM-4 Giza++	31.04	68.98	32.3	67.72	29.08	70.94	82.98	61.96

Table B.47: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Romanian

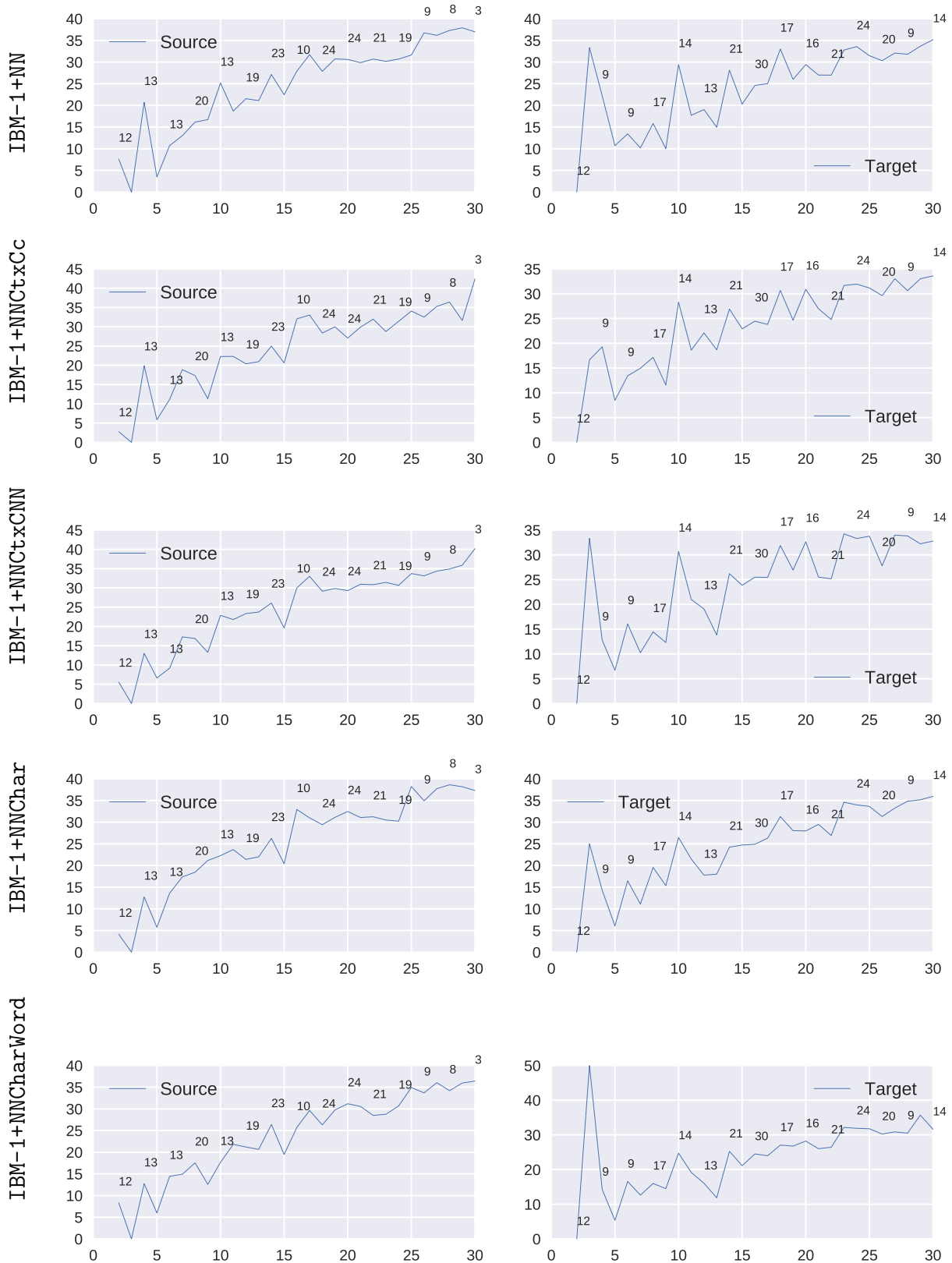
Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
IBM-1 Giza++	63.12	36.88	61.55	38.45	55.31	44.69	64	34.33
IBM-1+NN	62.64	37.36	56.9	43.1	48.3	51.7	47.99	56.02
IBM-1+NNCtxCc	61.62	38.38	55.69	44.31	49.94	50.06	44.35	57.46
IBM-1+NNCtxCNN	60.21	39.79	54.01	45.99	47.63	52.37	45.18	62.28
IBM-1+NNChar	57.96	42.04	54.91	45.09	45.16	54.84	50.4	60.16
IBM-1+NNCharWord	60.92	39.08	53.54	46.46	47.68	52.32	45.64	61.29
Fastalign	50.67	49.33	49.37	50.63	44.79	55.21	64.99	47.99
HMM Giza++	57.01	42.99	54.41	45.59	52.18	47.82	62.37	38.78
HMM+NN	49.68	50.32	47.09	52.91	35.97	64.03	60.63	67.84
HMM+NNCtxCc	47.93	52.07	45.67	54.33	34.87	65.13	60.58	70.43
HMM+NNCtxCNN	46.82	53.18	44.79	55.21	32.06	67.94	70.09	65.93
HMM+NNCharTgt	43.92	56.08	43.6	56.4	31.55	68.45	64.21	73.29
HMM+NNCharWord	45.35	54.65	40.16	59.84	30.71	69.29	64.1	75.4
HMM+NNCharBoth	43.71	56.29	40.26	59.74	30.54	69.46	66	73.31
HMM+NNCharJT	39.69	60.31	39.22	60.78	24.92	75.08	76.56	73.66
HMM+NNCharJB	40.92	59.08	37.71	62.29	25.52	74.48	72.96	76.07
IBM-4 Giza++	52.52	47.48	49.23	50.77	46.68	53.32	74.58	41.49

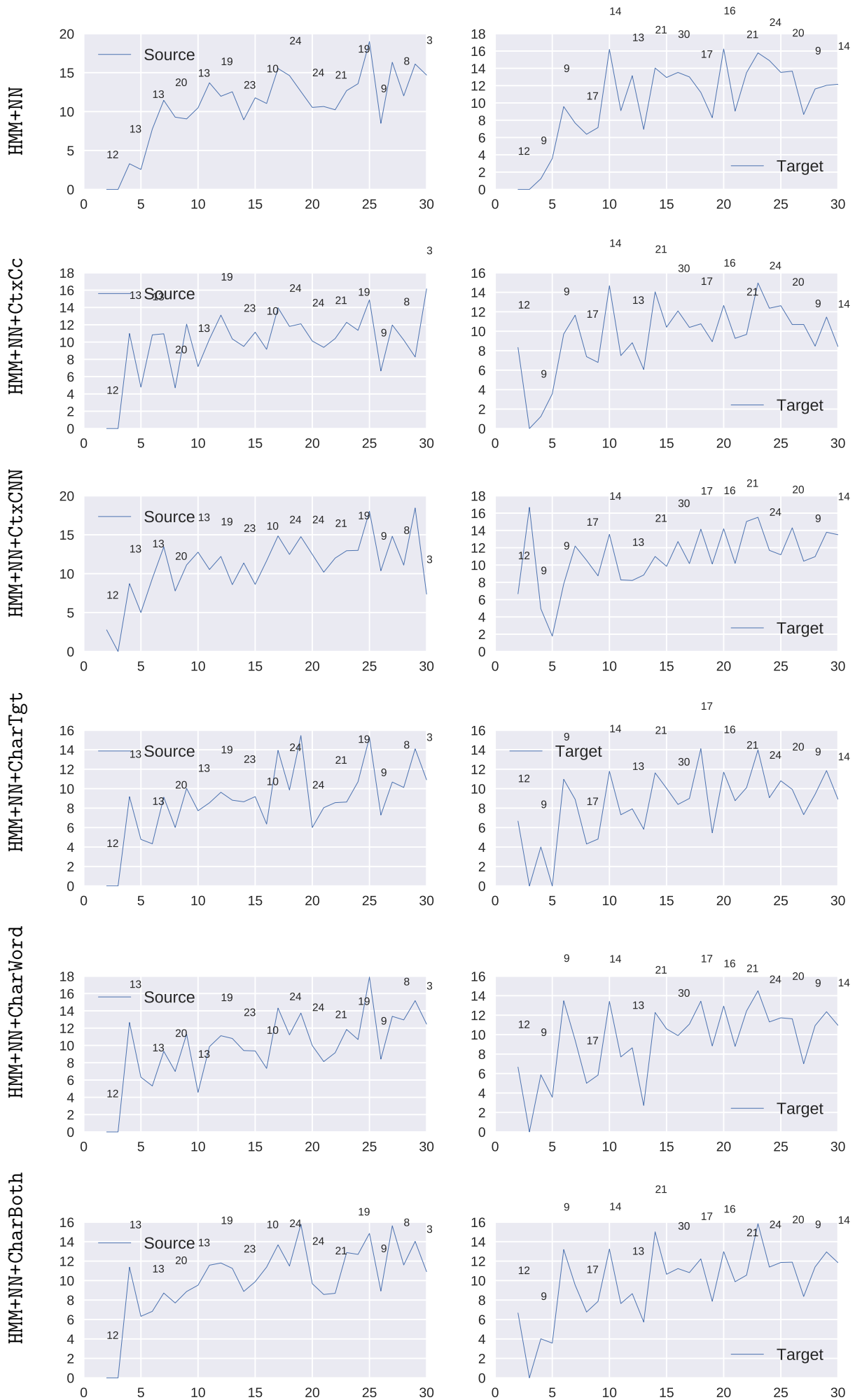
Table B.48: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Japanese

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
IBM-1 Giza++	69.43	30.58	46.45	53.56	47.91	52.09	70.78	41.21
IBM-1+NN	58.87	41.13	42.25	57.75	47.3	52.7	88.34	37.56
IBM-1+NNCtxCc	59.4	40.6	37.5	62.5	39.85	60.16	82.85	47.22
IBM-1+NNCtxCNN	54.55	45.45	38.45	61.56	39.41	60.6	85.33	46.98
IBM-1+NNChar	55.06	44.94	41.15	58.86	45.52	54.49	86.59	39.75
IBM-1+NNCharWord	53.2	46.8	35.45	64.55	43.61	56.4	94.49	40.2
Fastalign	48.89	51.11	32.82	67.19	32.9	67.11	64.41	70.05
HMM Giza++	57.86	42.14	37.57	62.43	39.79	60.21	61.51	58.97
HMM+NN	49.27	50.73	31.45	68.56	28.74	71.27	90.09	58.95
HMM+NNCtxCc	49.56	50.44	33	67	32.42	67.58	91.04	53.74
HMM+NNCtxCNN	48.62	51.39	29.23	70.78	26.93	73.08	89.45	61.77
HMM+NNCharTgt	47.52	52.49	30.94	69.06	31.98	68.03	91.09	54.29
HMM+NNCharWord	44.37	55.64	30.92	69.09	29.12	70.89	93.33	57.15
HMM+NNCharBoth	48.1	51.91	30.72	69.28	33.69	66.32	92.68	51.63
HMM+NNCharJT	43.78	56.22	28.86	71.15	25.79	74.21	93.05	61.72
HMM+NNCharJB	43.28	56.73	27.59	72.42	25.32	74.69	93.47	62.19
IBM-4 Giza++	51.91	48.1	33.19	66.82	33.29	66.71	70.25	63.51

Table B.49: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Vietnamese

B.10 AER as a function of sentence length





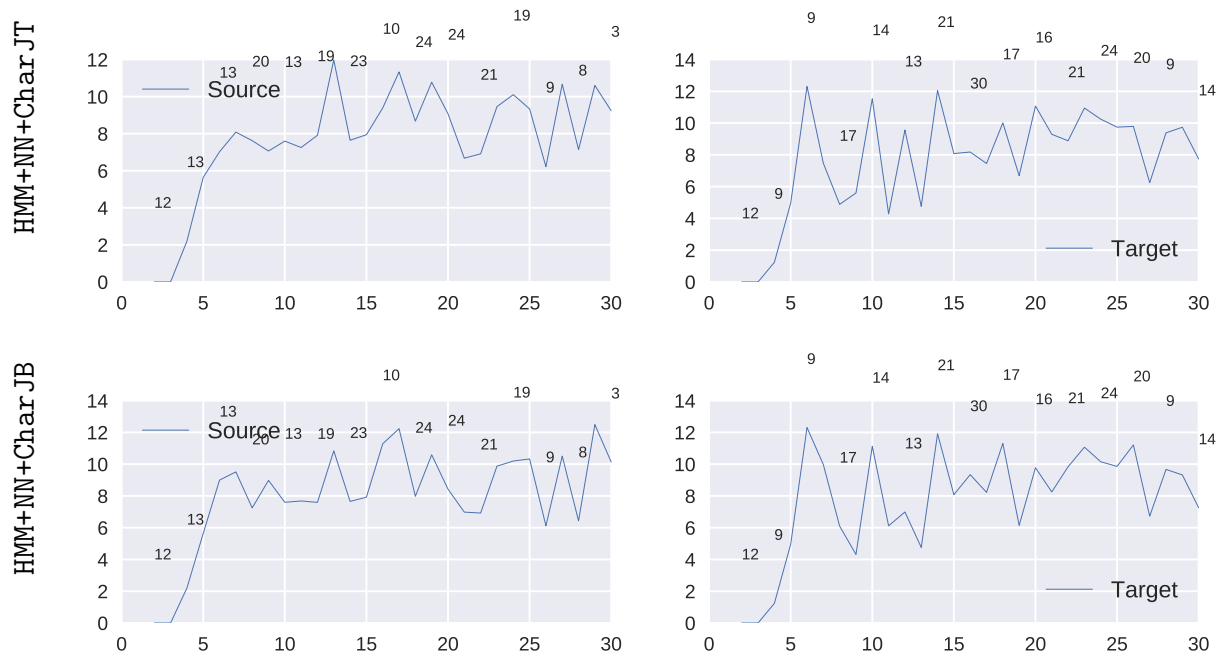
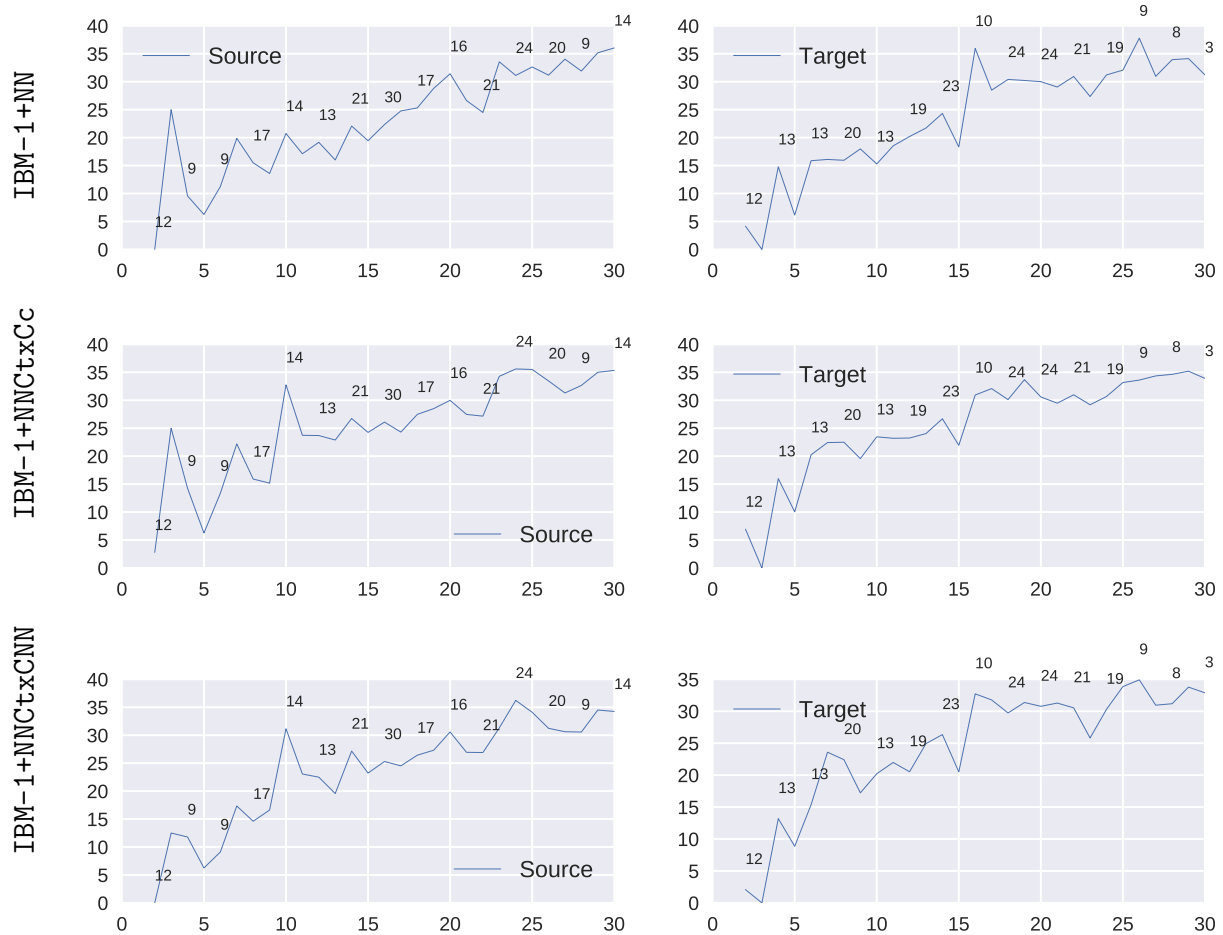
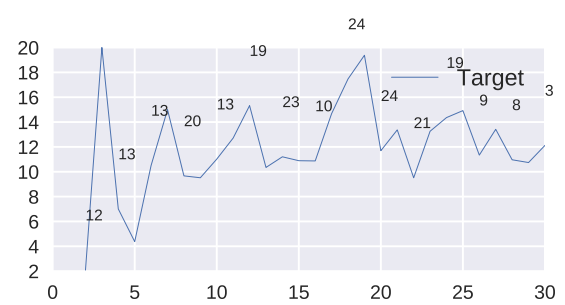
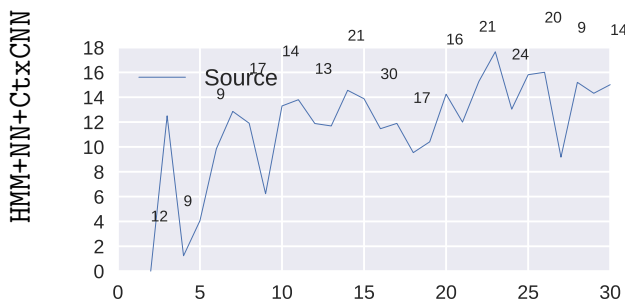
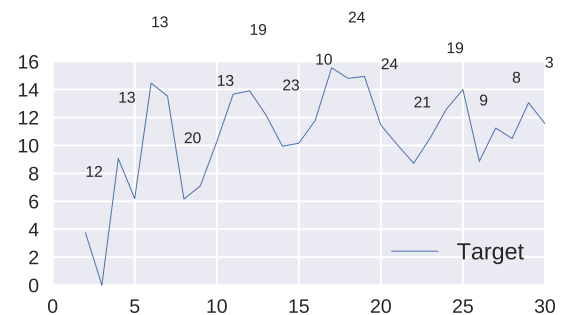
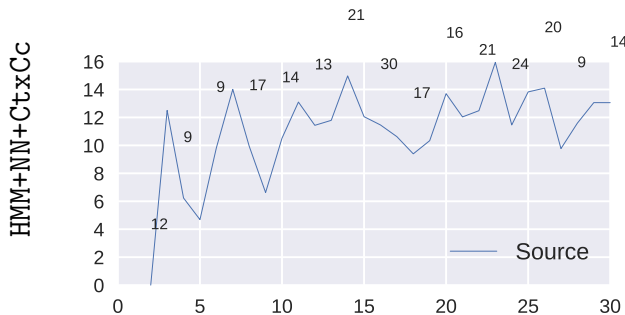
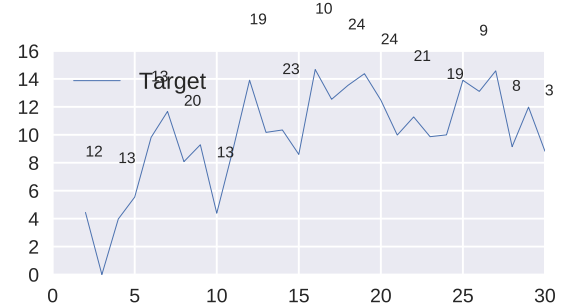
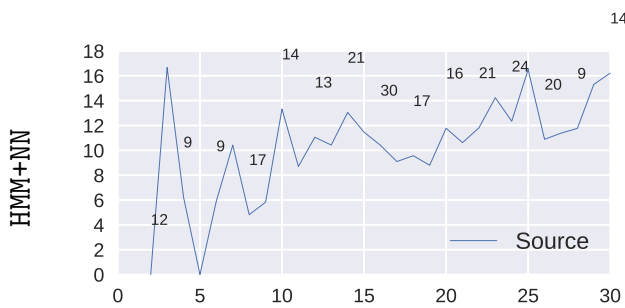
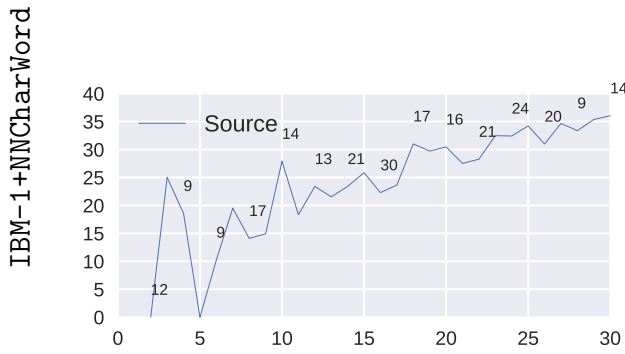
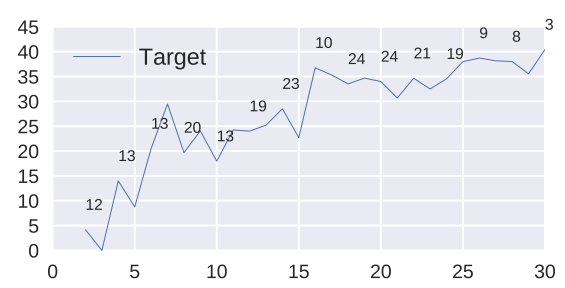
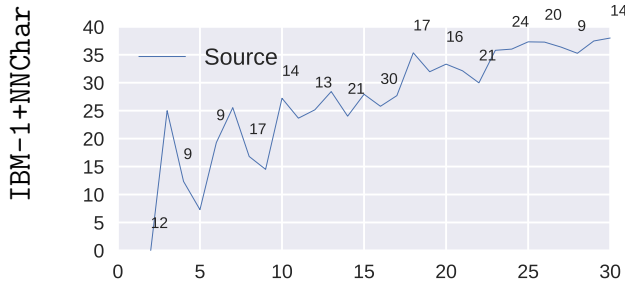


Figure B.49: Neural models for the direction English-French: AER score as a function of sentence length





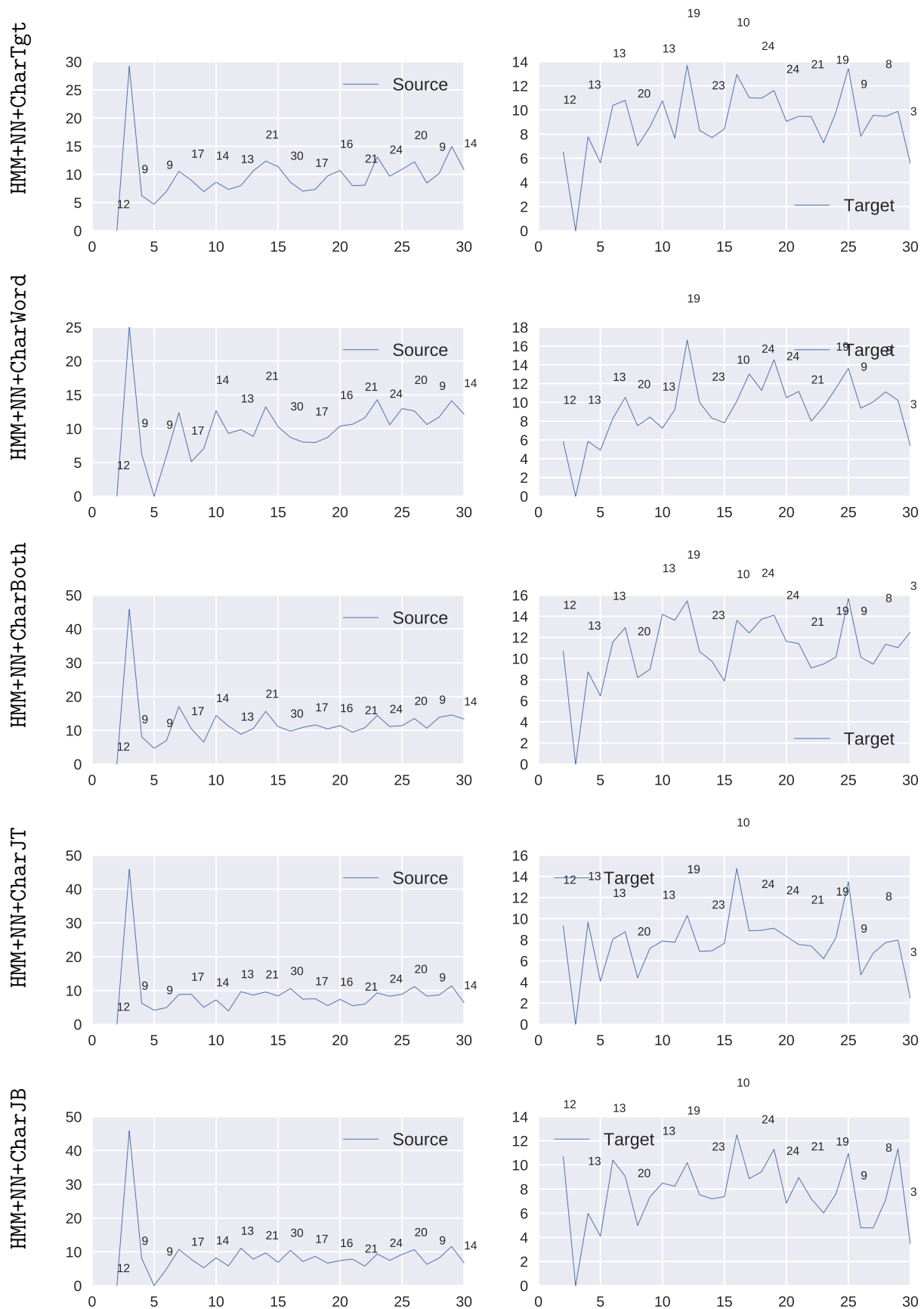
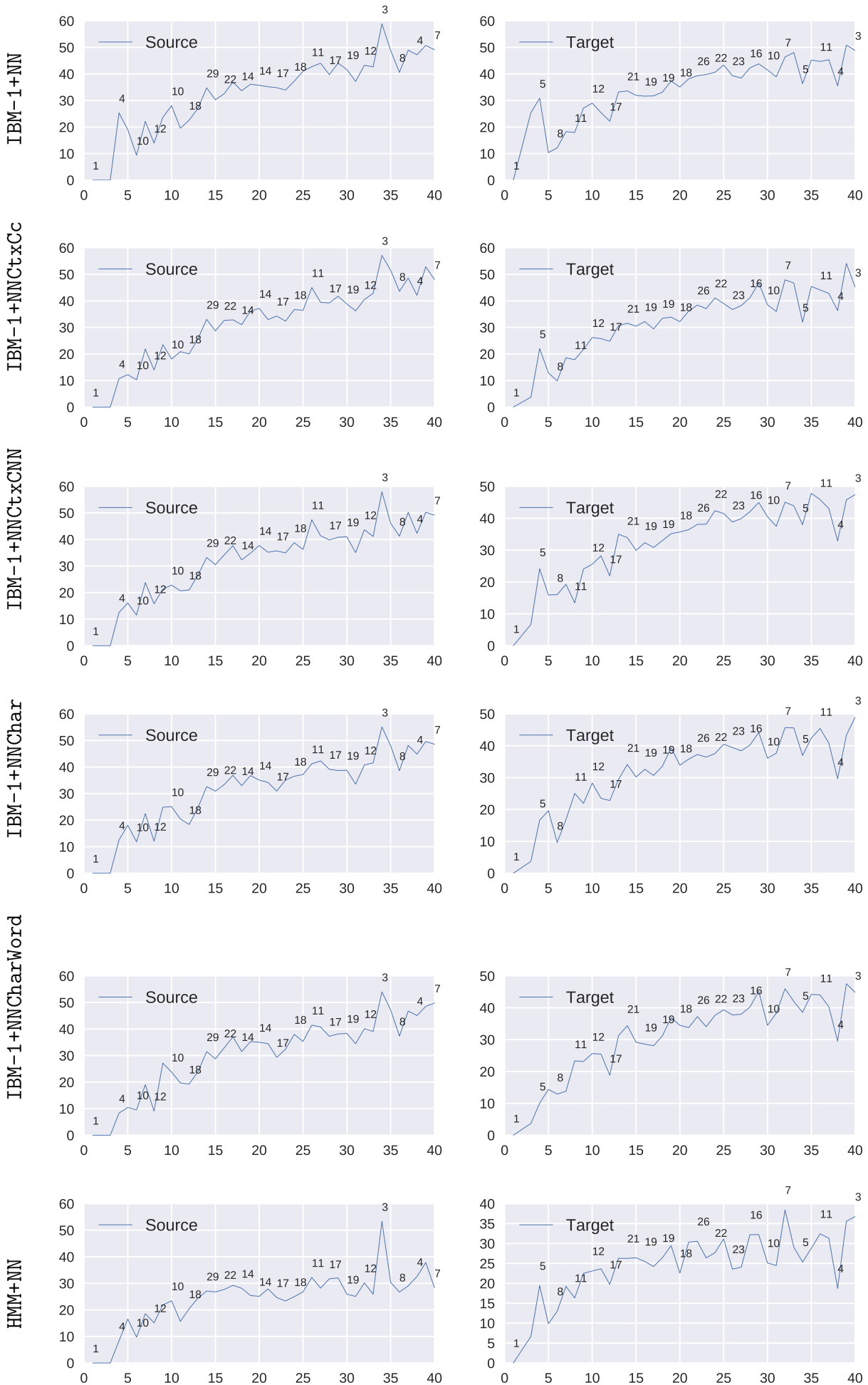
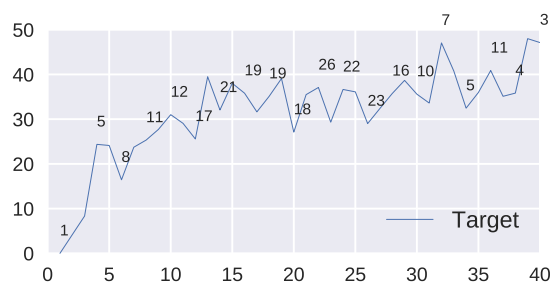
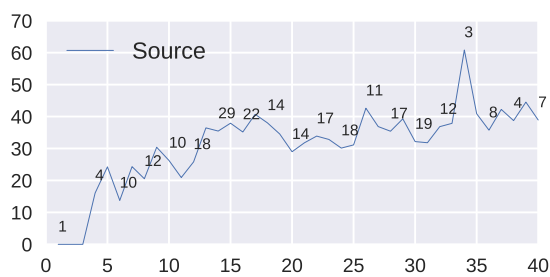


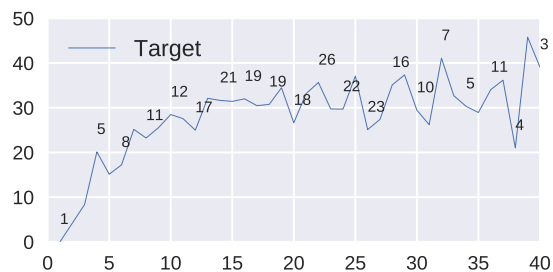
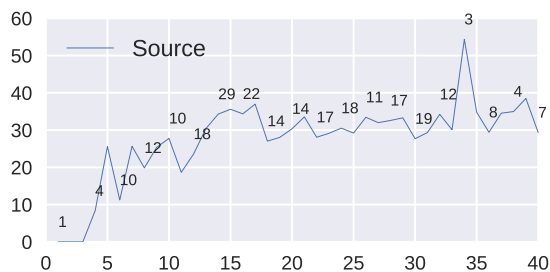
Figure B.50: Neural models for the direction French-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



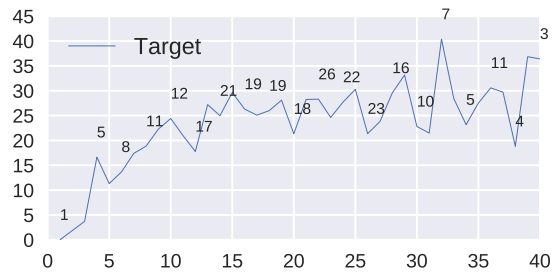
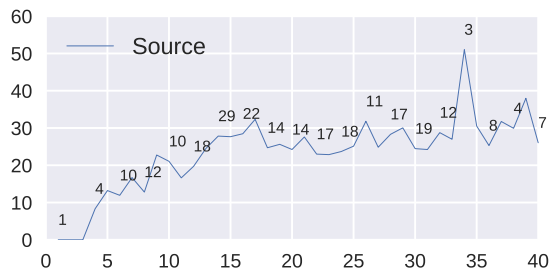
HMM+NN+CtxCc



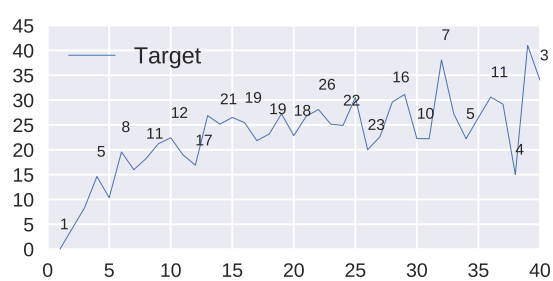
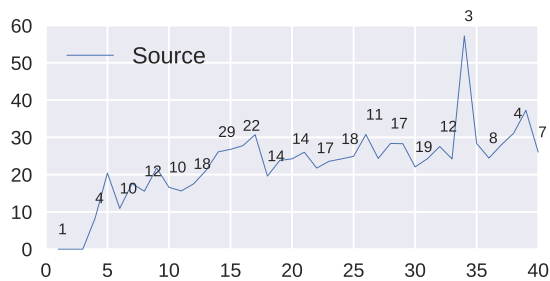
HMM+NN+CtxCNN



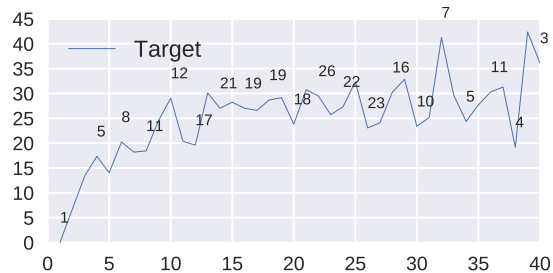
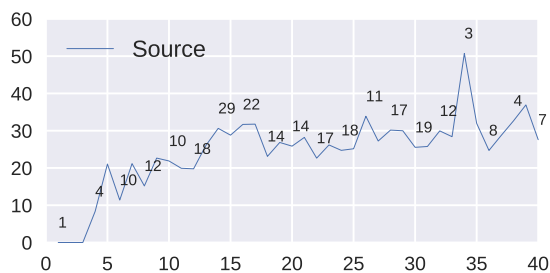
HMM+NN+CharTgt



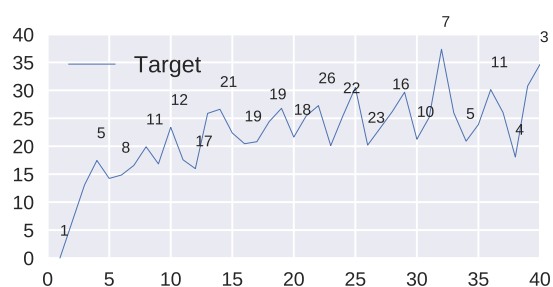
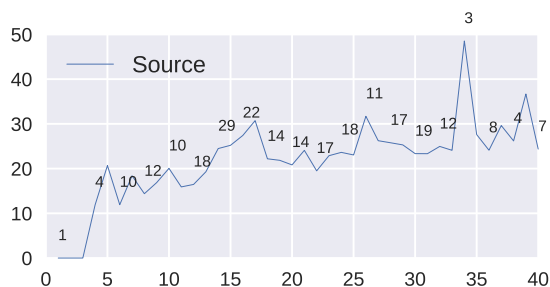
HMM+NN+CharWord



HMM+NN+CharBoth



HMM+NN+CharJT



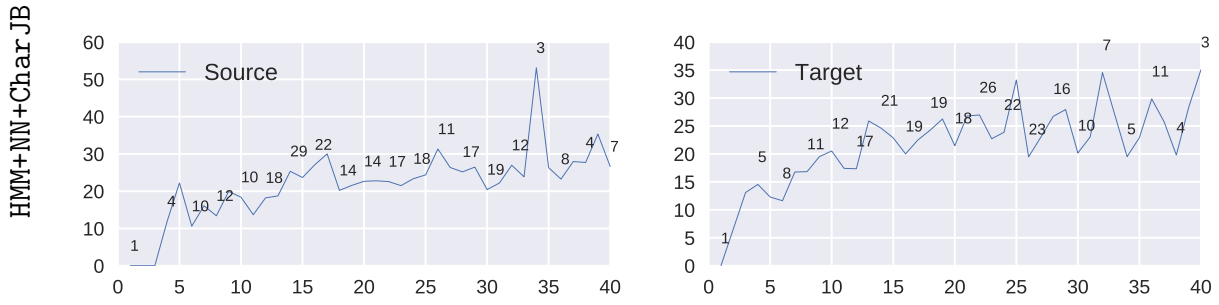
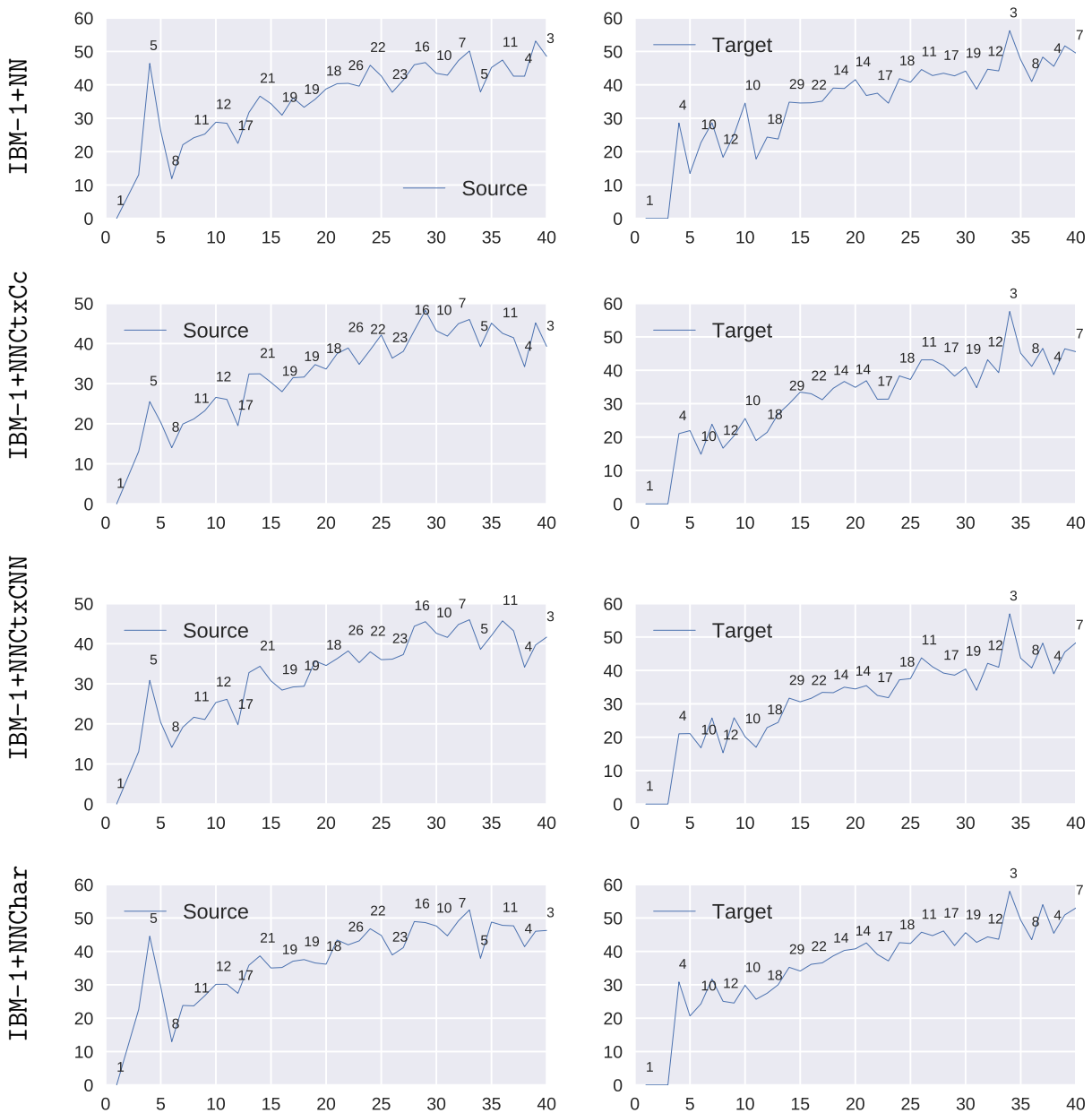
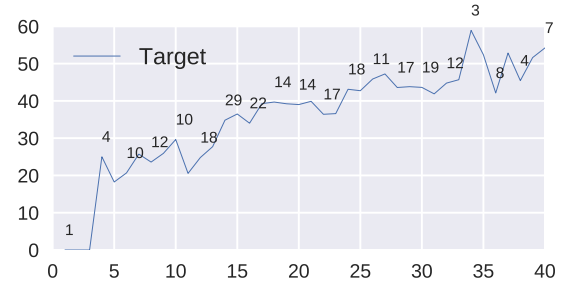
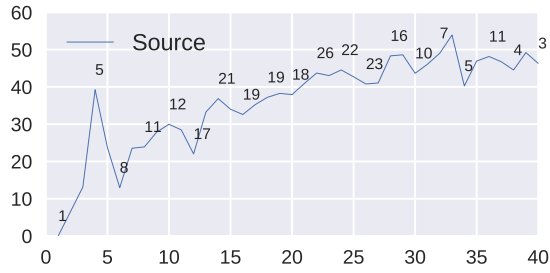


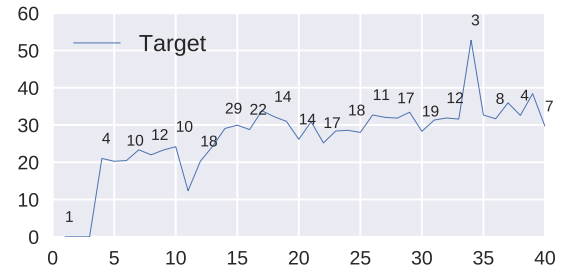
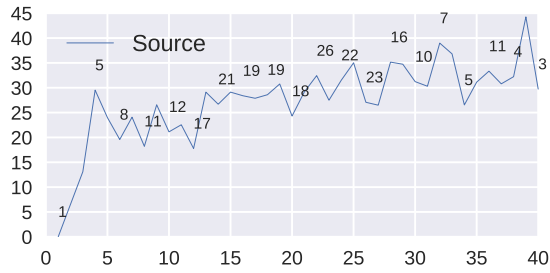
Figure B.51: Neural models for the direction English-German: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



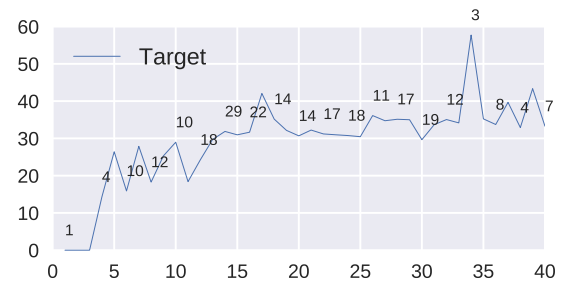
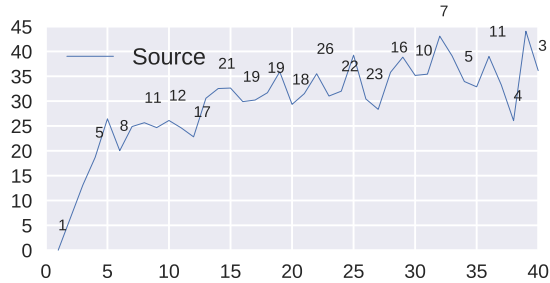
IBM-1+NNCharWord



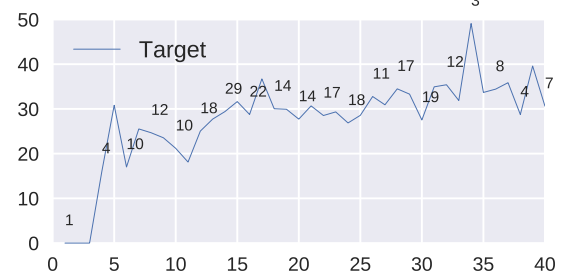
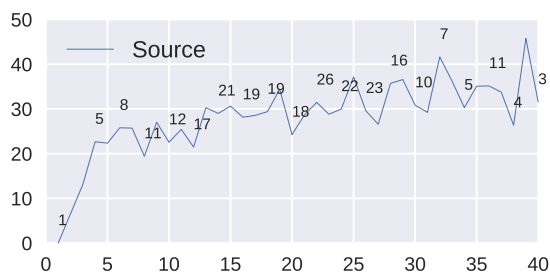
HMM+NN



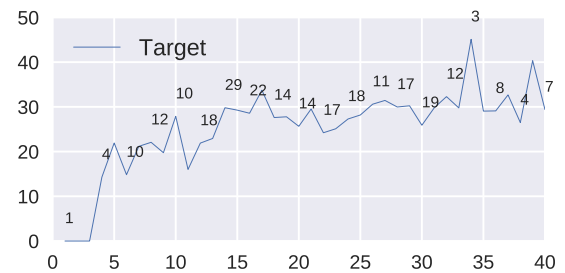
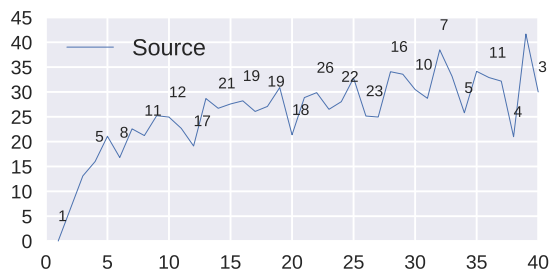
HMM+NN+CtxCc



HMM+NN+CtxCNN



HMM+NN+CharTgt



HMM+NN+CharWord

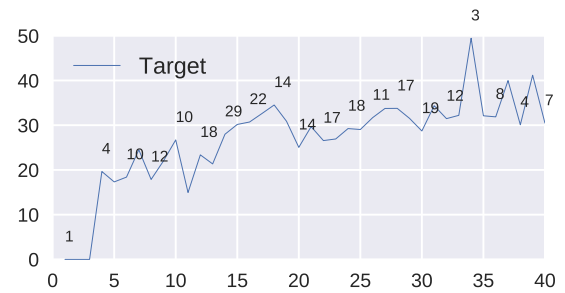
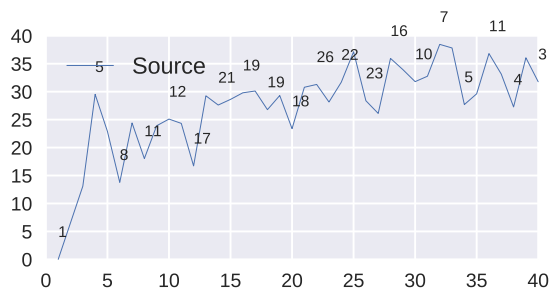
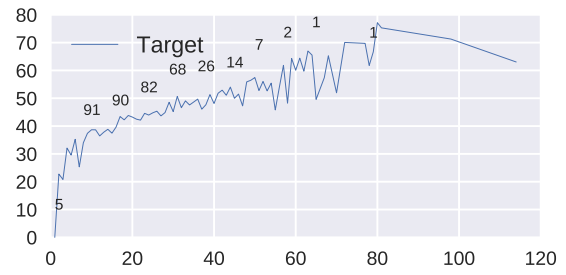
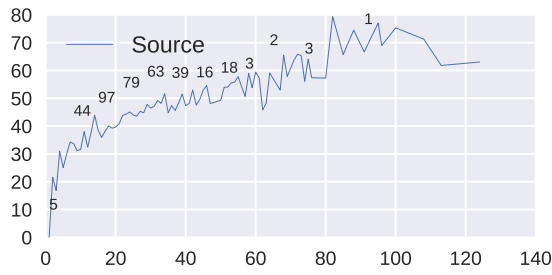




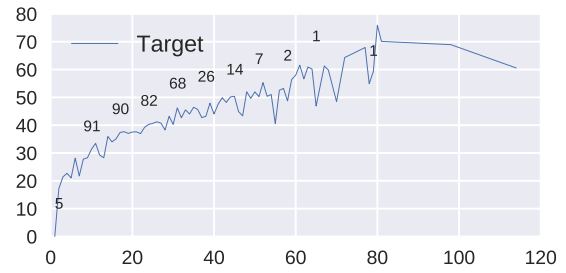
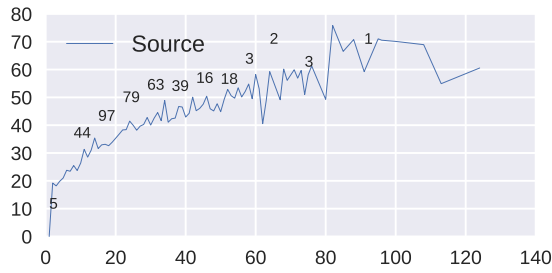
Figure B.52: Neural models for the direction German-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



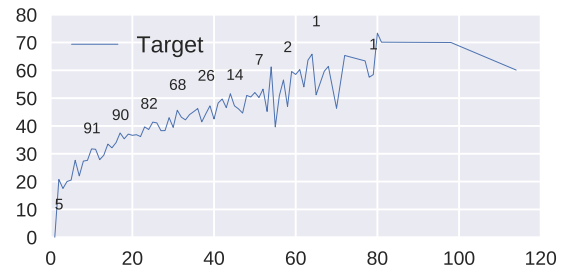
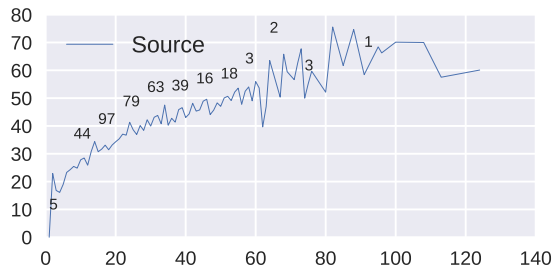
IBM-1+NNctxCNN



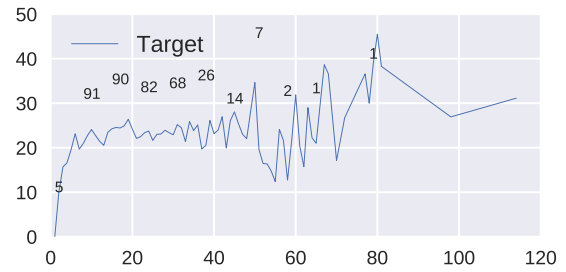
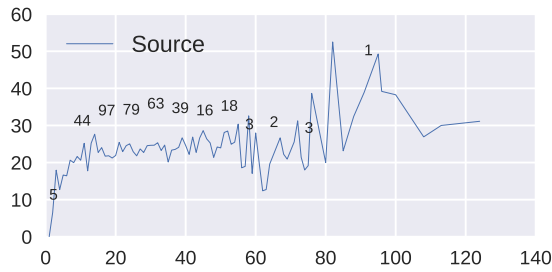
IBM-1+NNChar



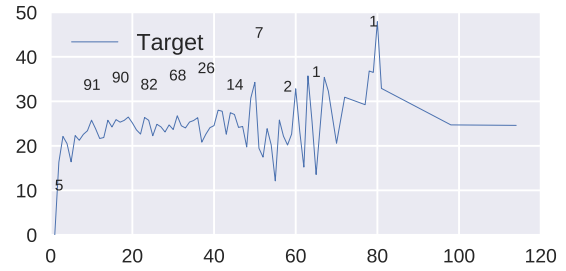
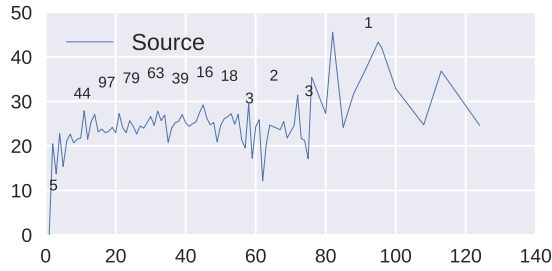
IBM-1+NNCharWord



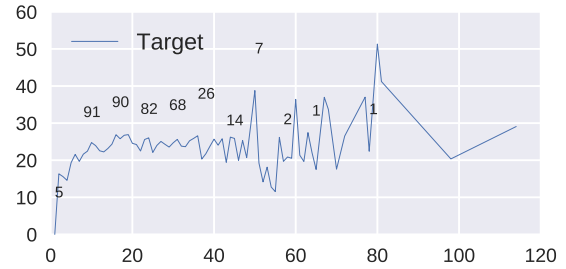
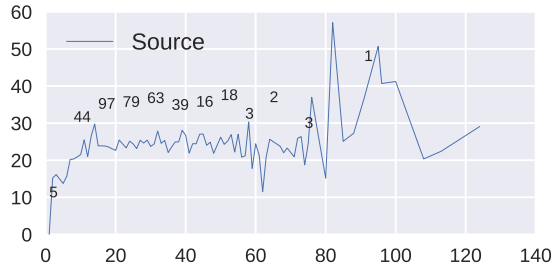
HMM+NN



HMM+NN+CtxCc



HMM+NN+CtxCNN



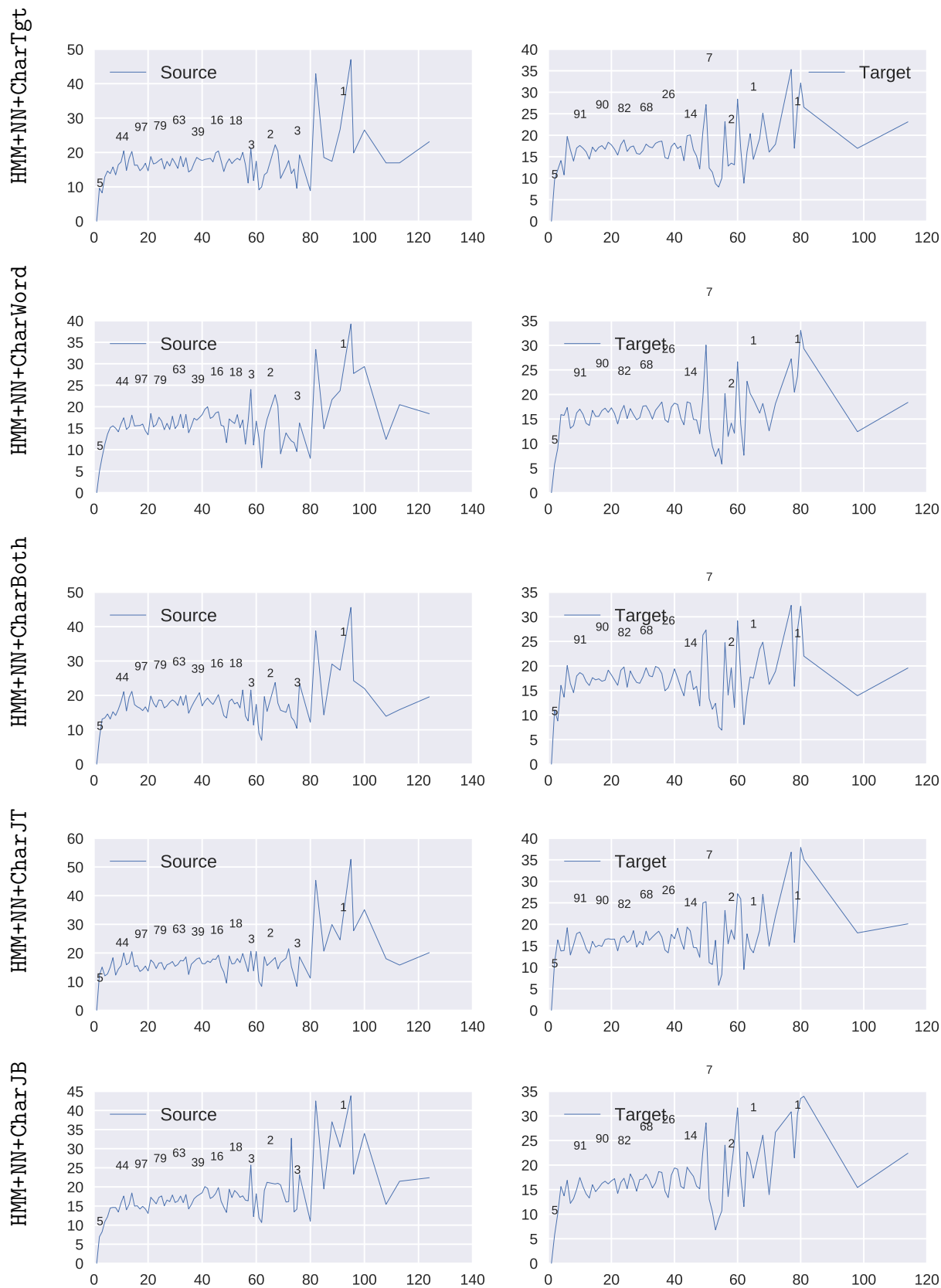
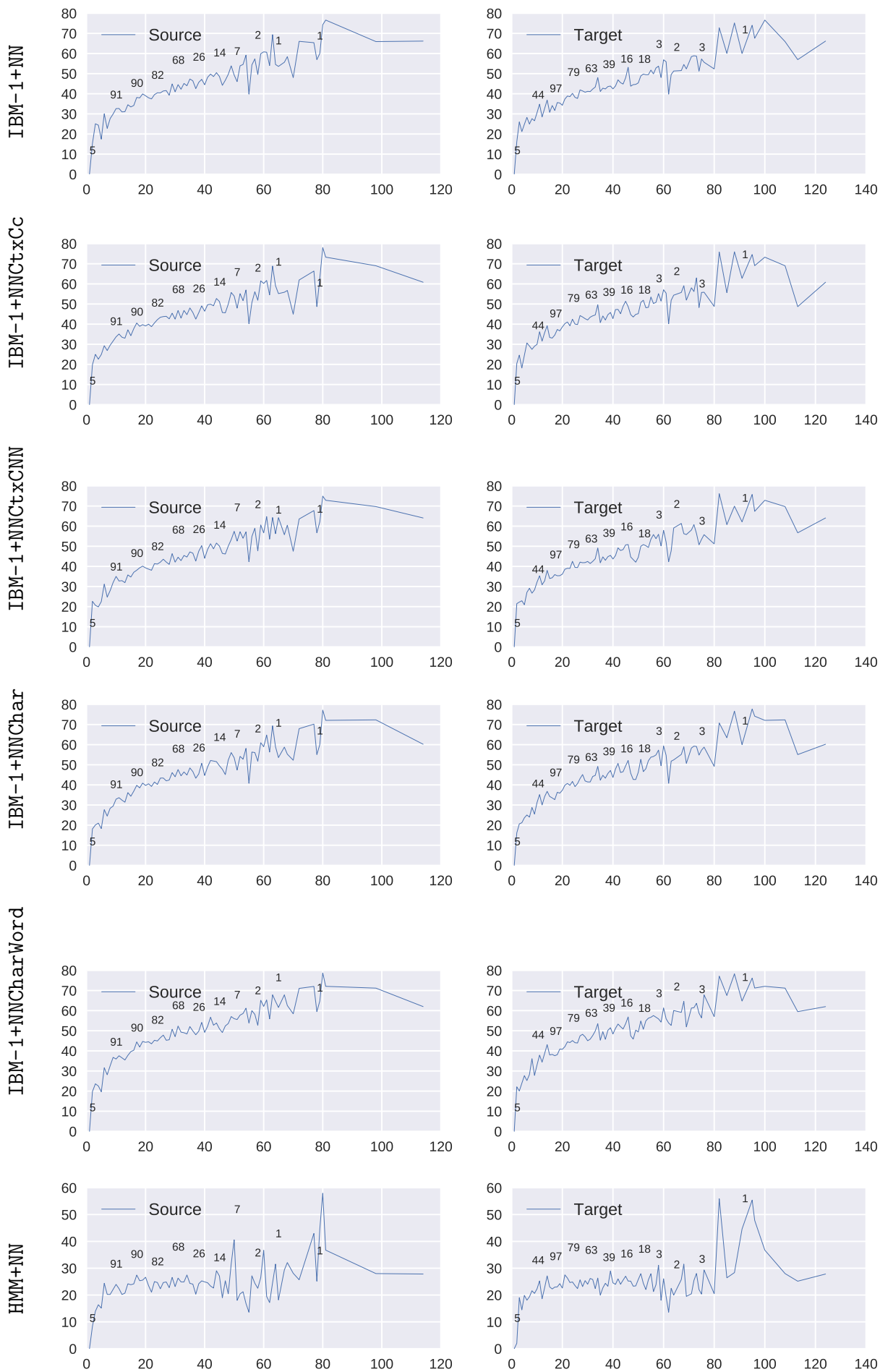
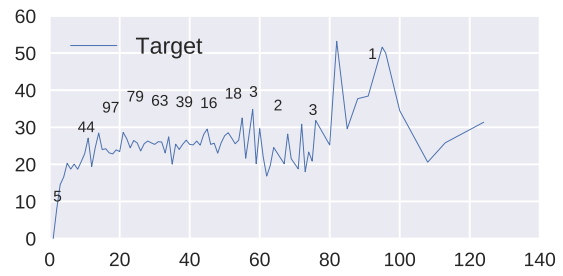
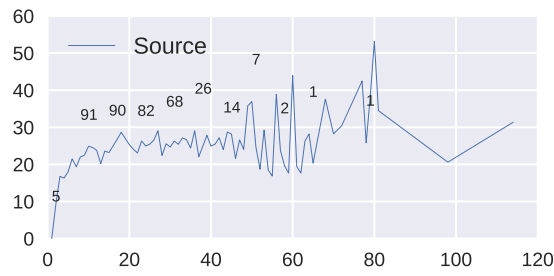


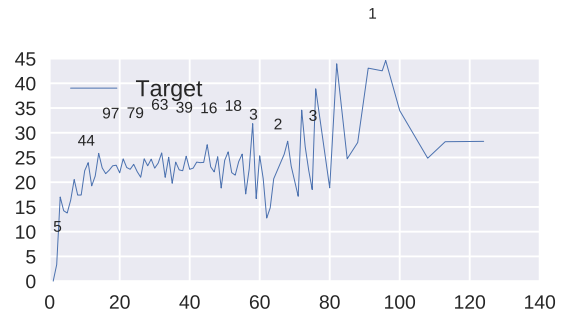
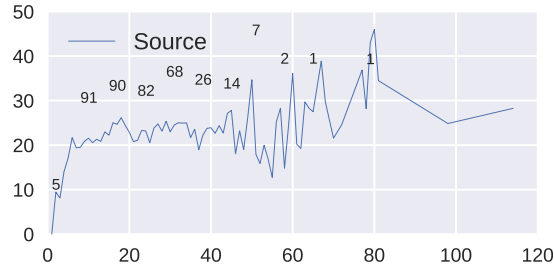
Figure B.53: Neural models for the direction English-Czech: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



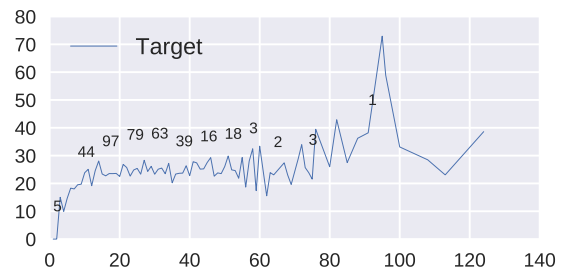
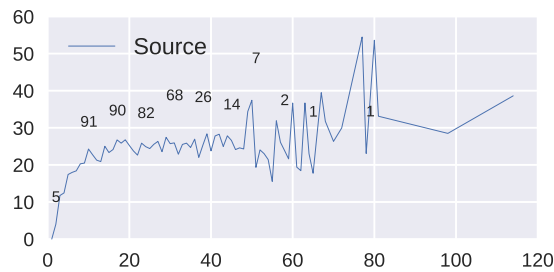
HMM+NN+CtxCc



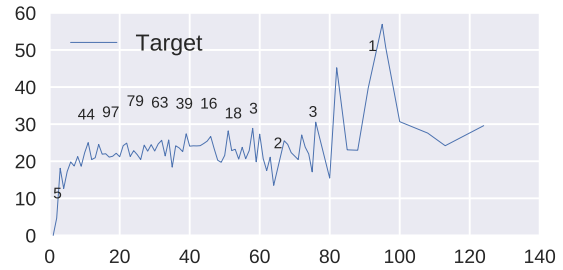
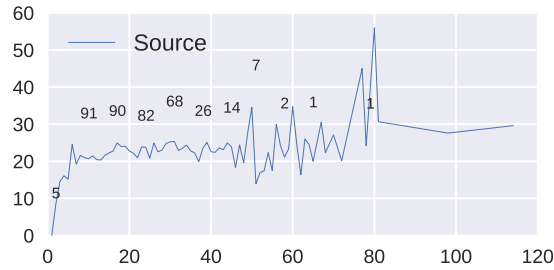
HMM+NN+CtxCNN



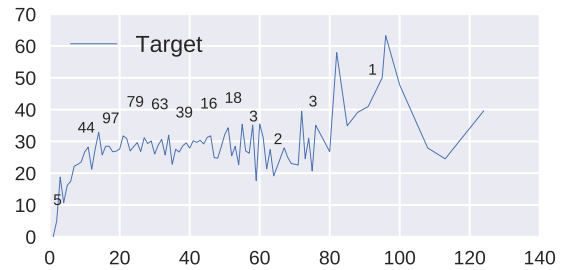
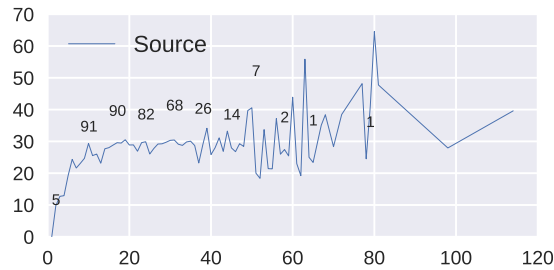
HMM+NN+CharTgt



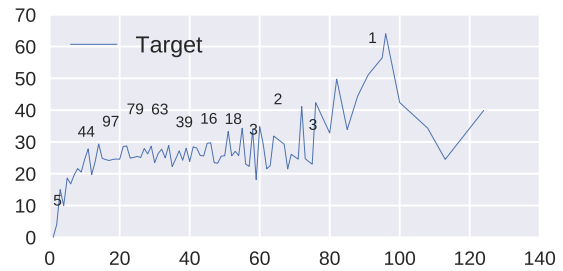
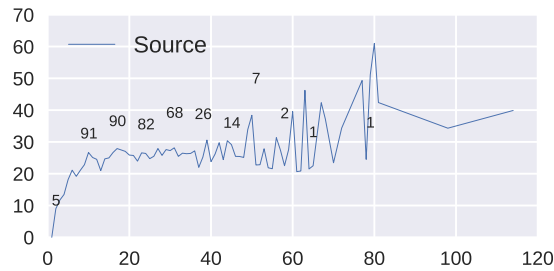
HMM+NN+CharWord



HMM+NN+CharBoth



HMM+NN+CharJT



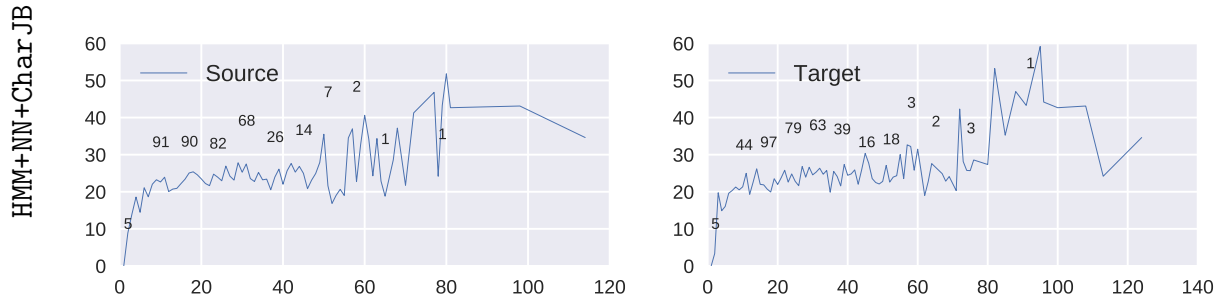
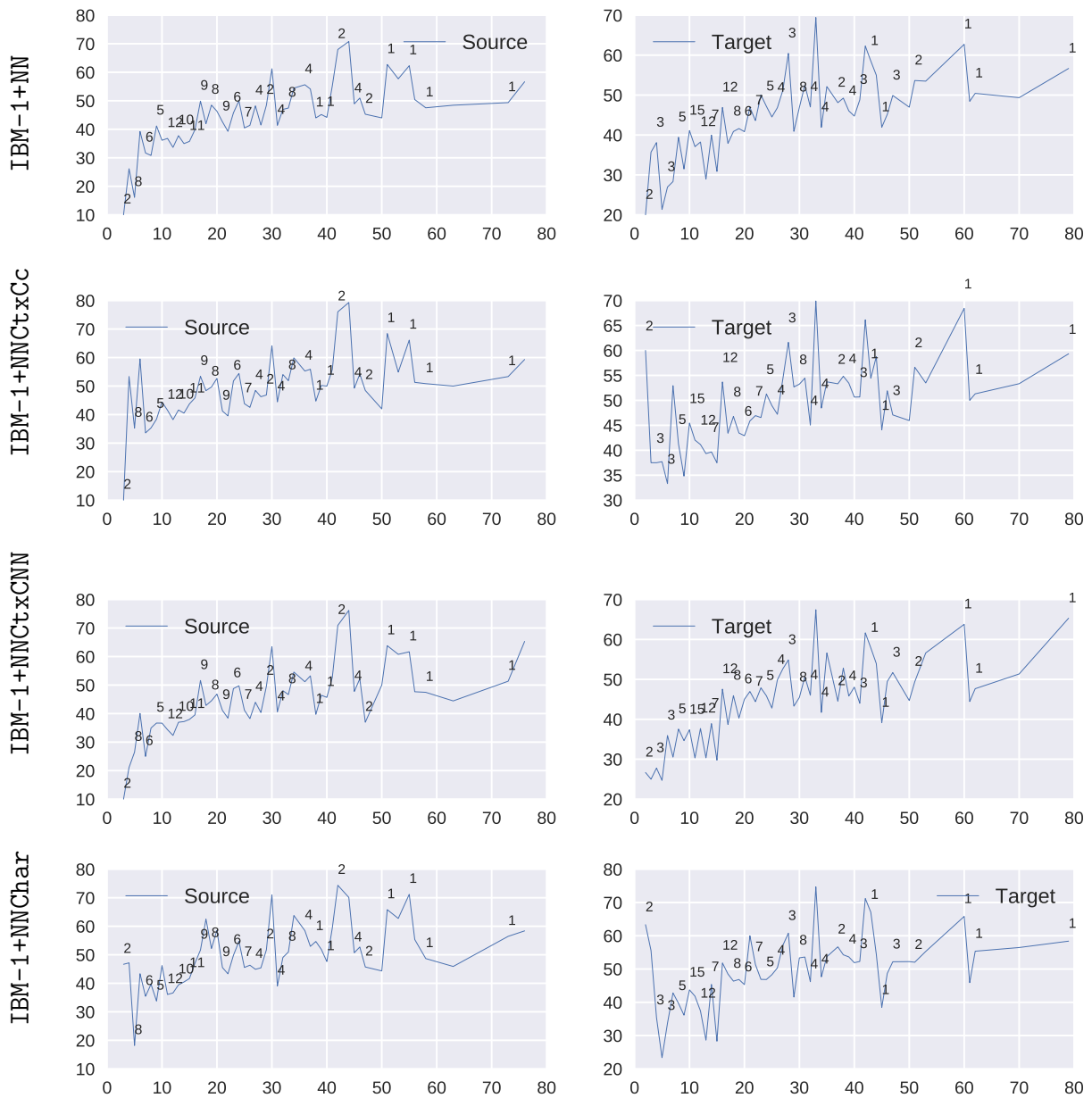
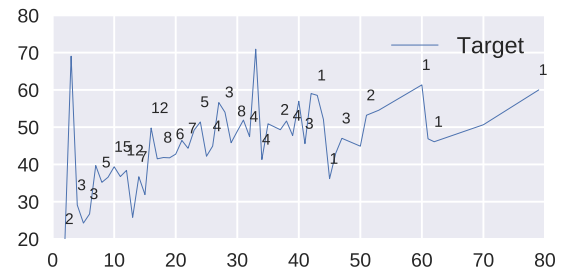
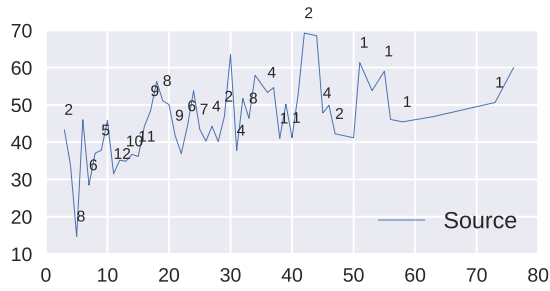


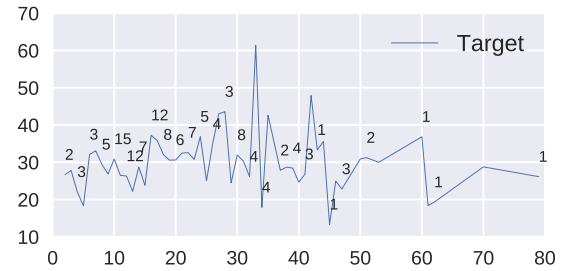
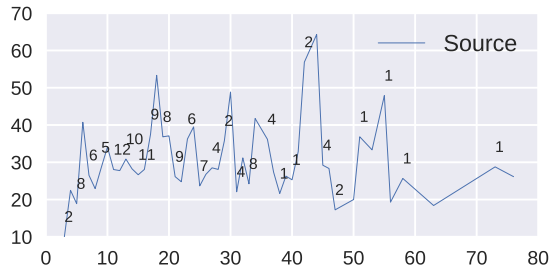
Figure B.54: Neural models for the direction Czech-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



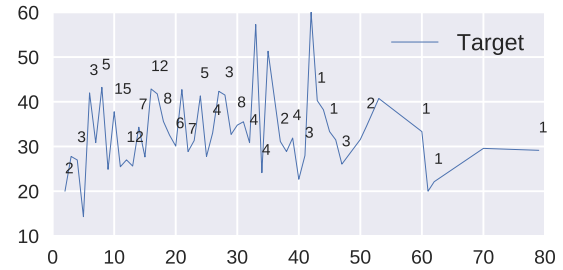
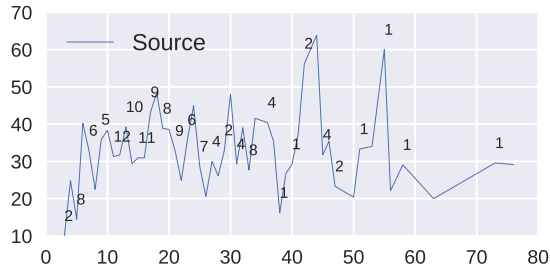
IBM-1+NNCharWord



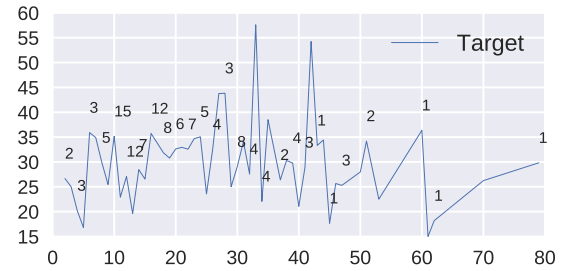
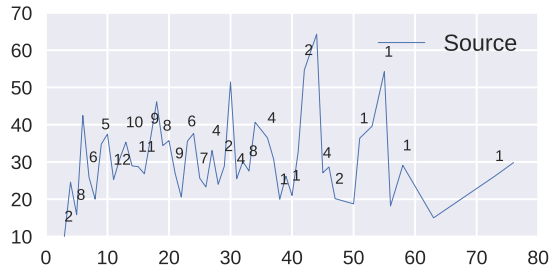
HMM+NN



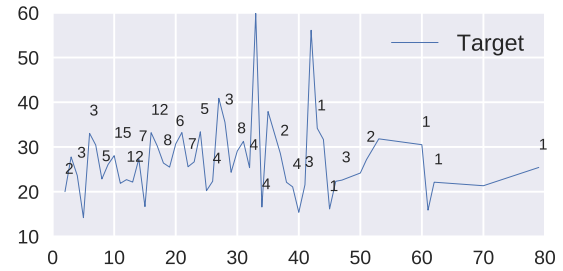
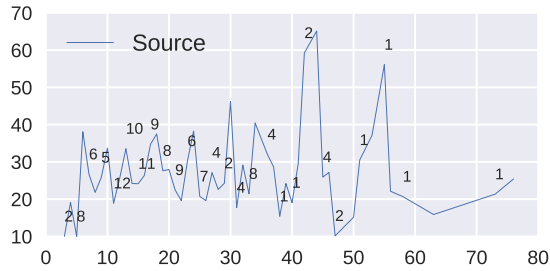
HMM+NN+CtxCc



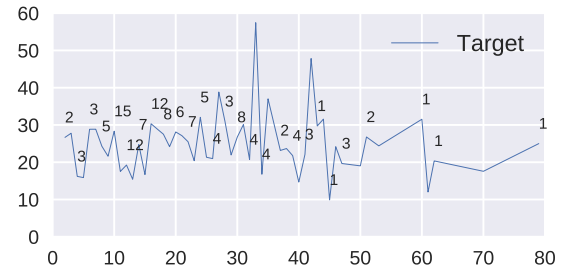
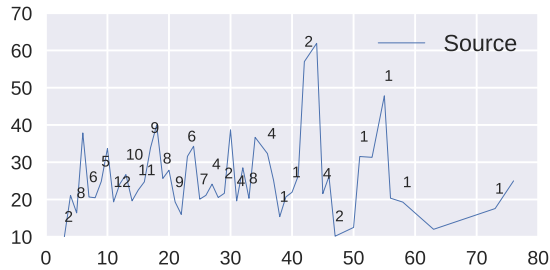
HMM+NN+CtxCNN



HMM+NN+CharTgt



HMM+NN+CharWord



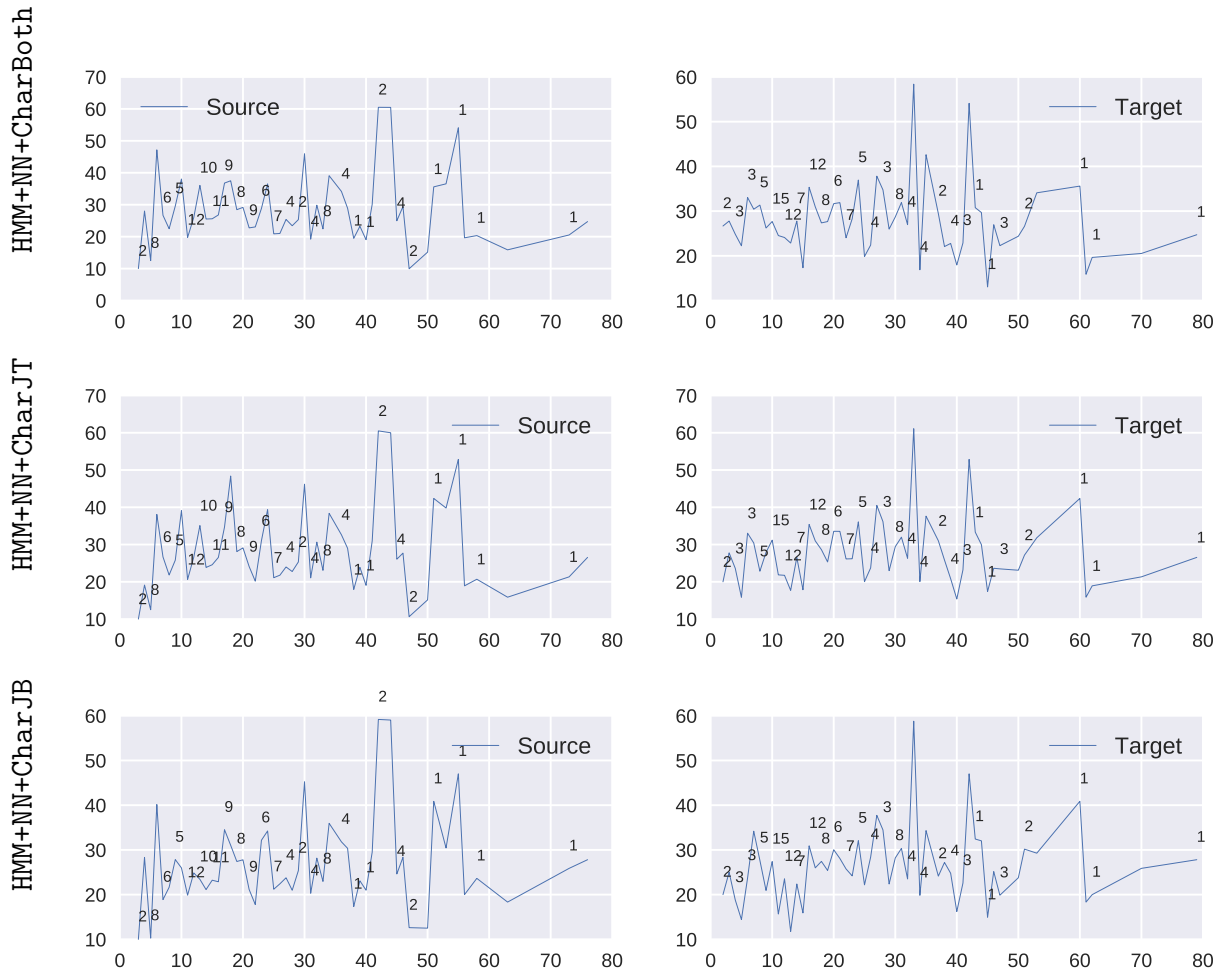
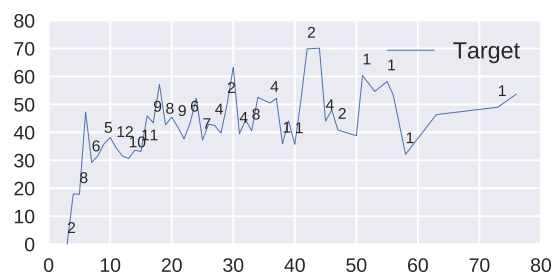
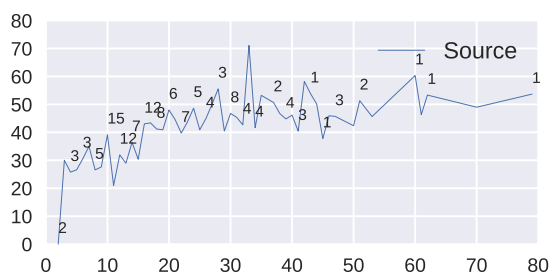


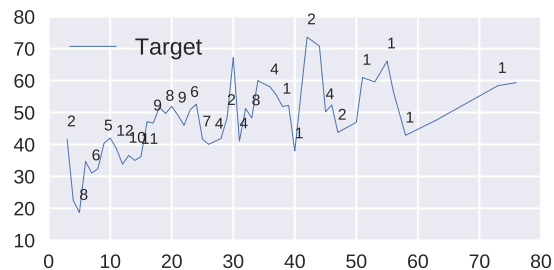
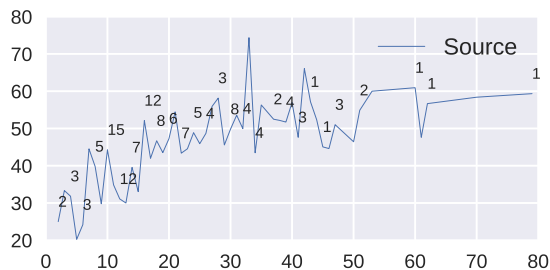
Figure B.55: Neural models for the direction English-Romanian: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



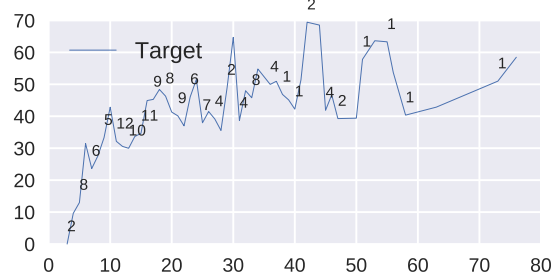
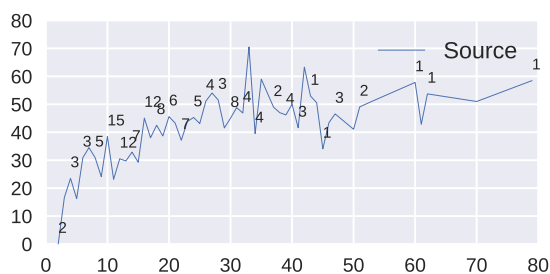
IBM-1+NNctxCNN



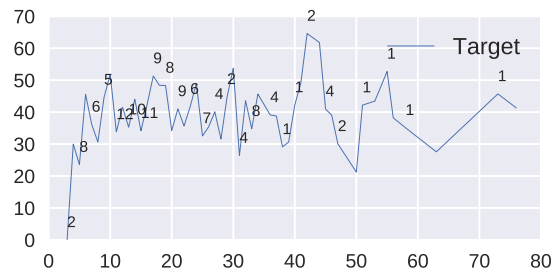
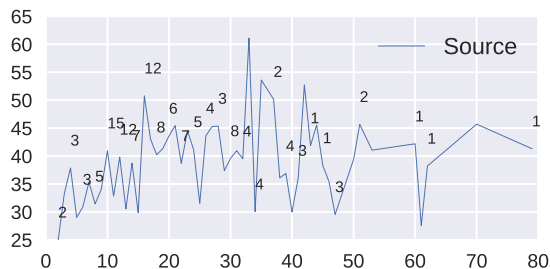
IBM-1+NNChar



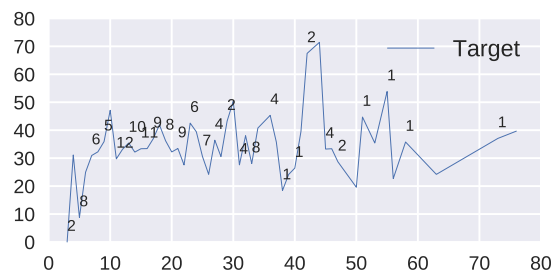
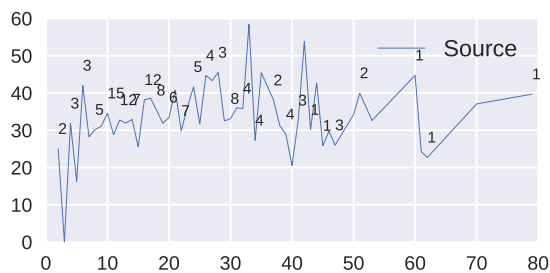
IBM-1+NNCharWord



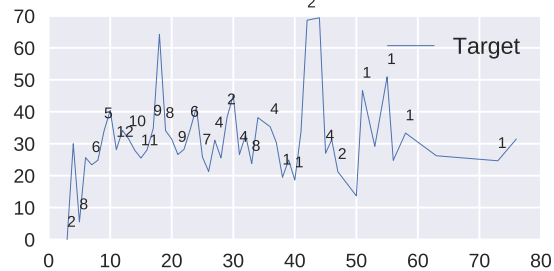
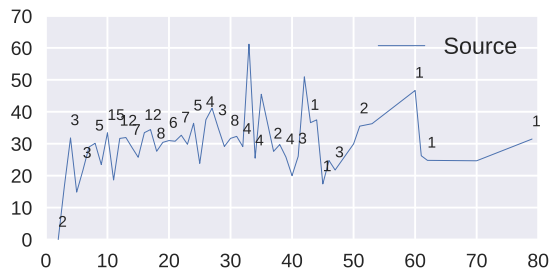
HMM+NN



HMM+NN+CtxCc



HMM+NN+CtxCNN



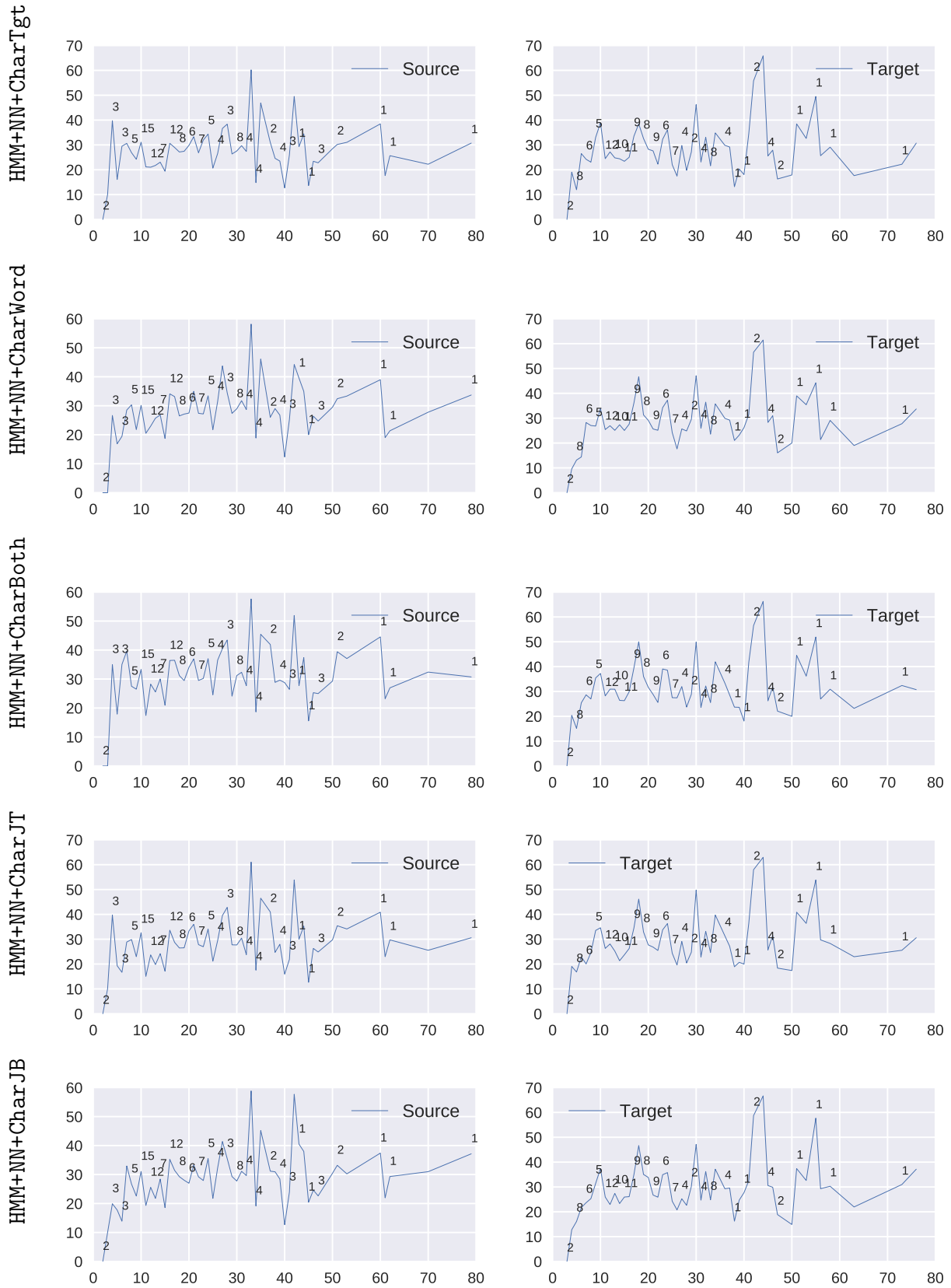
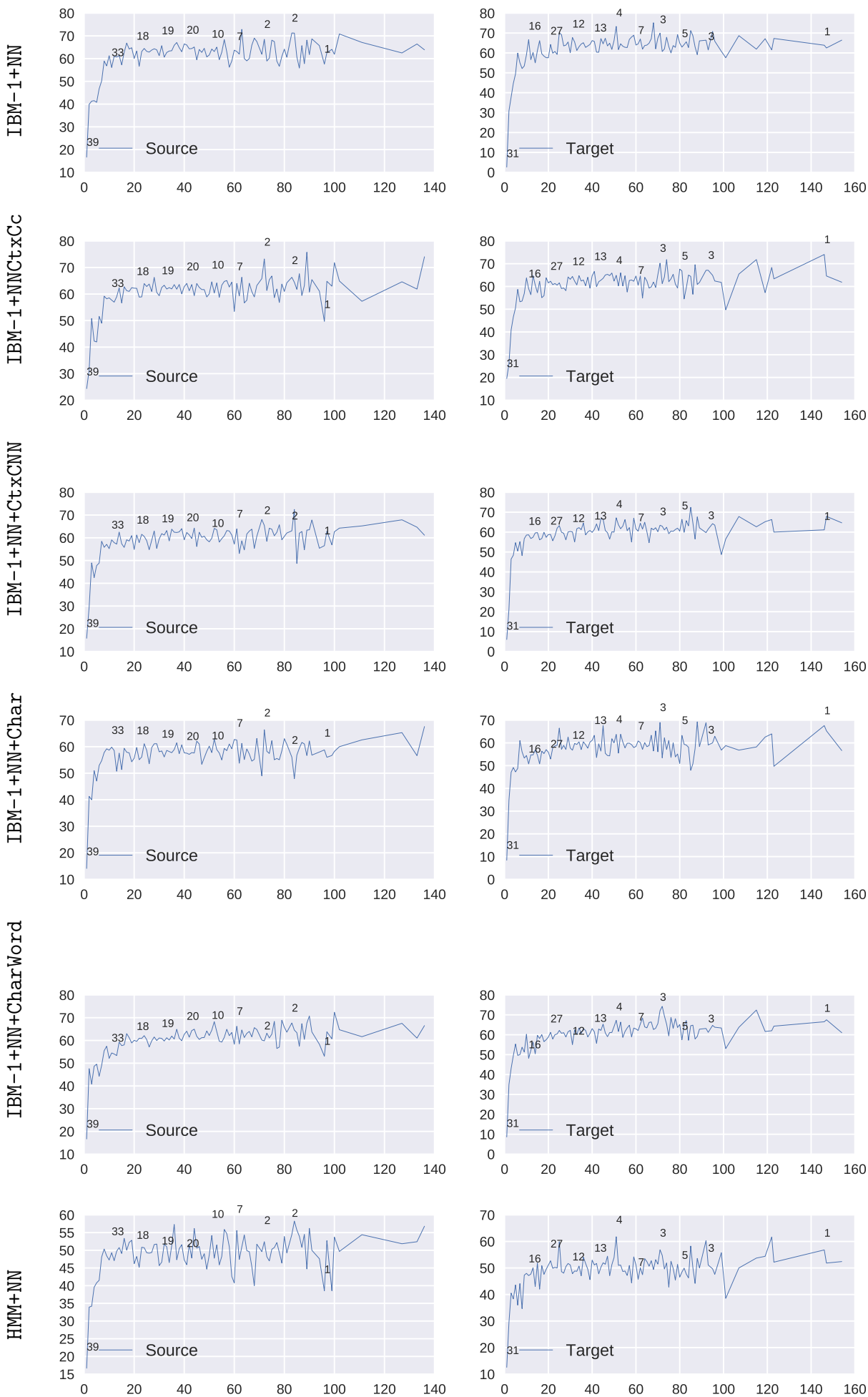
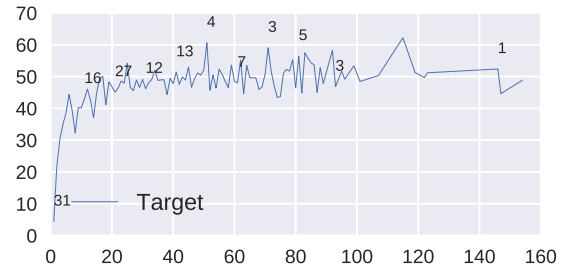
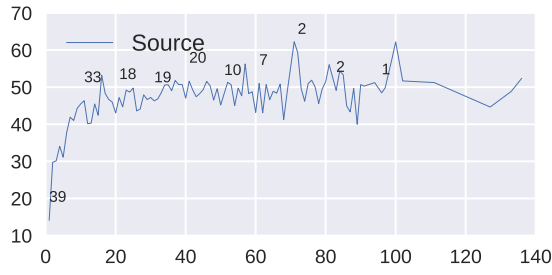


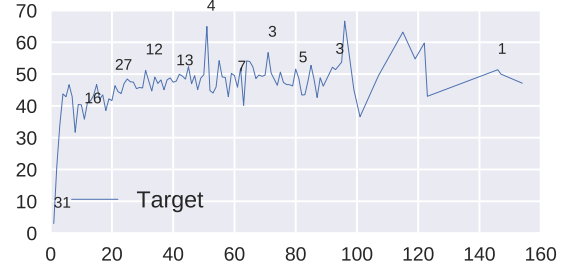
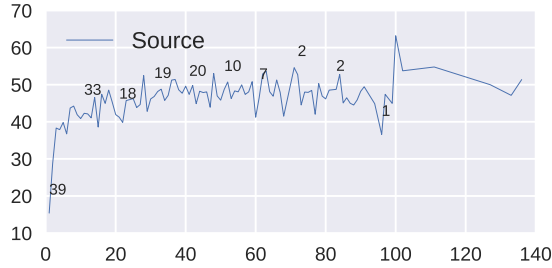
Figure B.56: Neural models for the direction Romanian-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



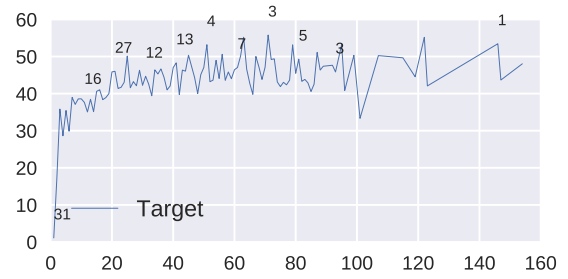
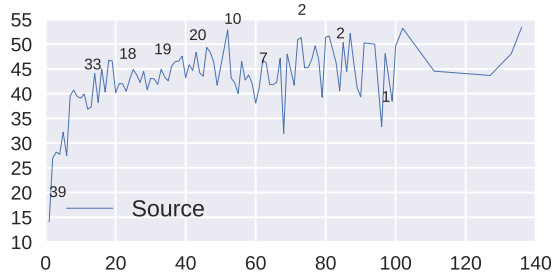
HMM+NN+CtxCc



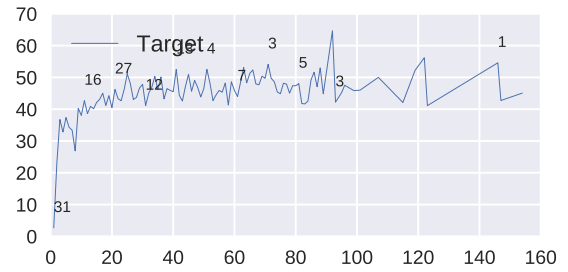
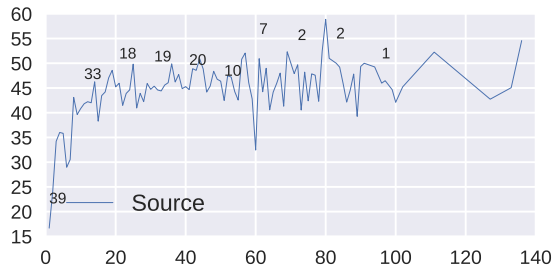
HMM+NN+CtxCNN



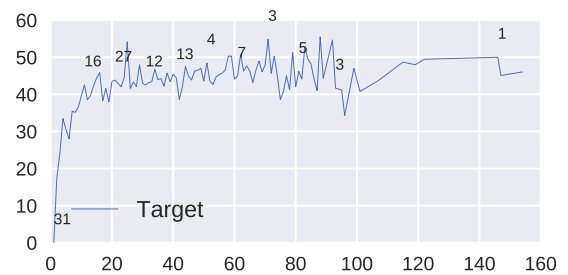
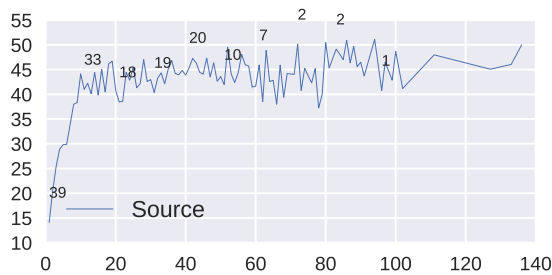
HMM+NN+CharTgt



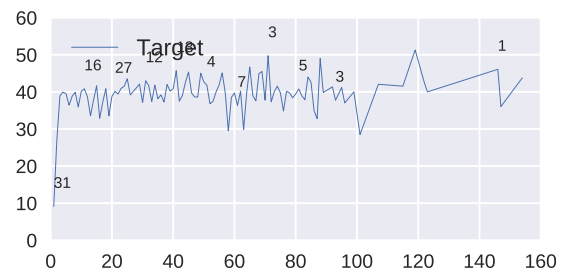
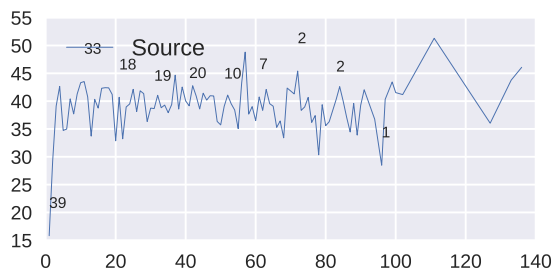
HMM+NN+CharWord



HMM+NN+CharBoth



HMM+NN+CharJT



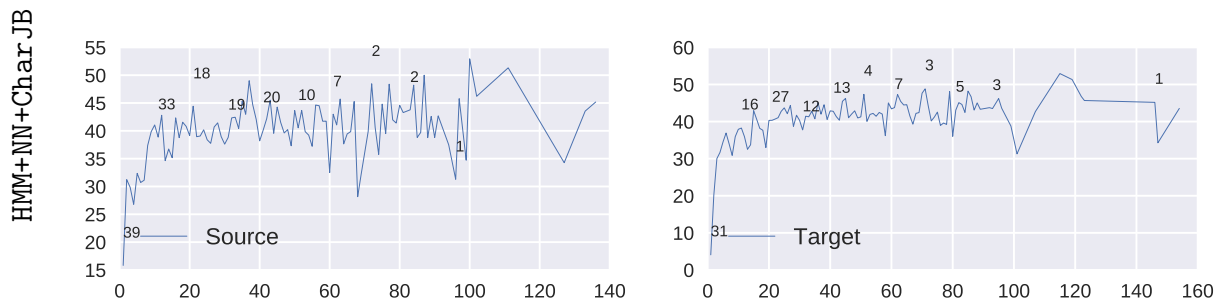
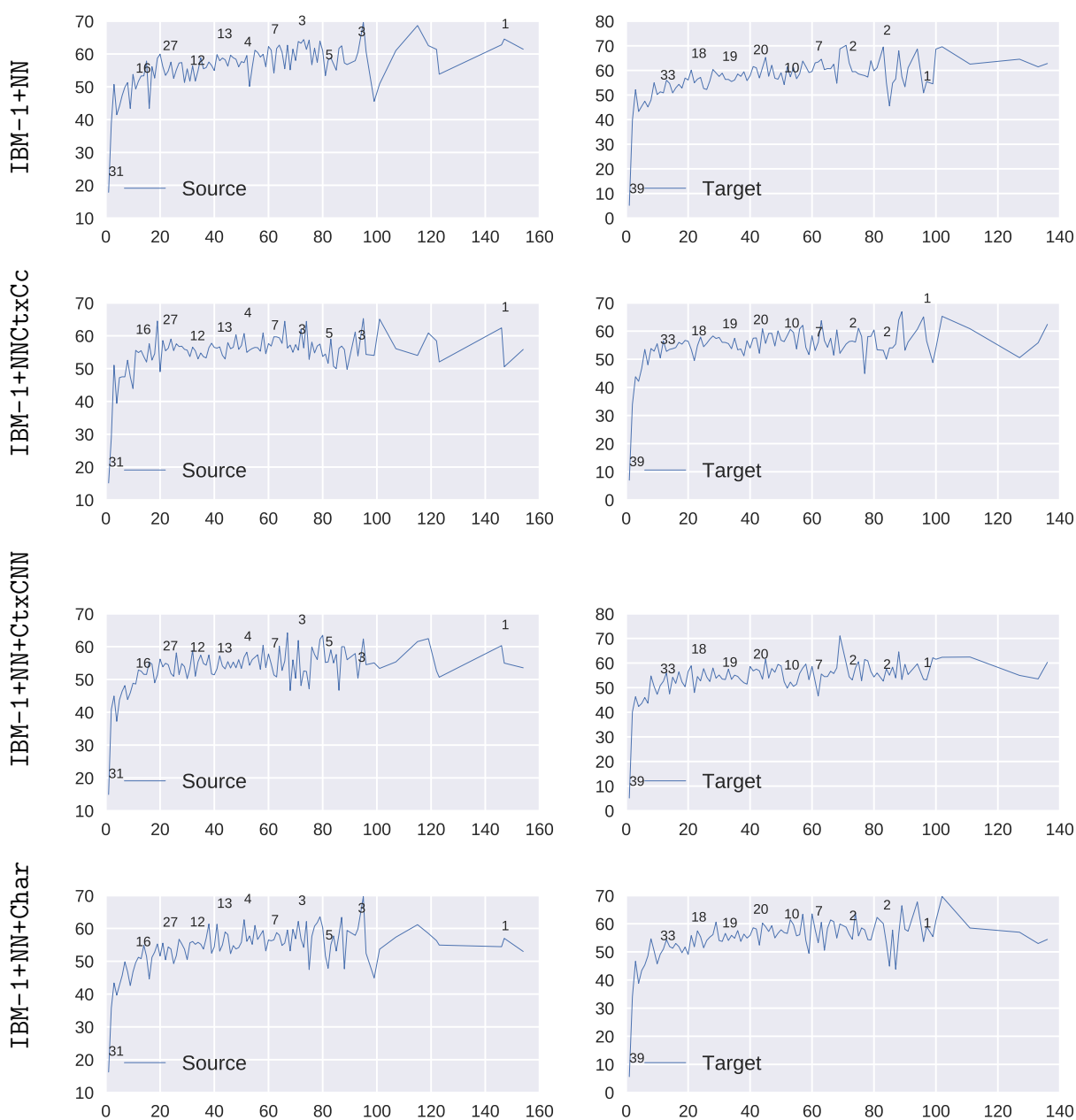
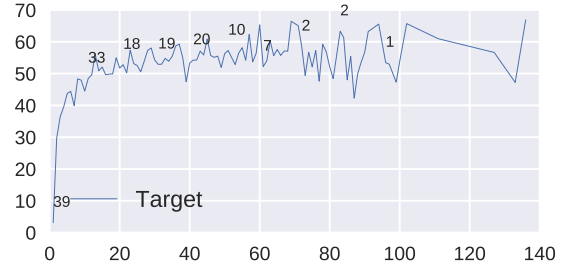
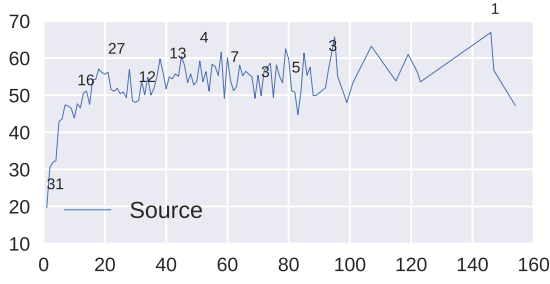


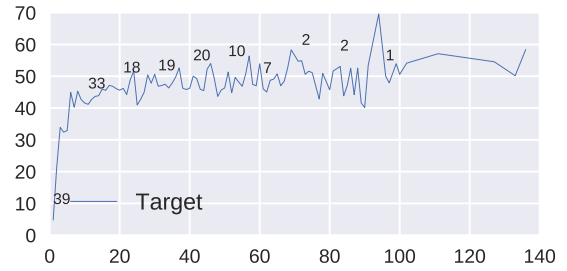
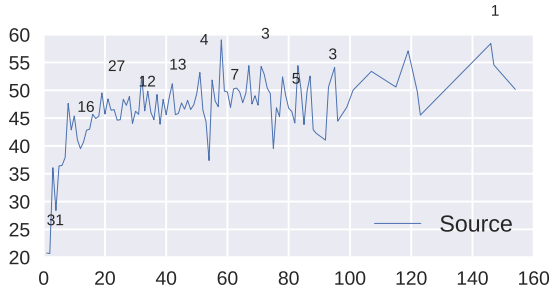
Figure B.57: Neural models for the direction English-Japanese: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



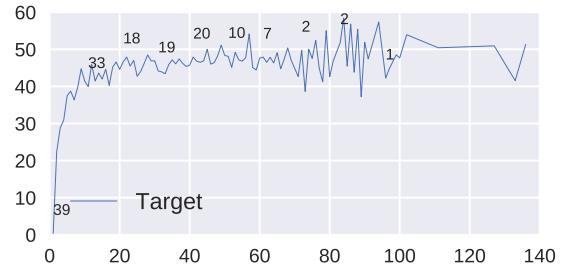
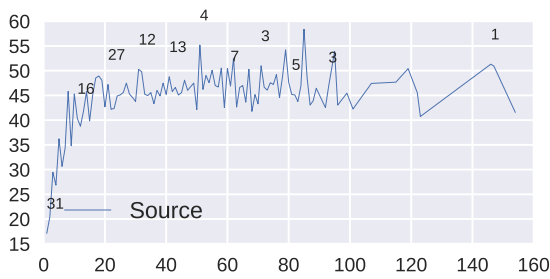
IBM-1+NN+CharWord



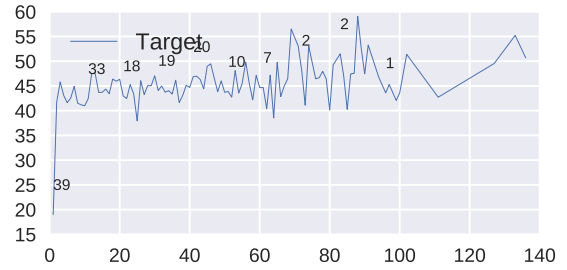
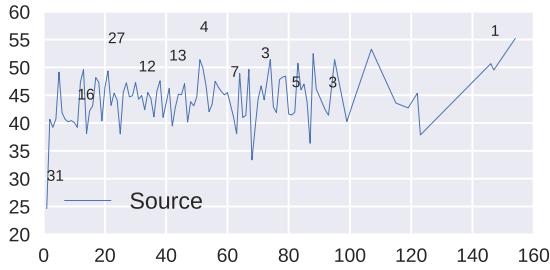
HMM+NN



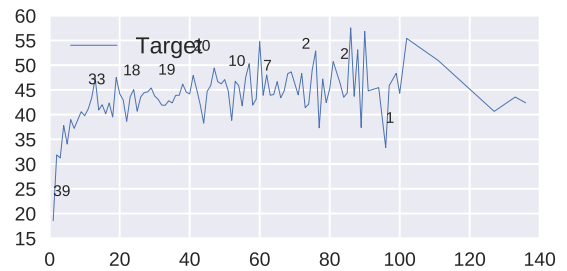
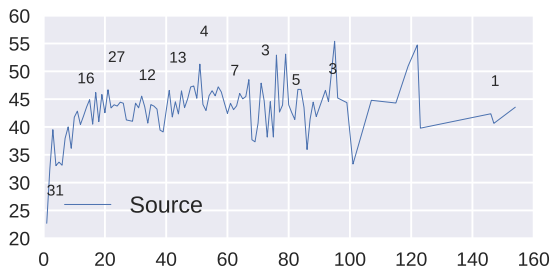
HMM+NN+CtxCc



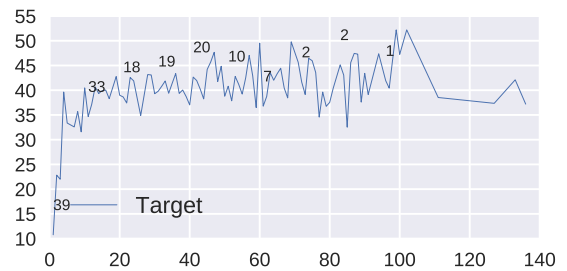
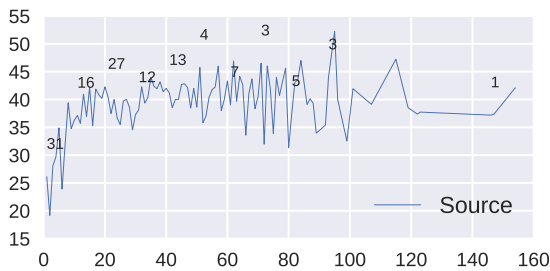
HMM+NN+CtxCNN



HMM+NN+CharTgt



HMM+NN+CharWord



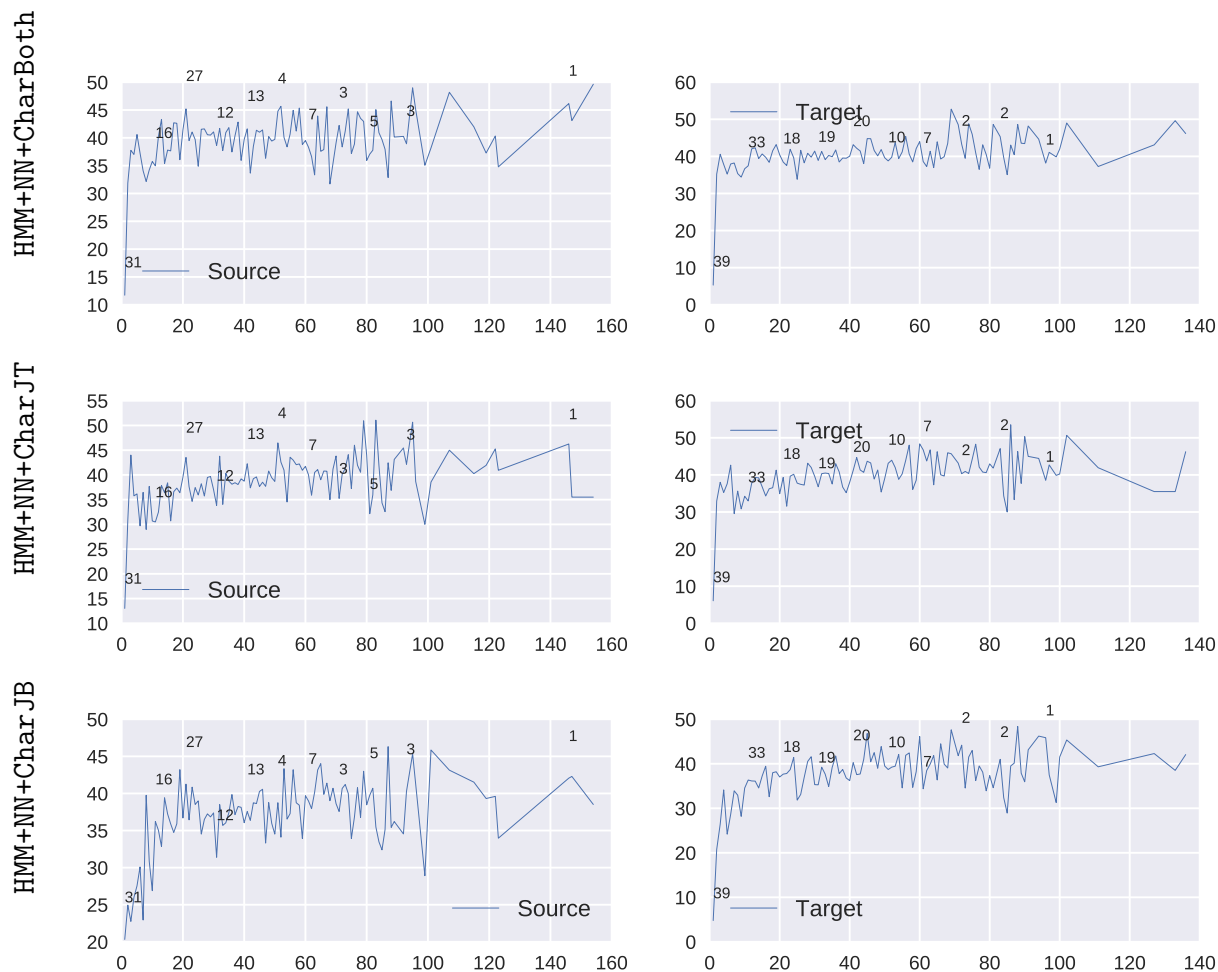
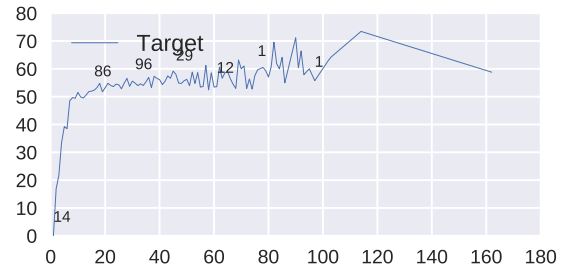
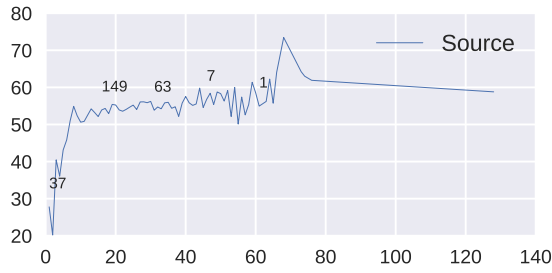


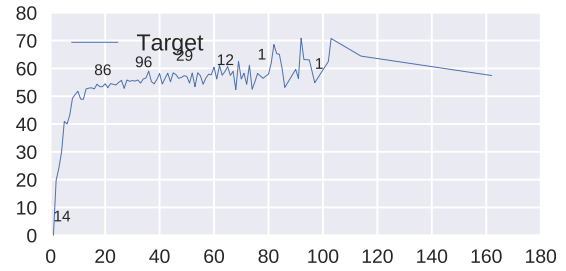
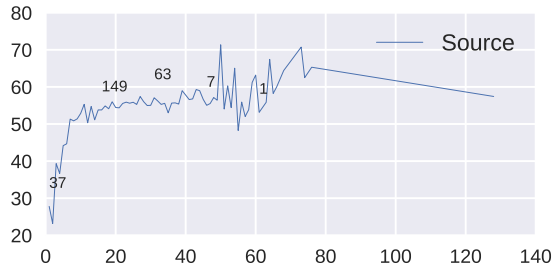
Figure B.58: Neural models for the direction Japanese-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



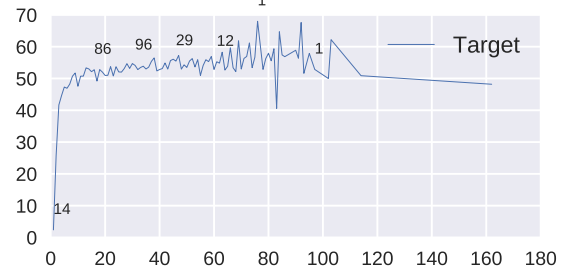
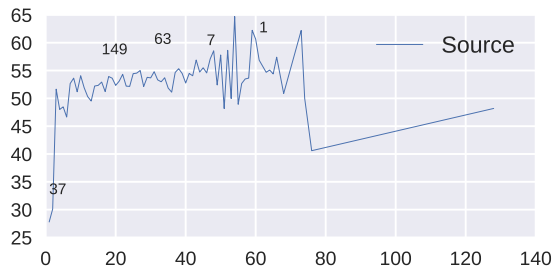
IBM-1+NN+CtxCNN



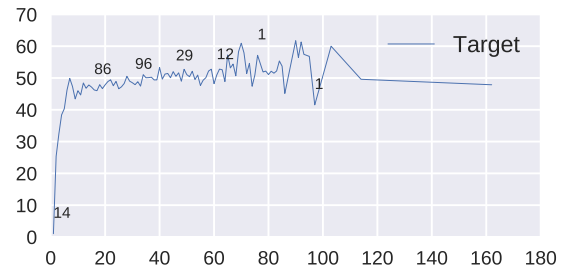
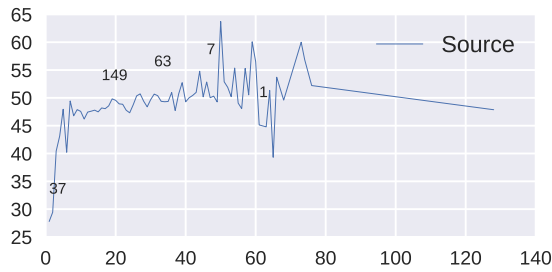
IBM-1+NN+Char



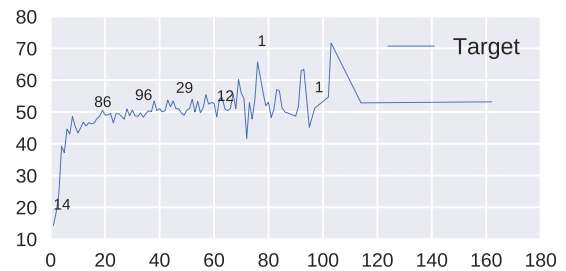
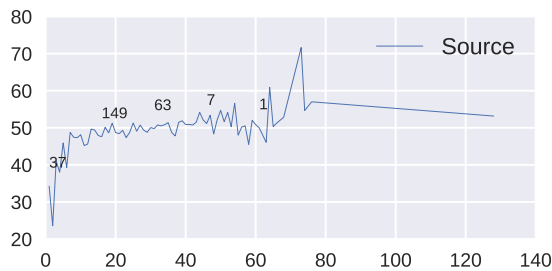
IBM-1+NN+CharWord



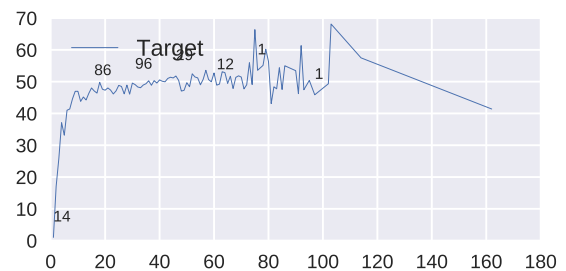
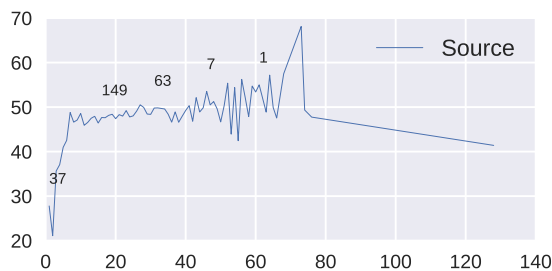
HMM+NN



HMM+NN+CtxCc



HMM+NN+CtxCNN



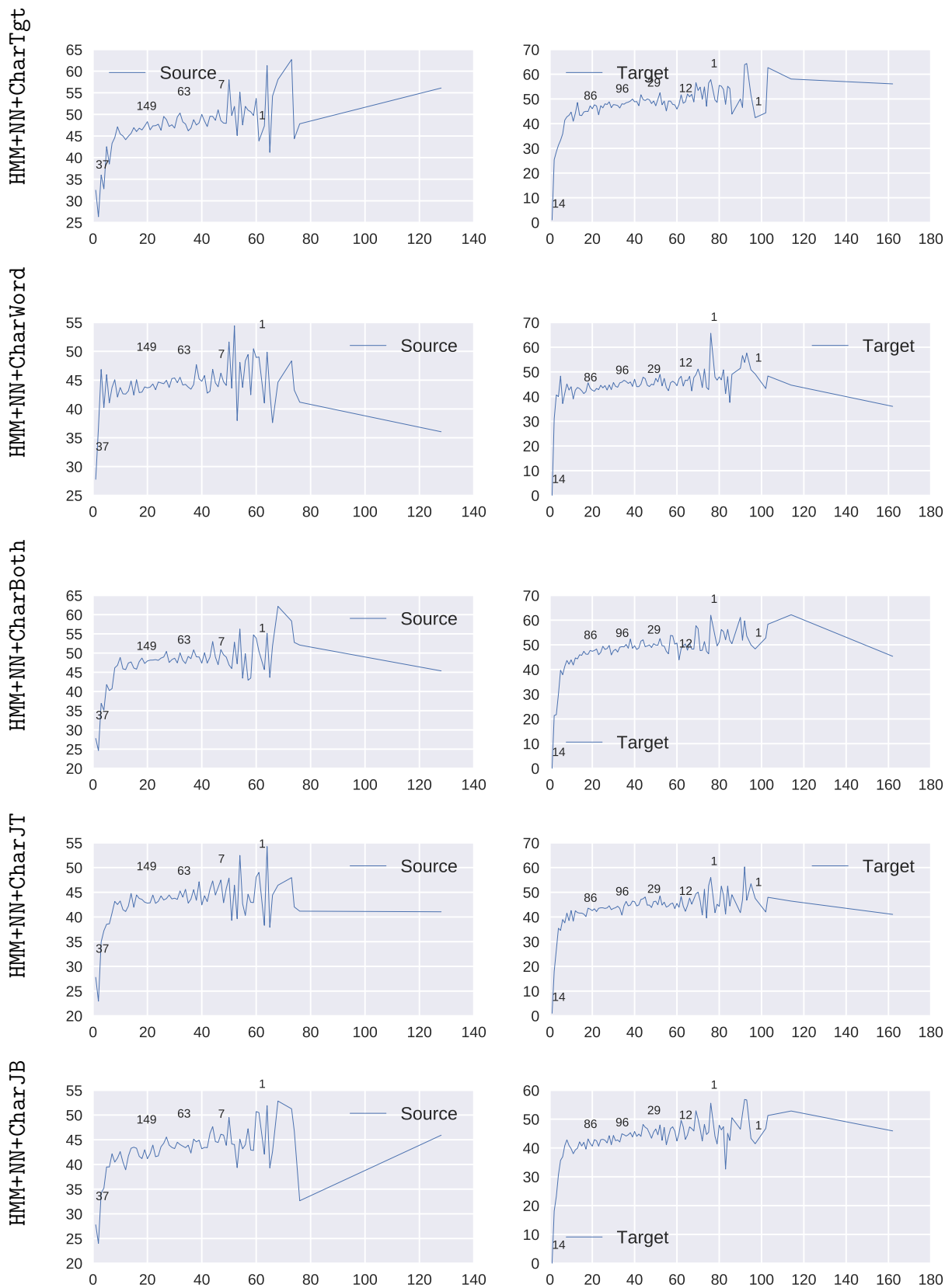
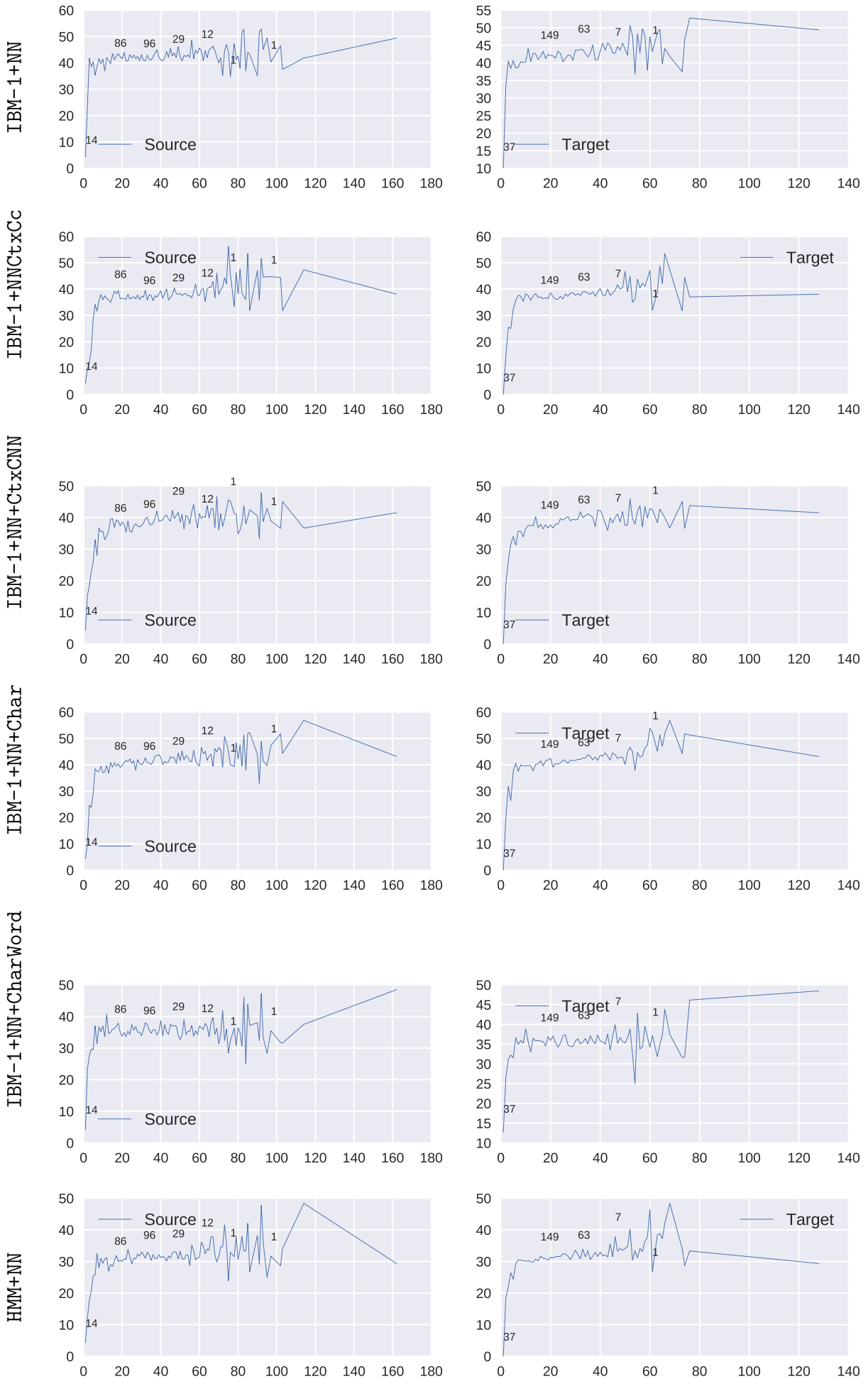
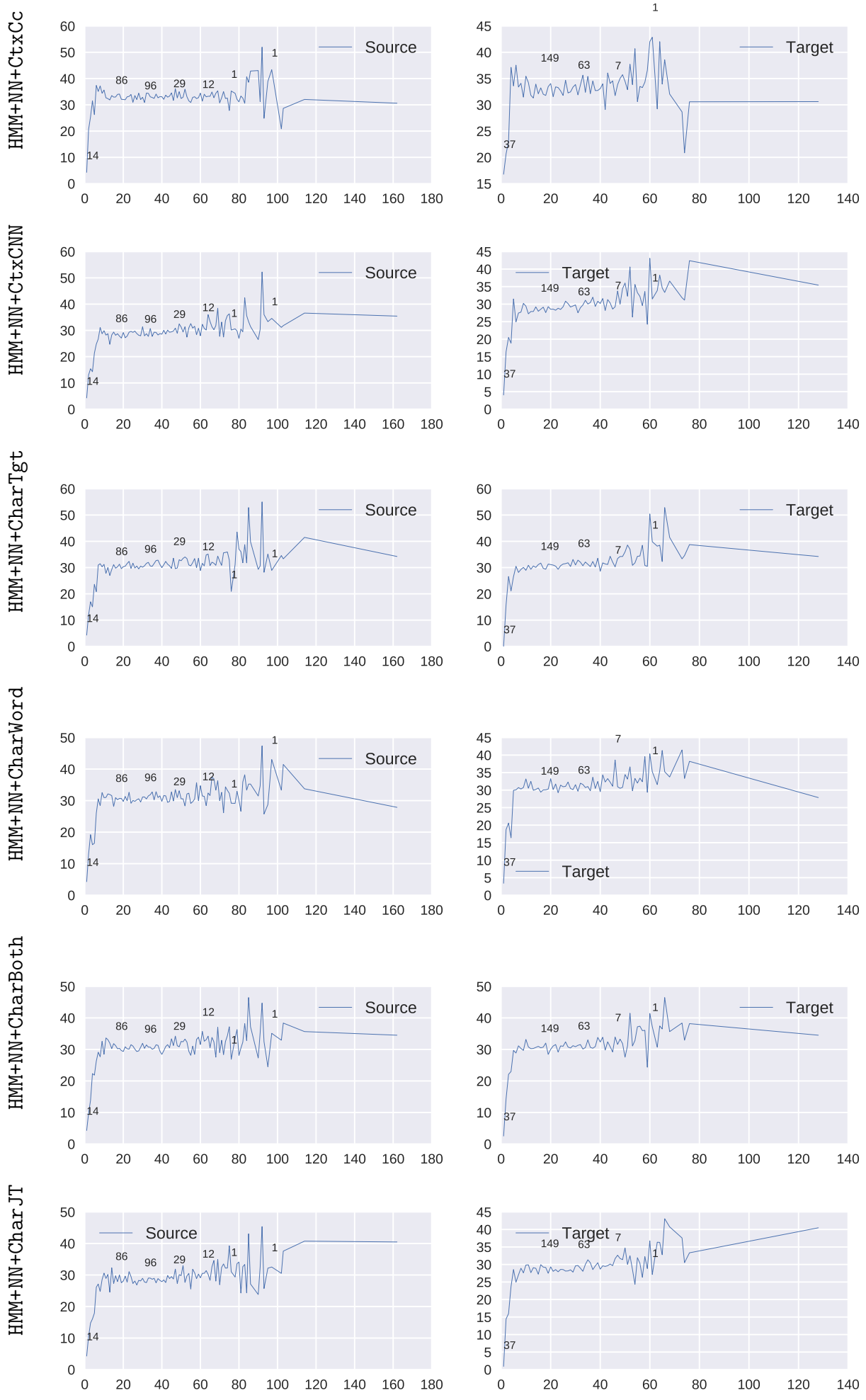


Figure B.59: Neural models for the direction English-Vietnamese: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.





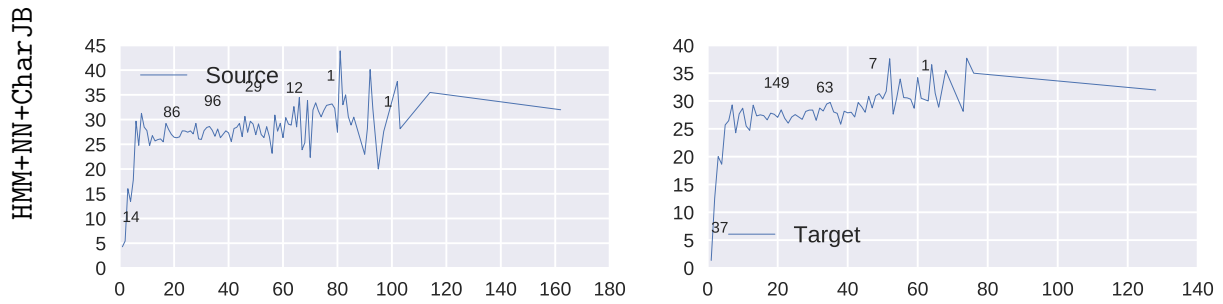
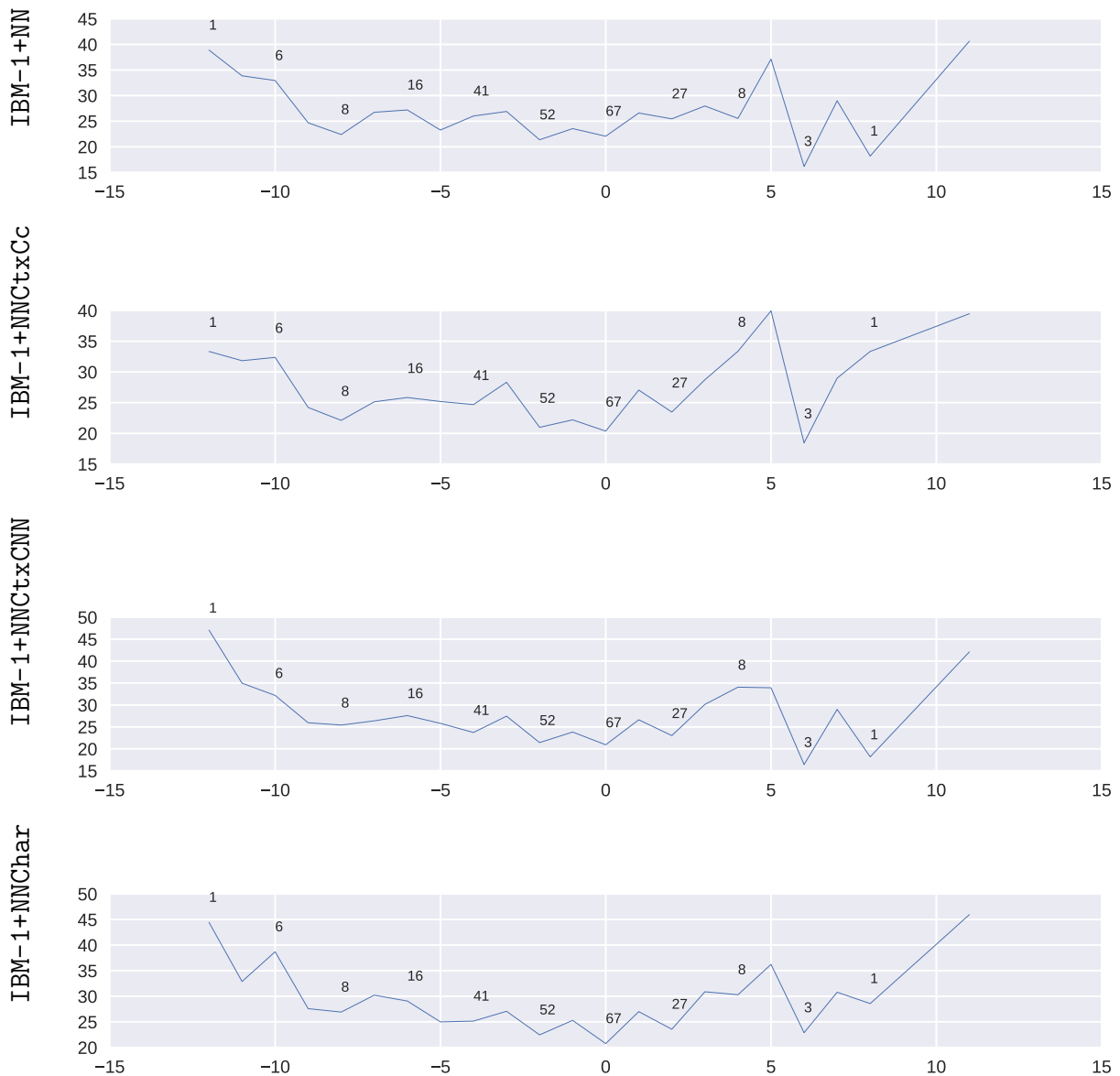
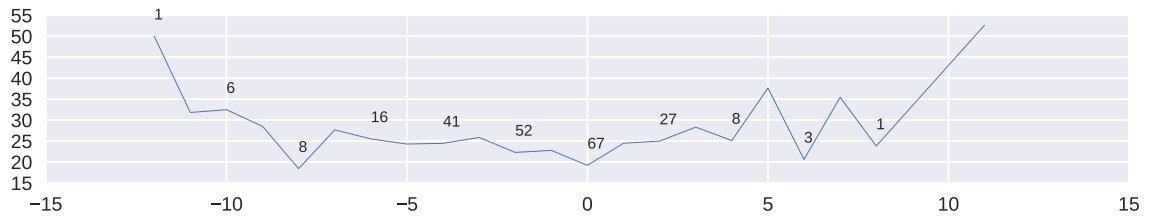


Figure B.60: Neural models for the direction Vietnamese-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.

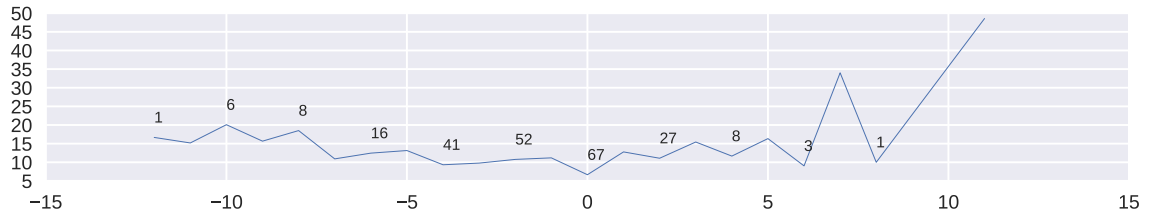
B.11 AER as a function of sentence length difference



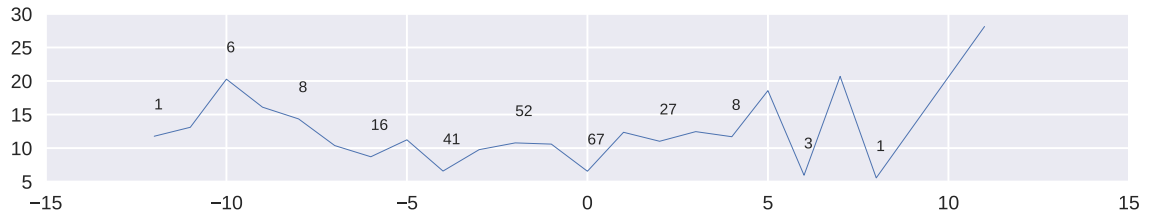
IBM-1+NNCharWord



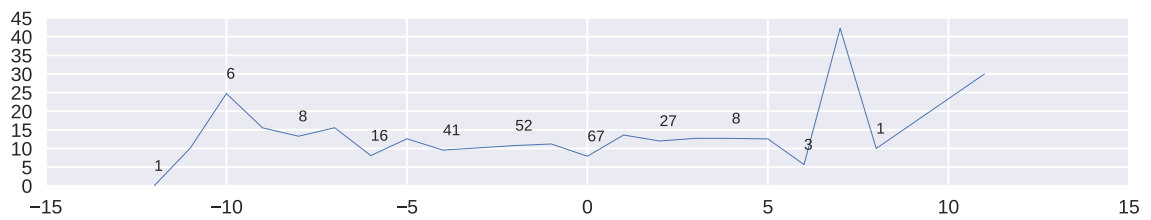
HMM+NN



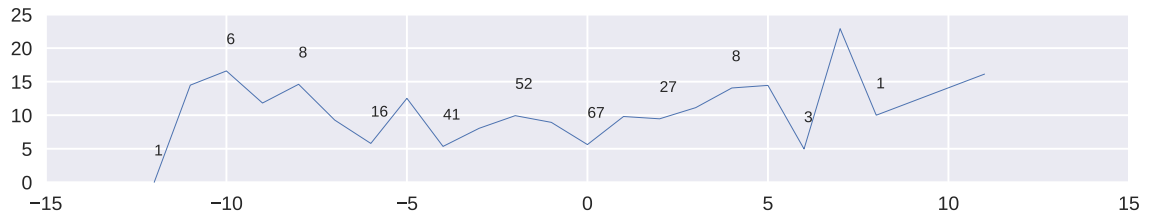
HMM+NNCtxCc



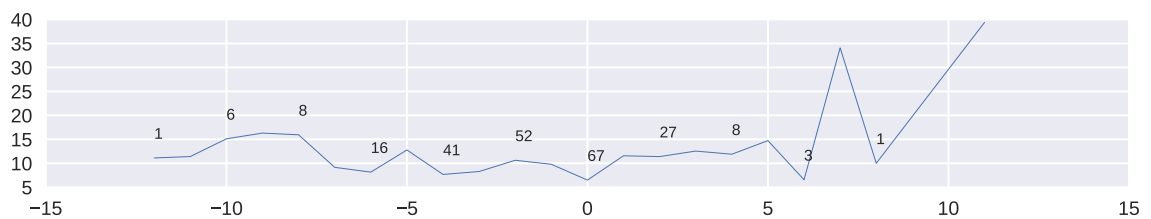
HMM+NNCtxCNN



HMM+NNCharTgt



HMM+NNCharWord



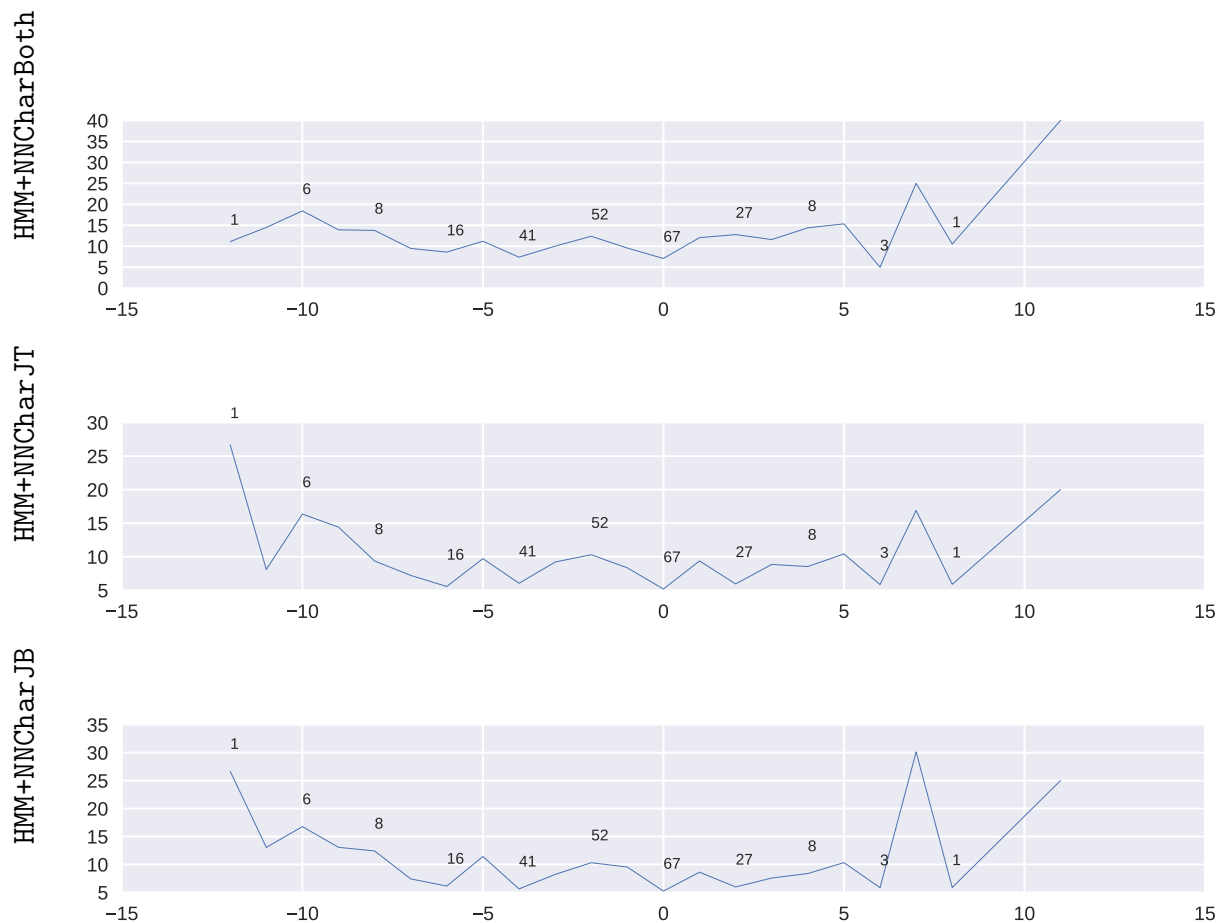
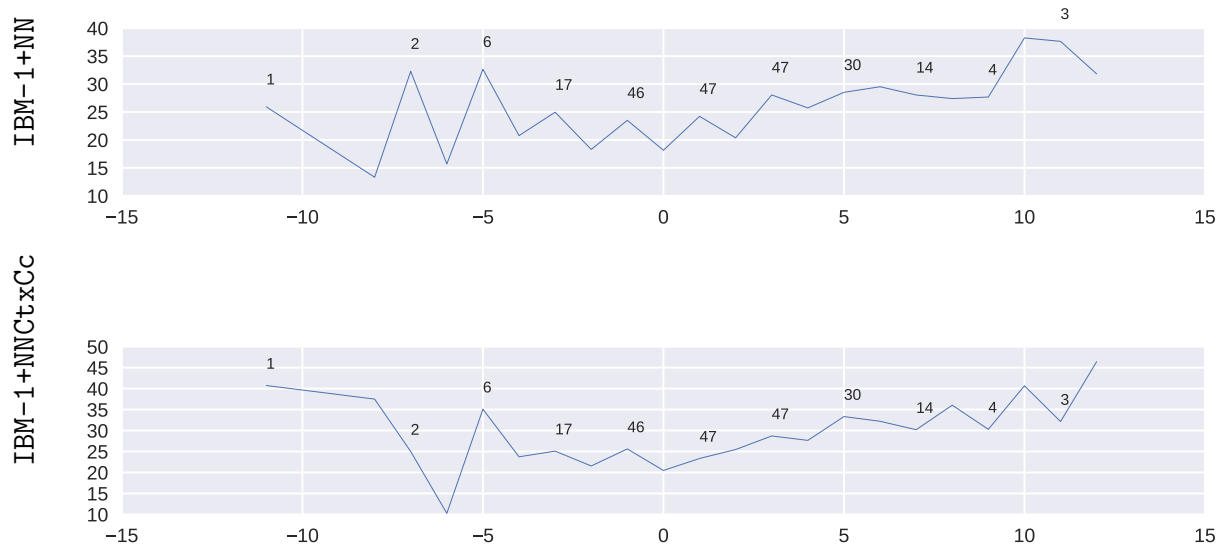
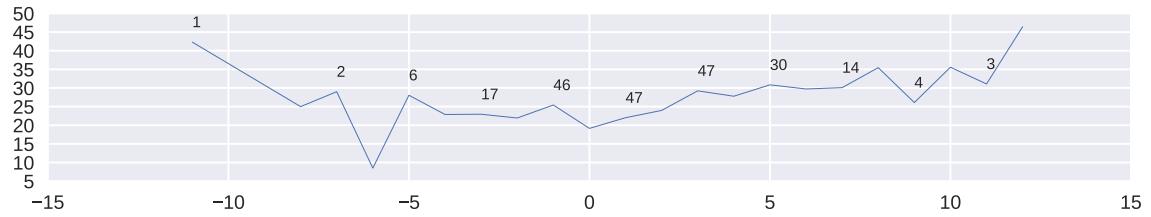


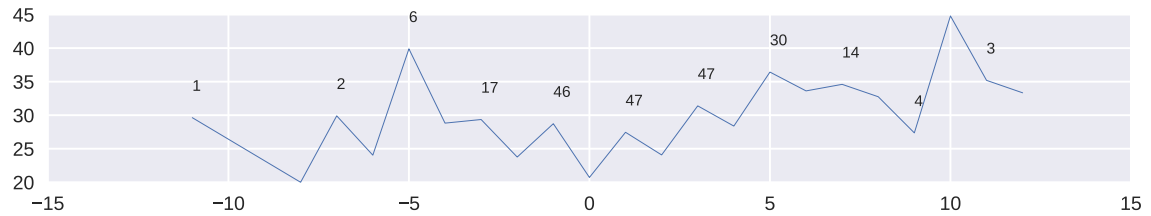
Figure B.61: Neural models for the direction English-French: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



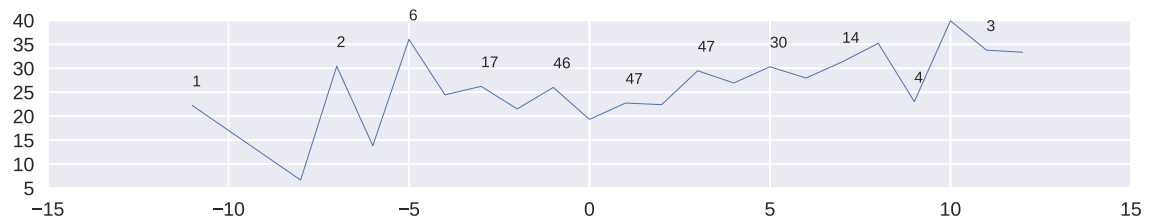
IBM-1+NNCtxCNN



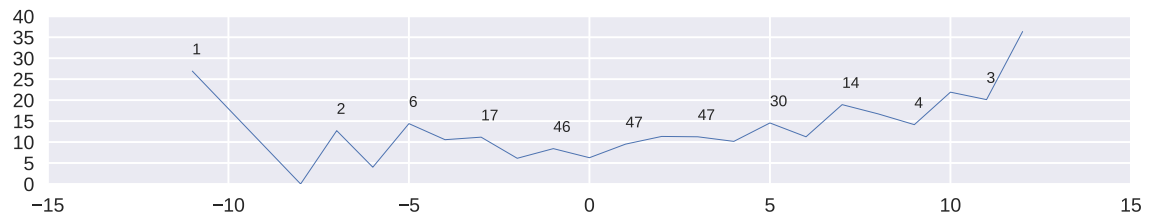
IBM-1+NNChar



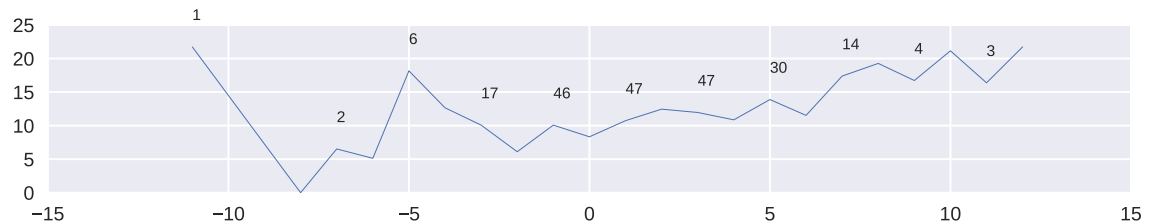
IBM-1+NNCharWord



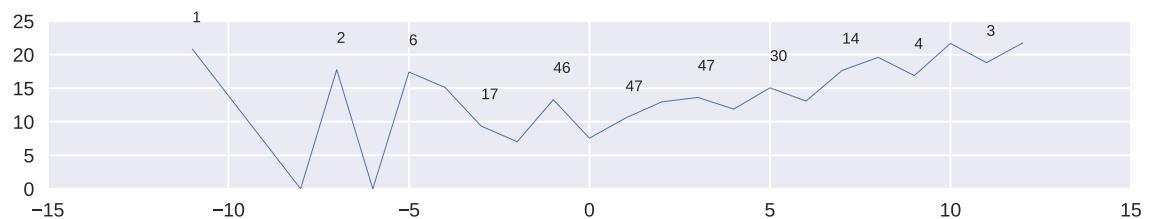
HMM+NN



HMM+NNCtxCc



HMM+NNCtxCNN



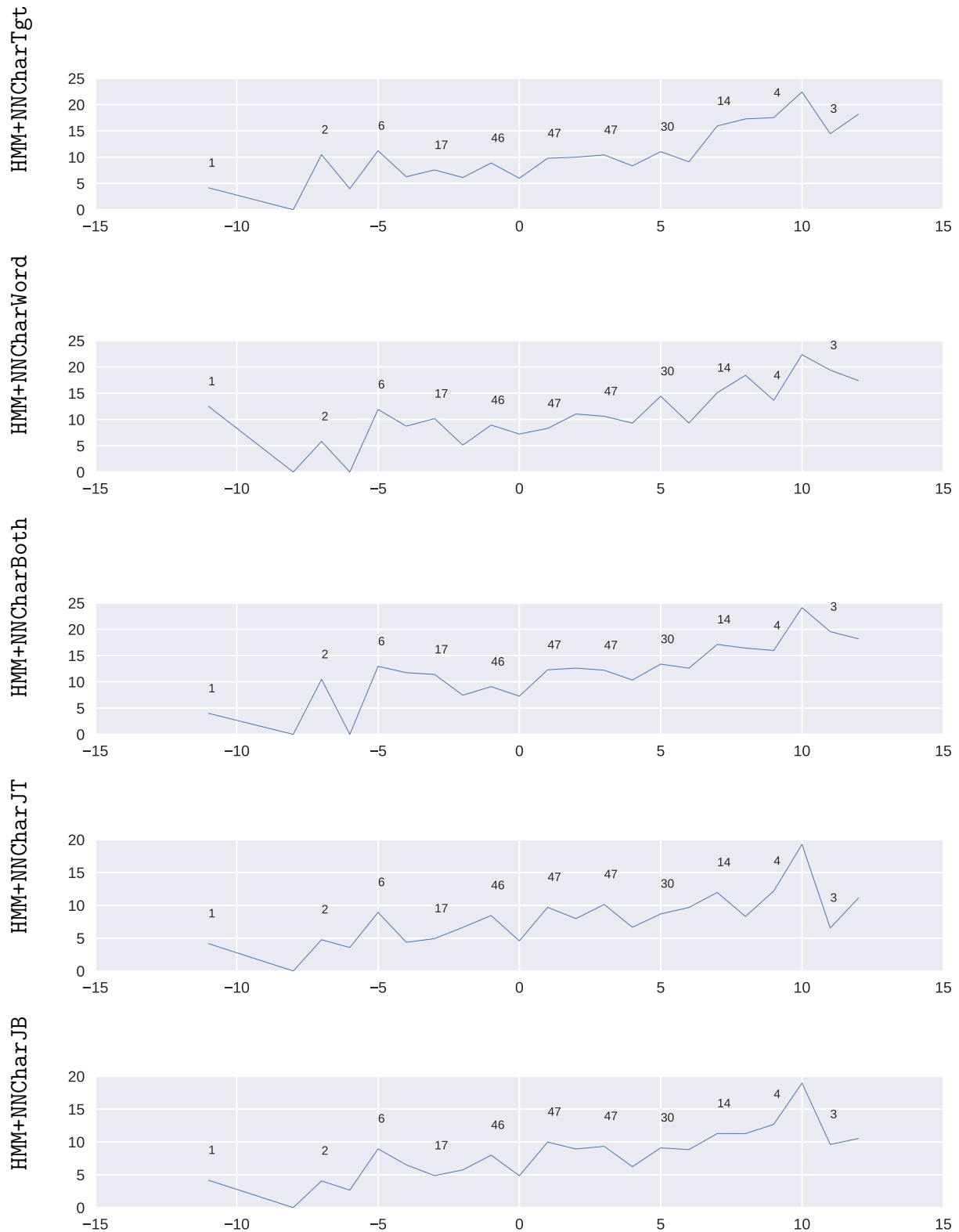
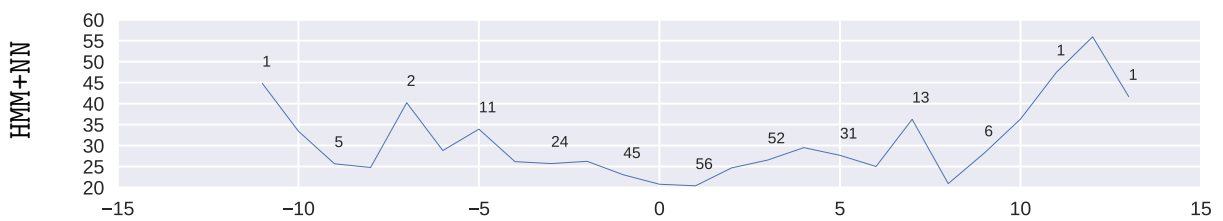
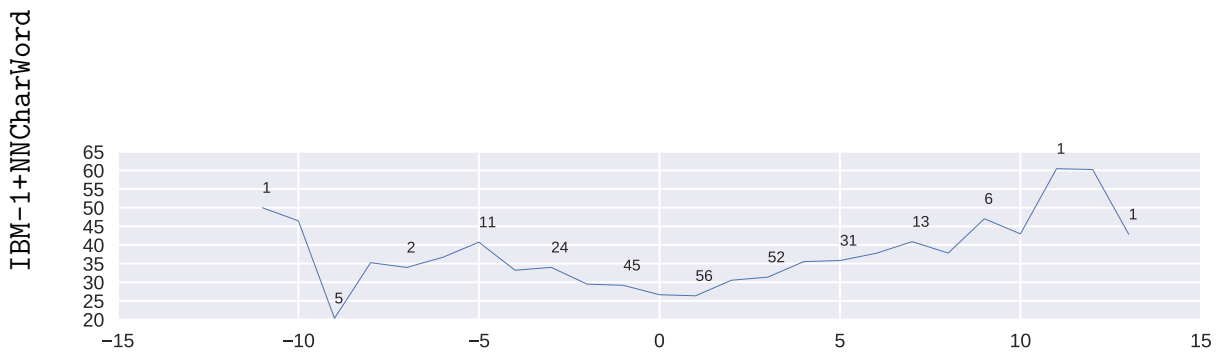
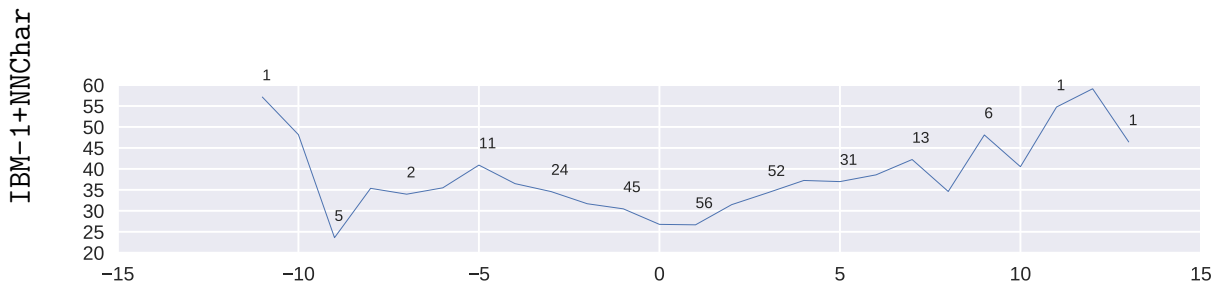
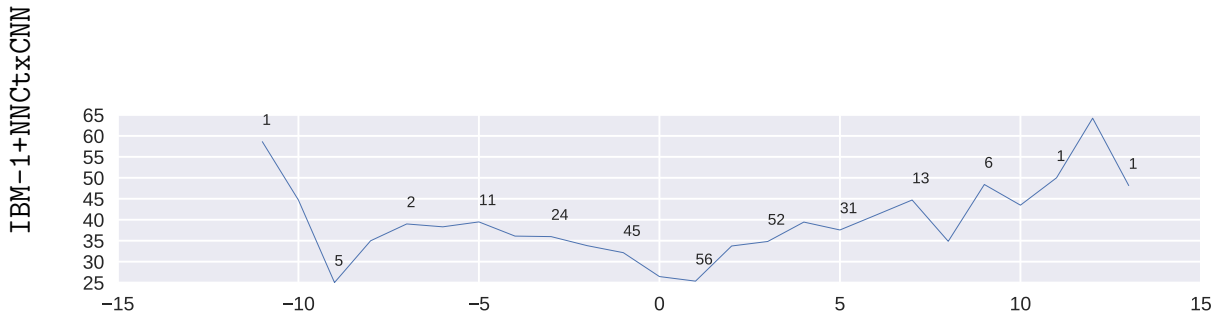
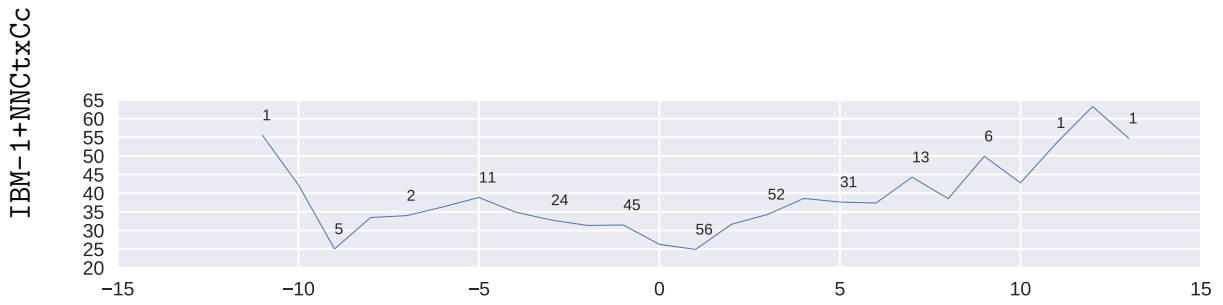
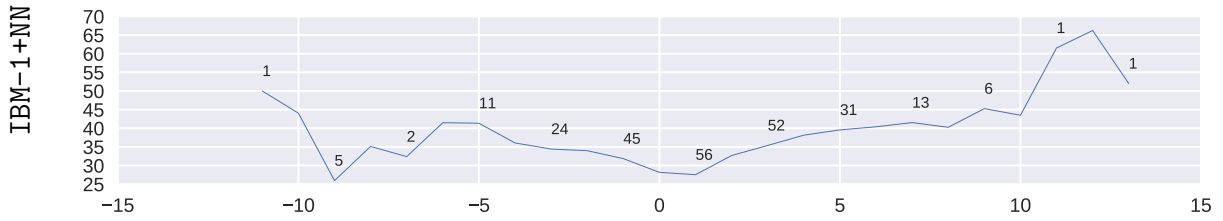
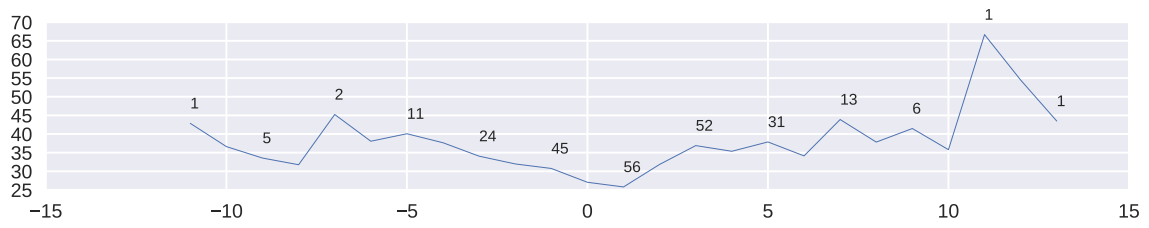


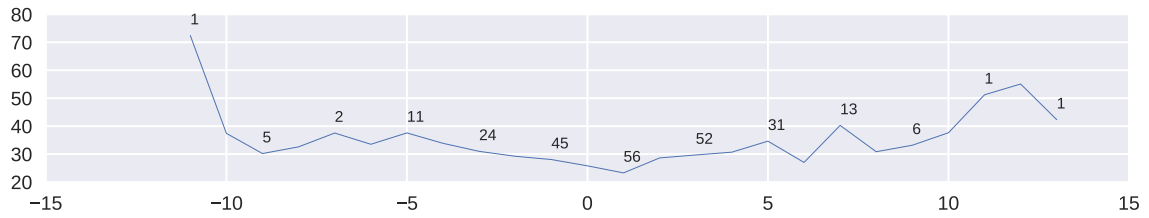
Figure B.62: Neural models for the direction French-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



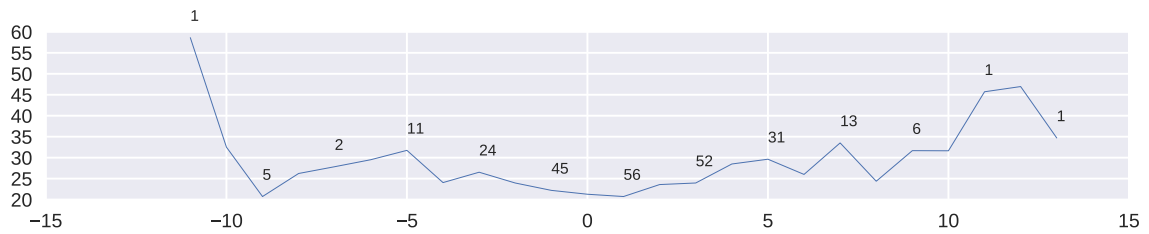
HMM+NNCtxCc



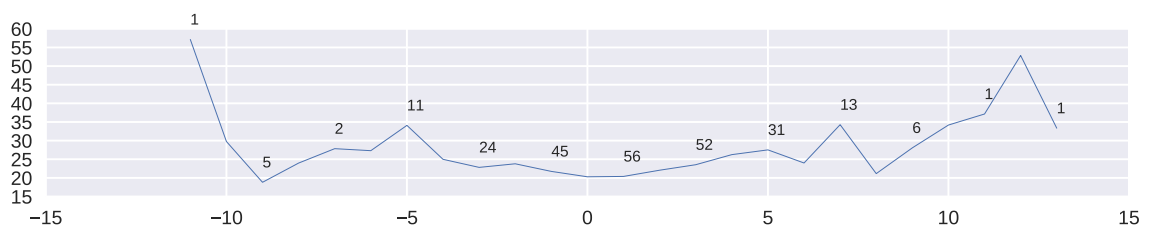
HMM+NNCtxCNN



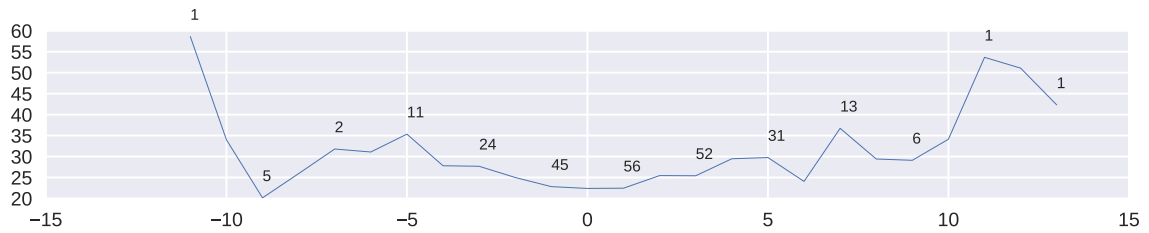
HMM+NNCharTgt



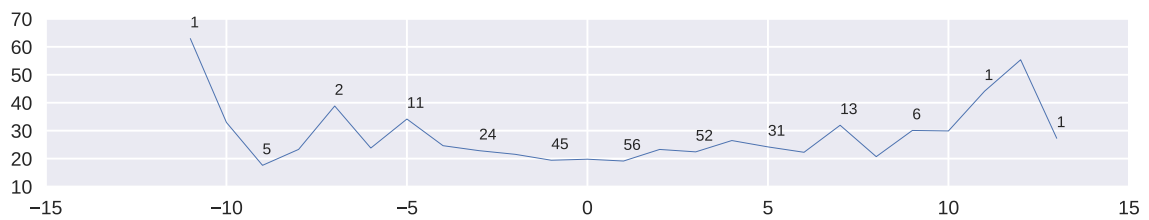
HMM+NNCharWord



HMM+NNCharBoth



HMM+NNCharJT



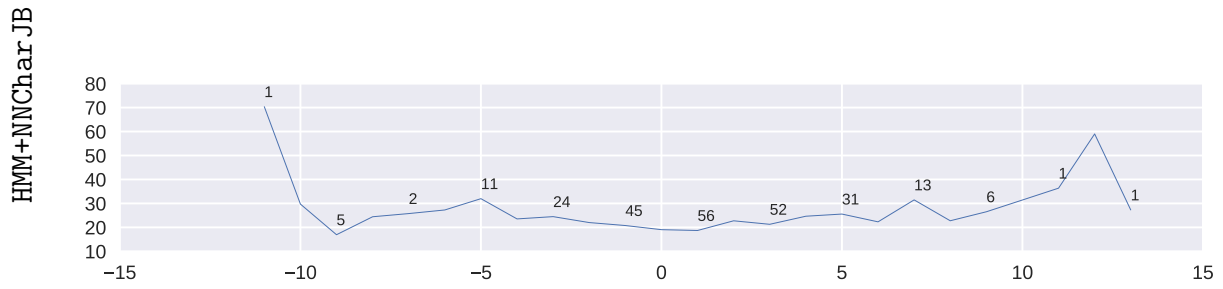
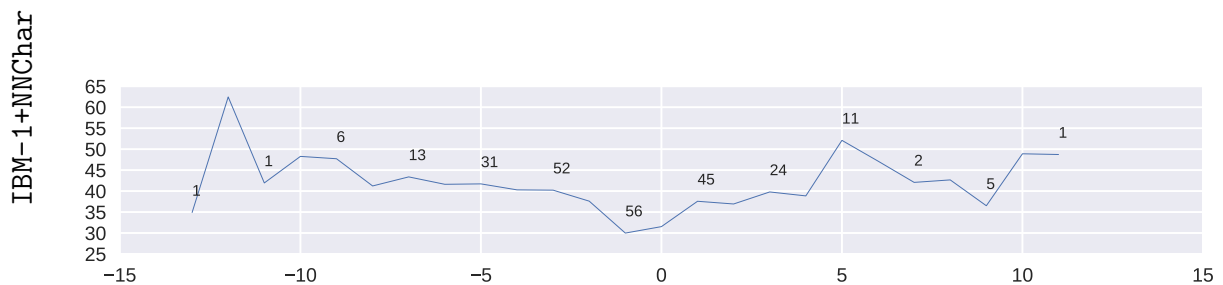
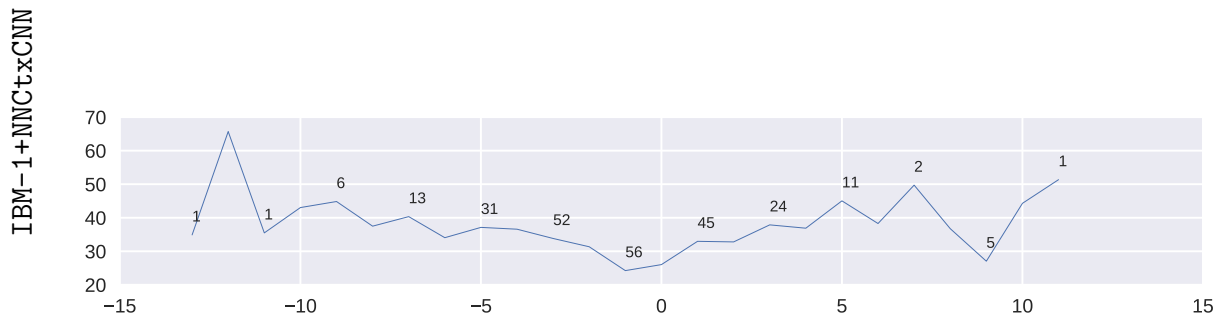
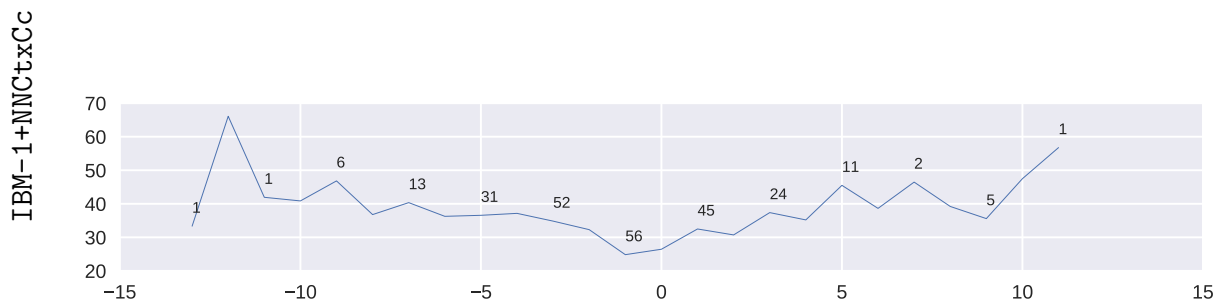
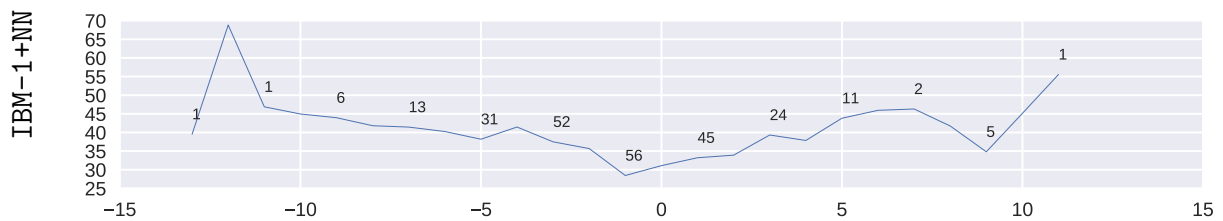
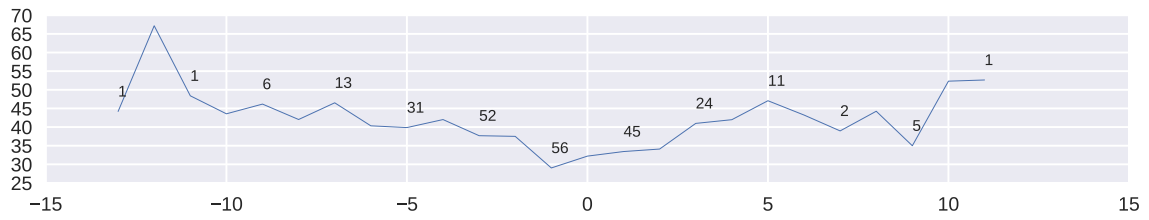


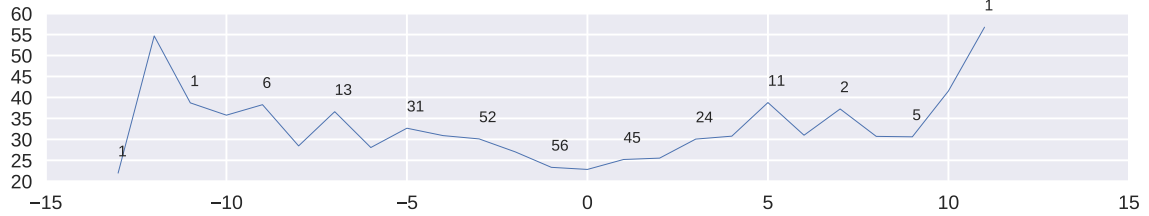
Figure B.63: Neural models for the direction English-German: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



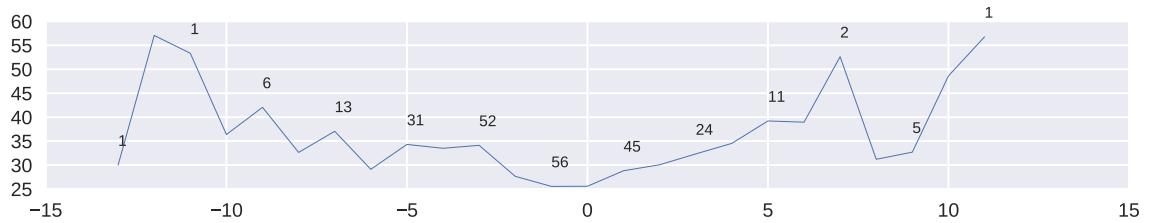
IBM-1+NNCharWord



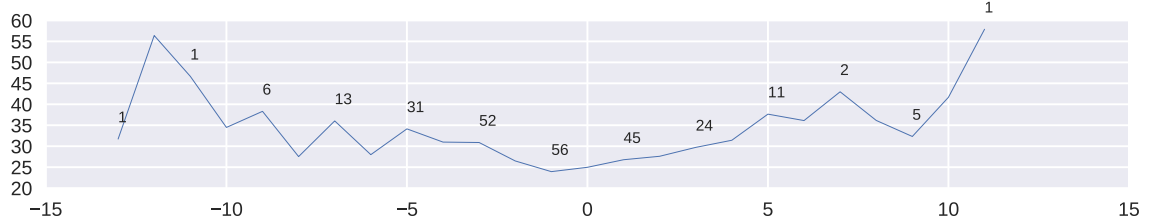
HMM+NN



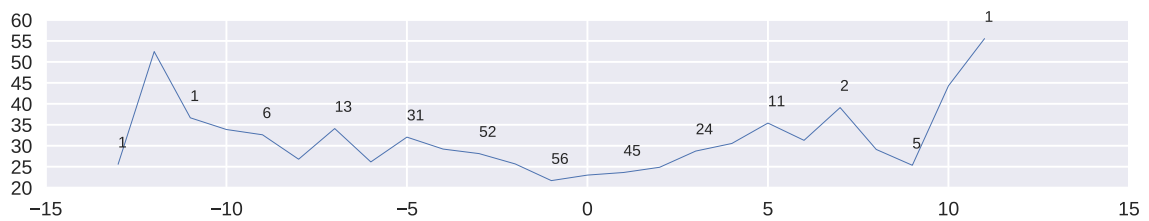
HMM+NNCtxCc



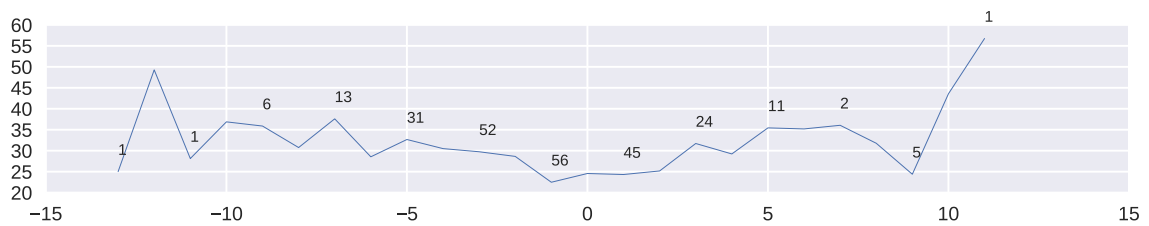
HMM+NNCtxCNN



HMM+NNCharTgt



HMM+NNCharWord



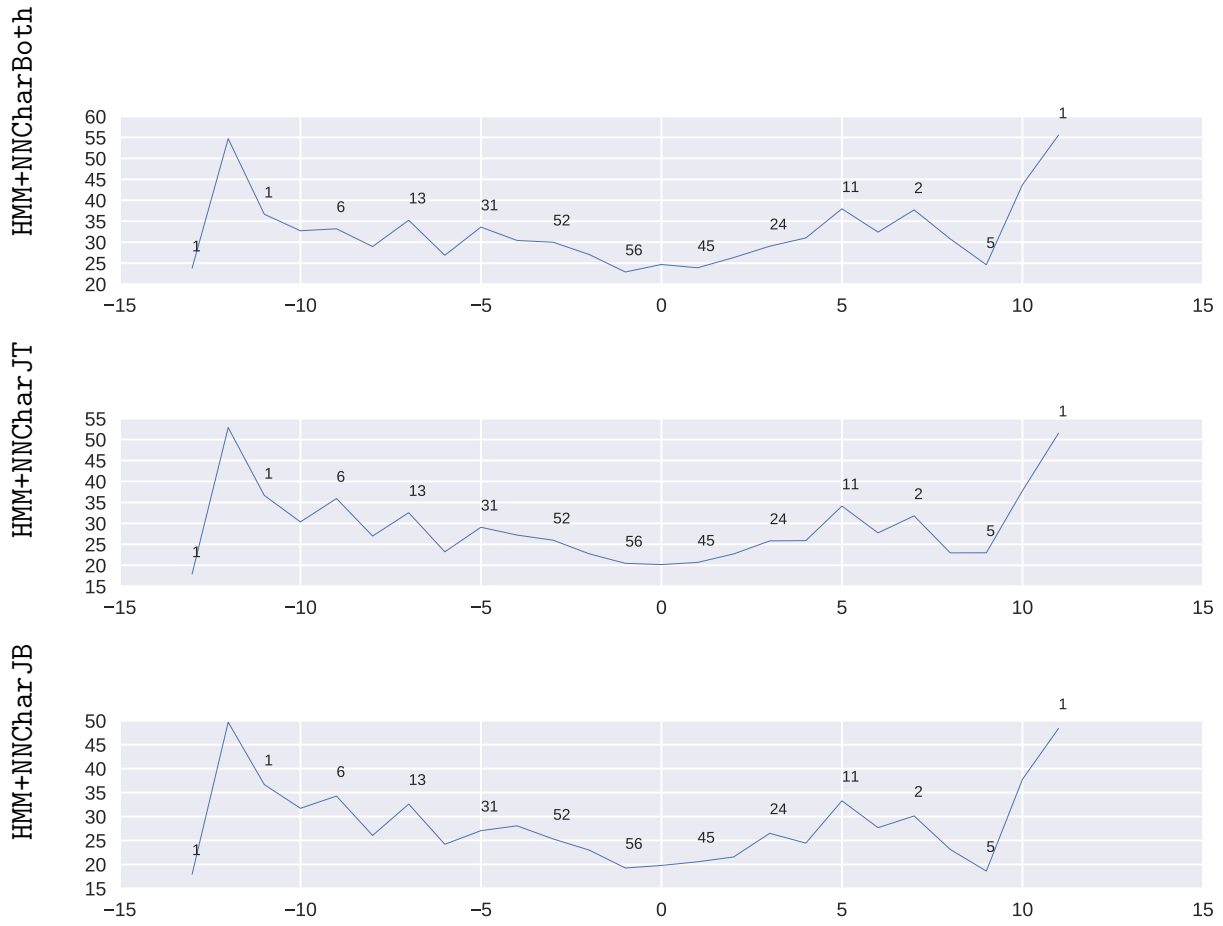
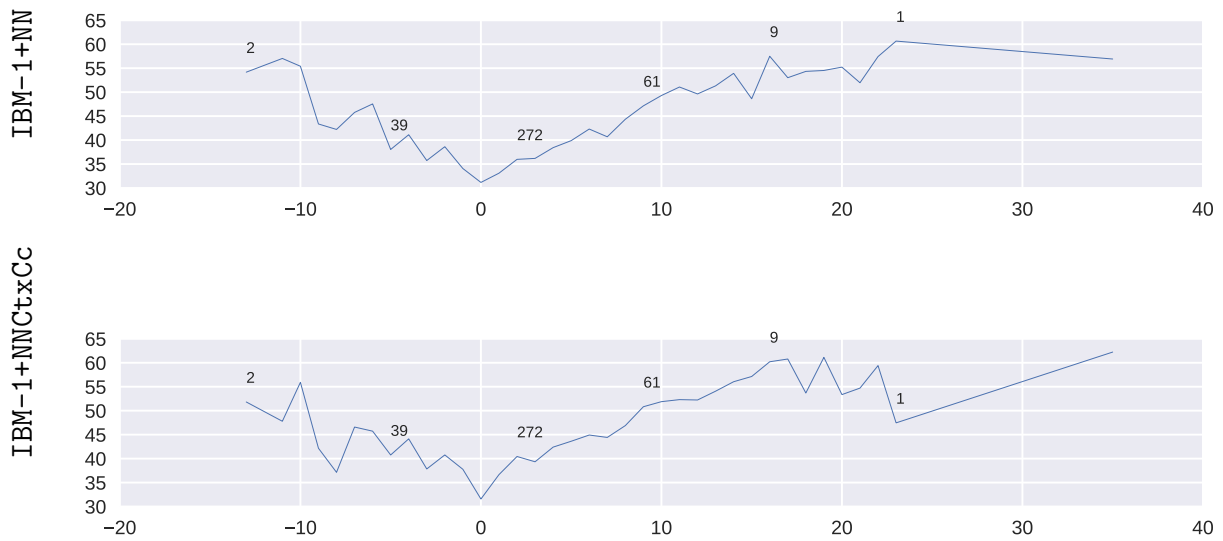
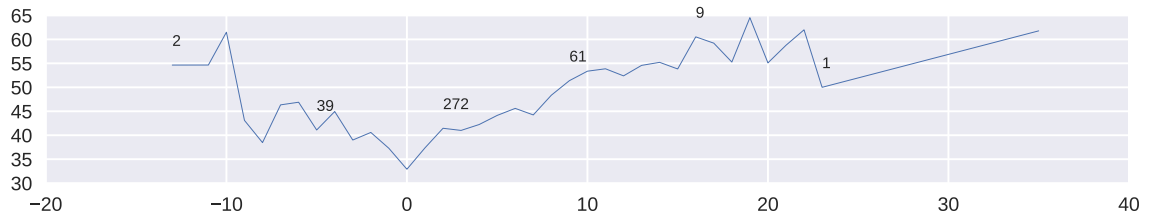


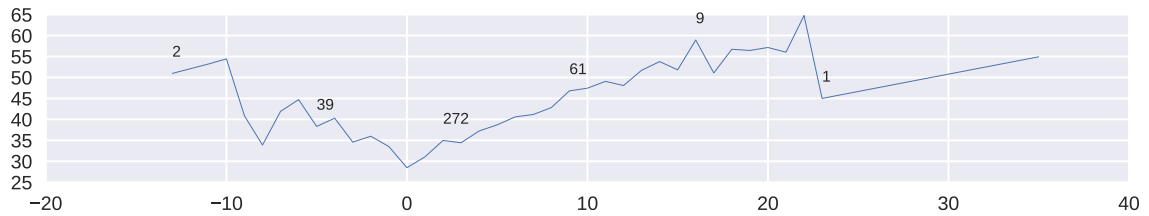
Figure B.64: Neural models for the direction German-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



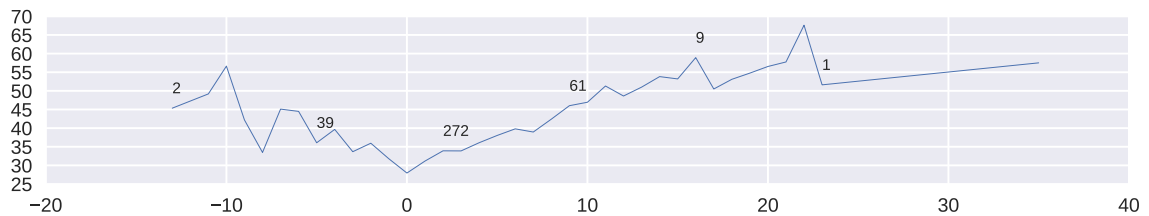
IBM-1+NNCtxCNN



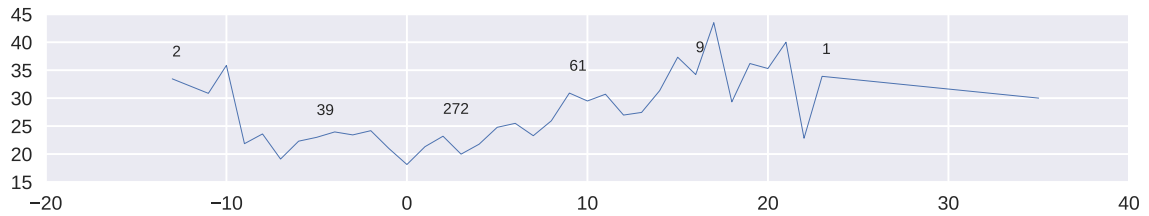
IBM-1+NNChar



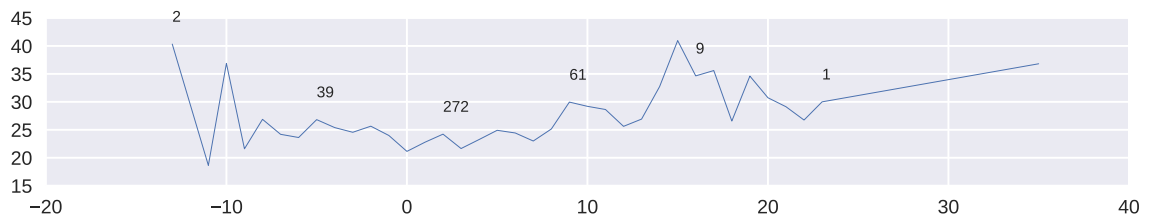
IBM-1+NNCharWord



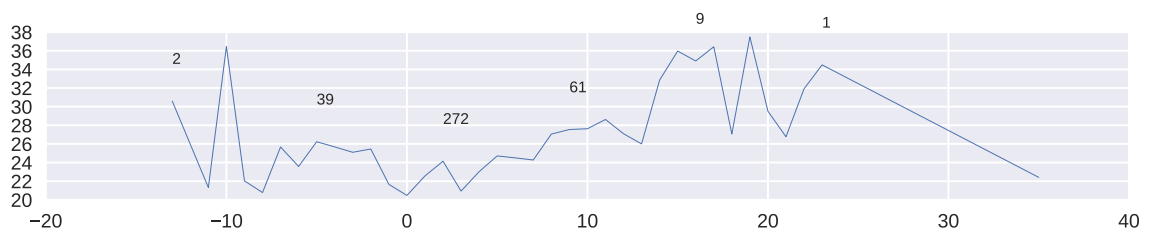
HMM+NN



HMM+NNCtxCc



HMM+NNCtxCNN



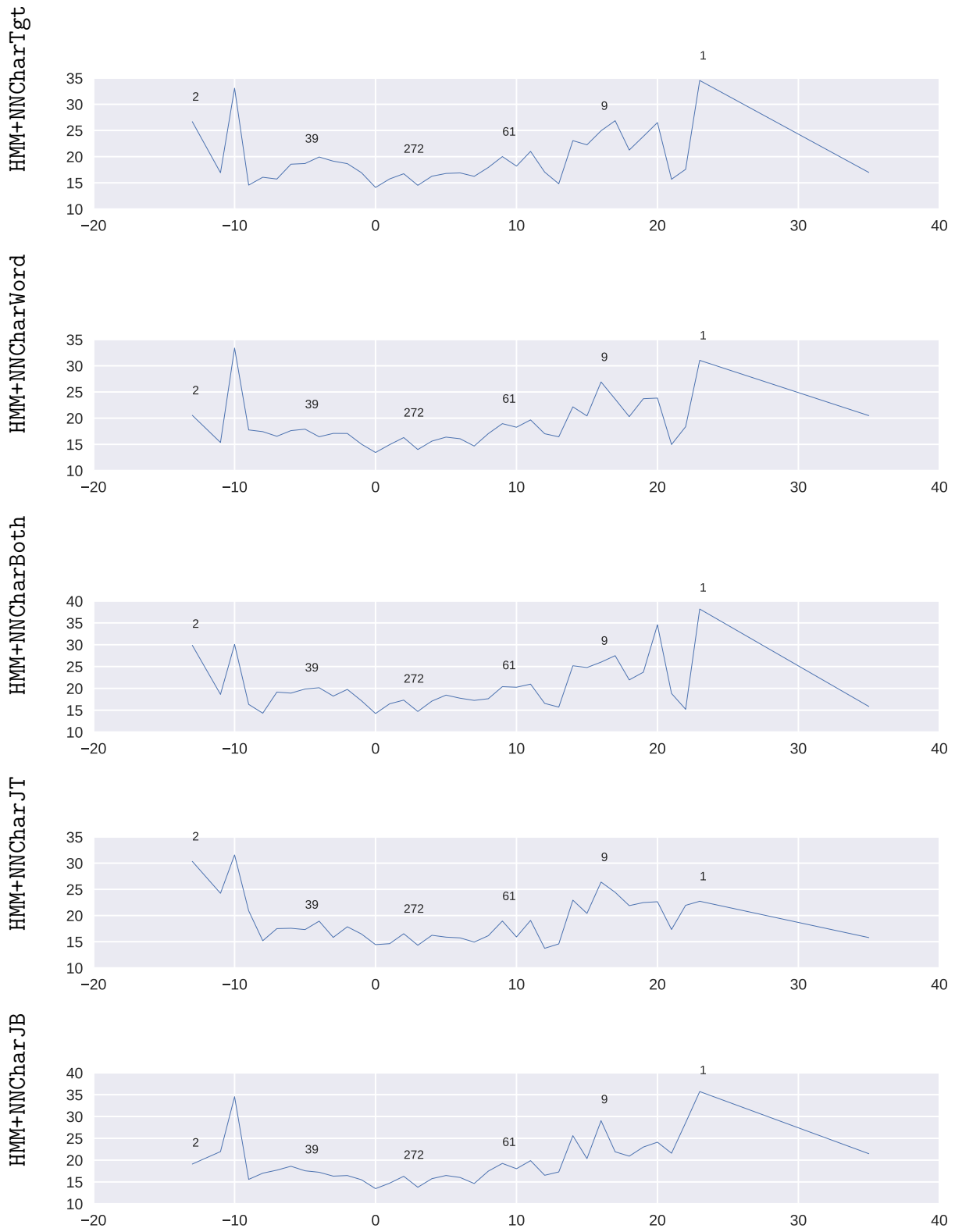
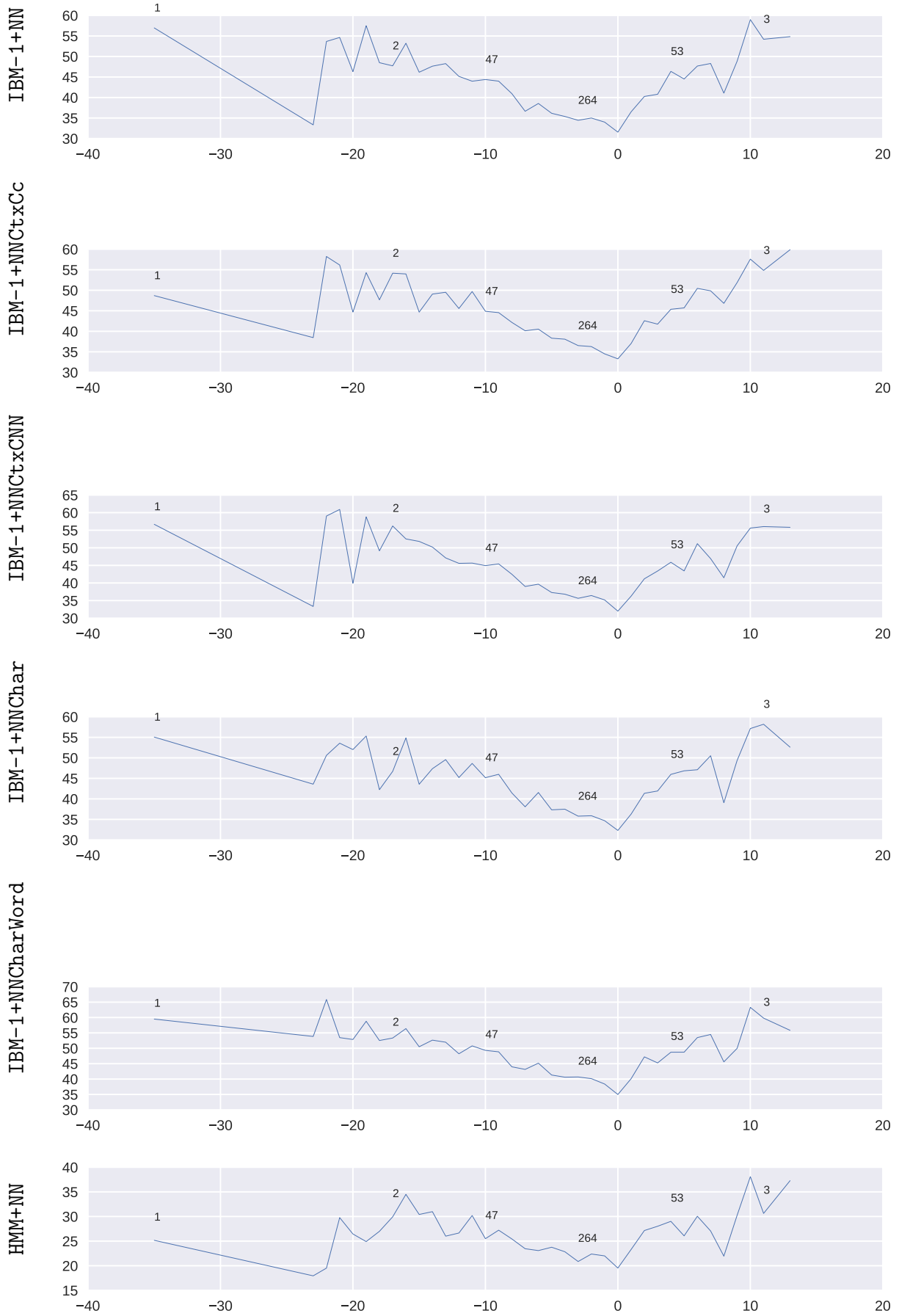
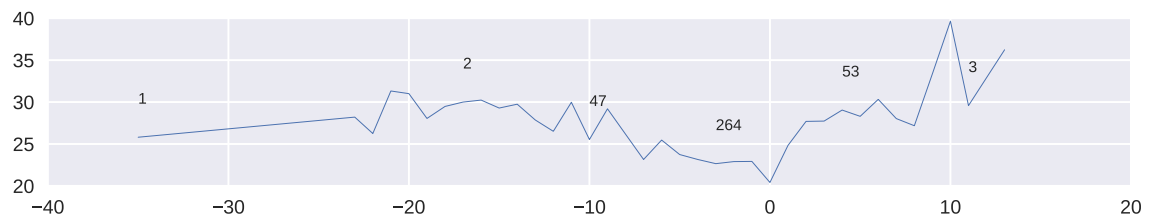


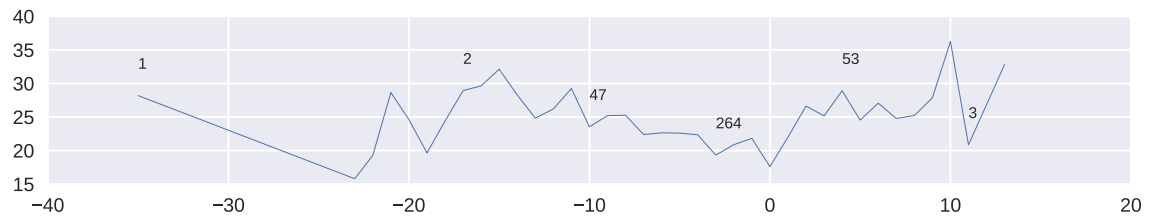
Figure B.65: Neural models for the direction English-Czech: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



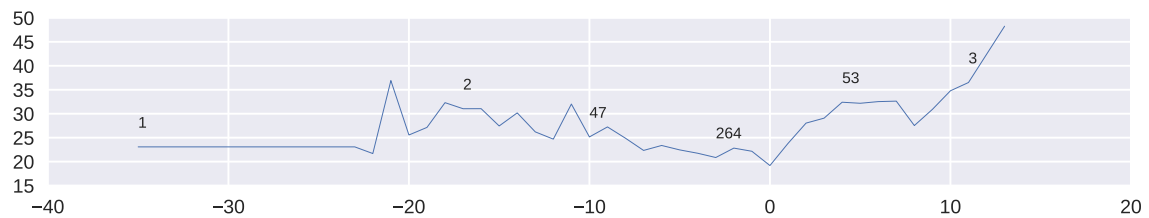
HMM+NNCtxCc



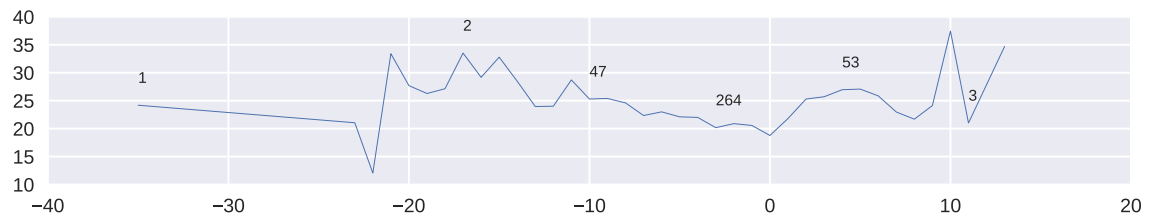
HMM+NNCtxCNN



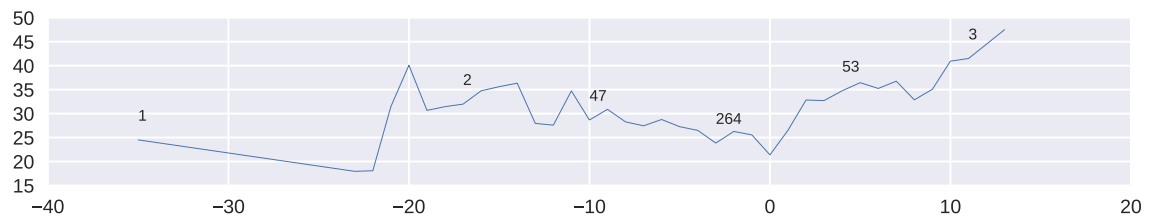
HMM+NNCharTgt



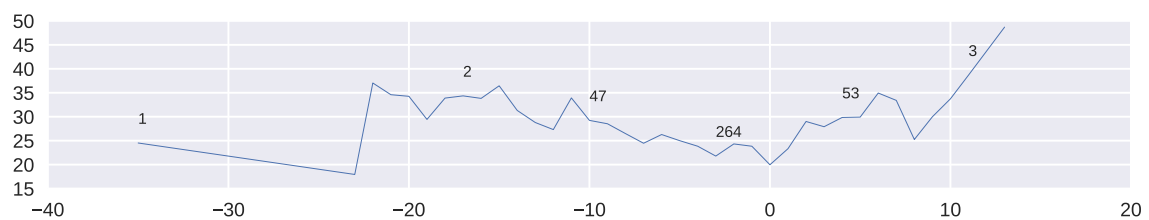
HMM+NNCharWord



HMM+NNCharBoth



HMM+NNCharJT



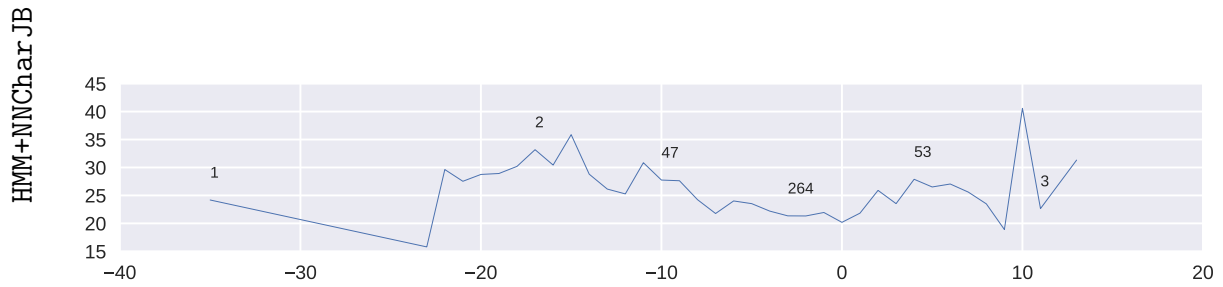
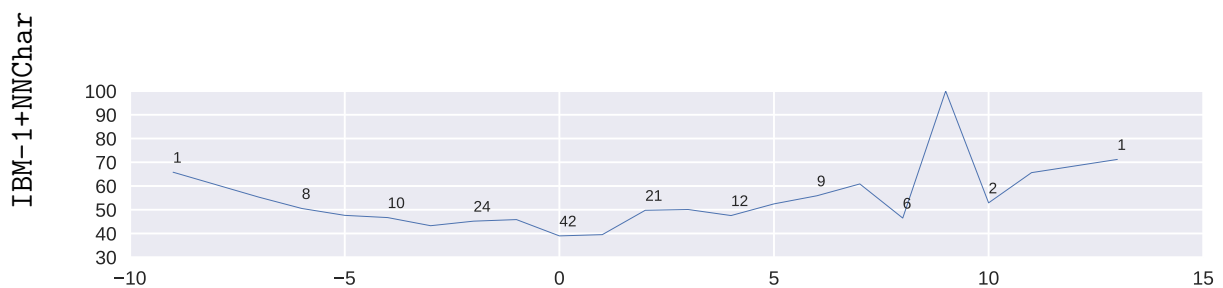
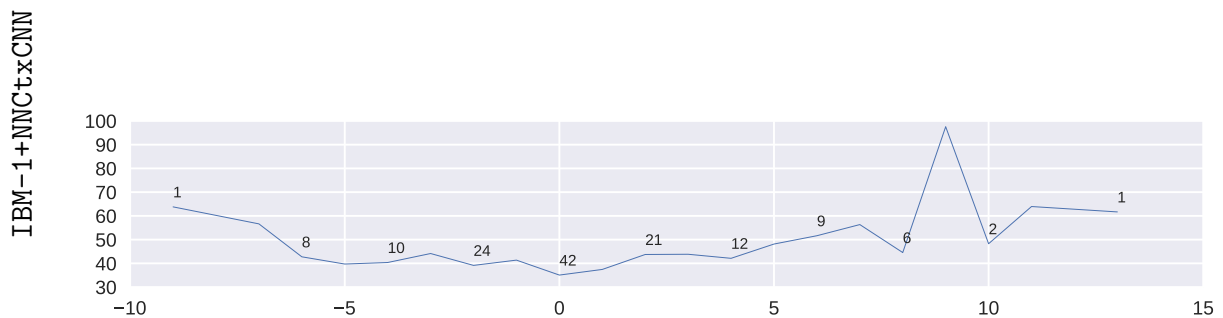
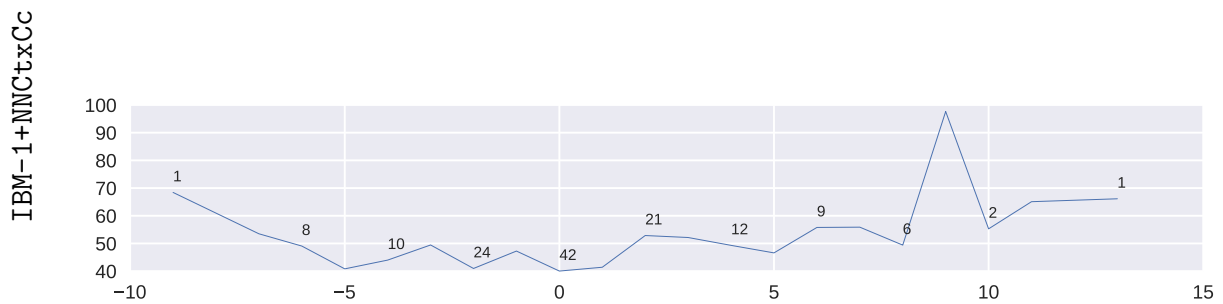
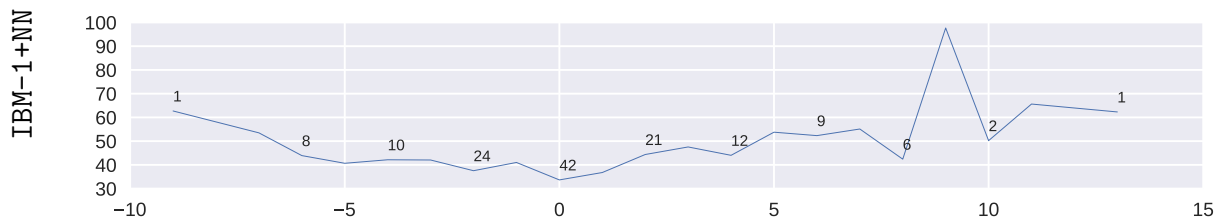
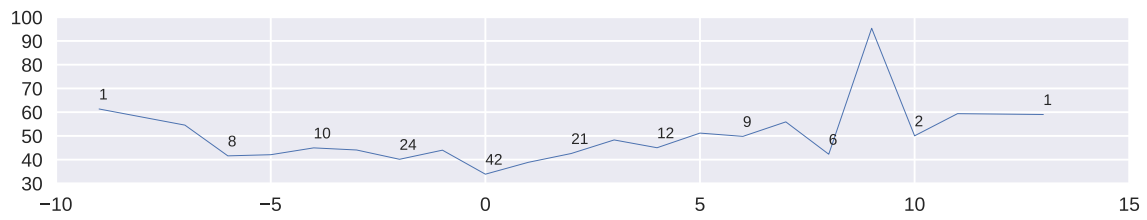


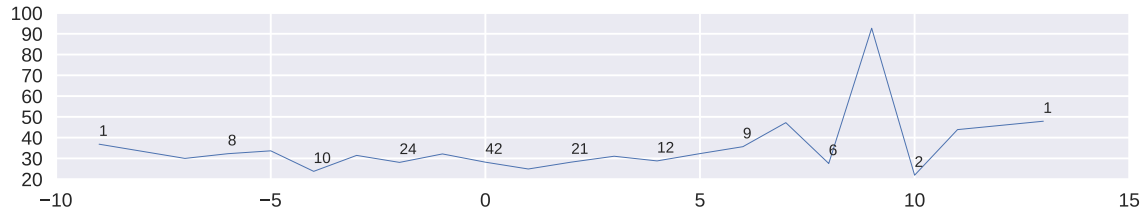
Figure B.66: Neural models for the direction Czech-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



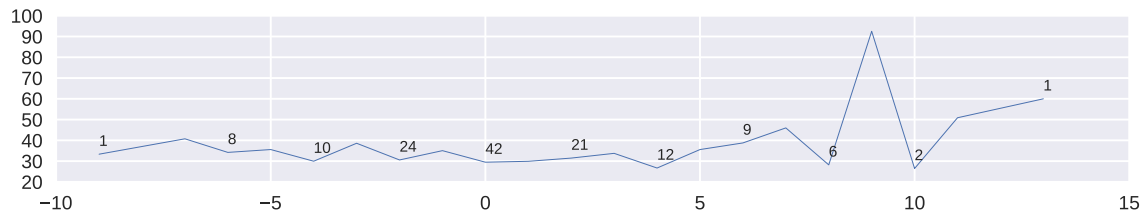
IBM-1+NNCharWord



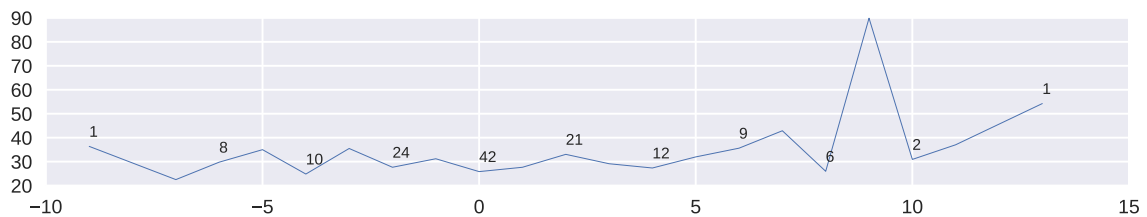
HMM+NN



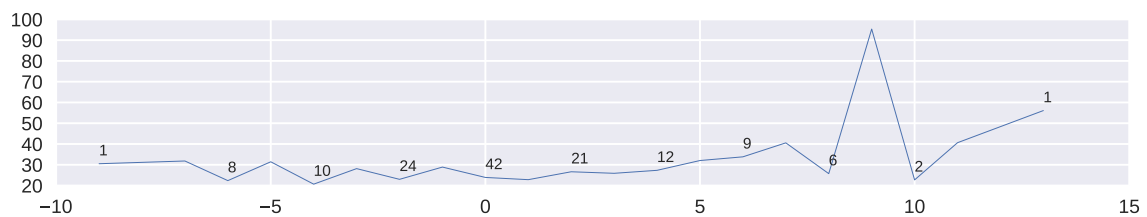
HMM+NNCtxCc



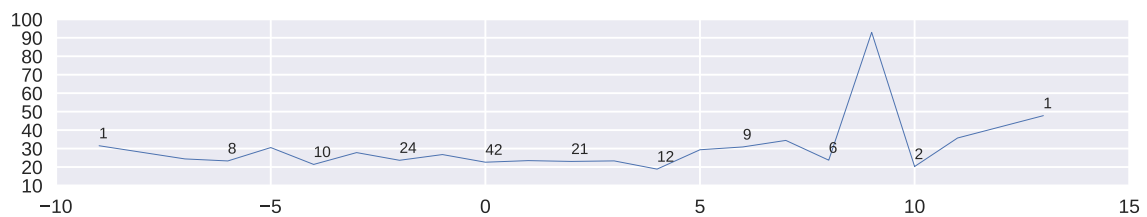
HMM+NNCtxCNN



HMM+NNCharTgt



HMM+NNCharWord



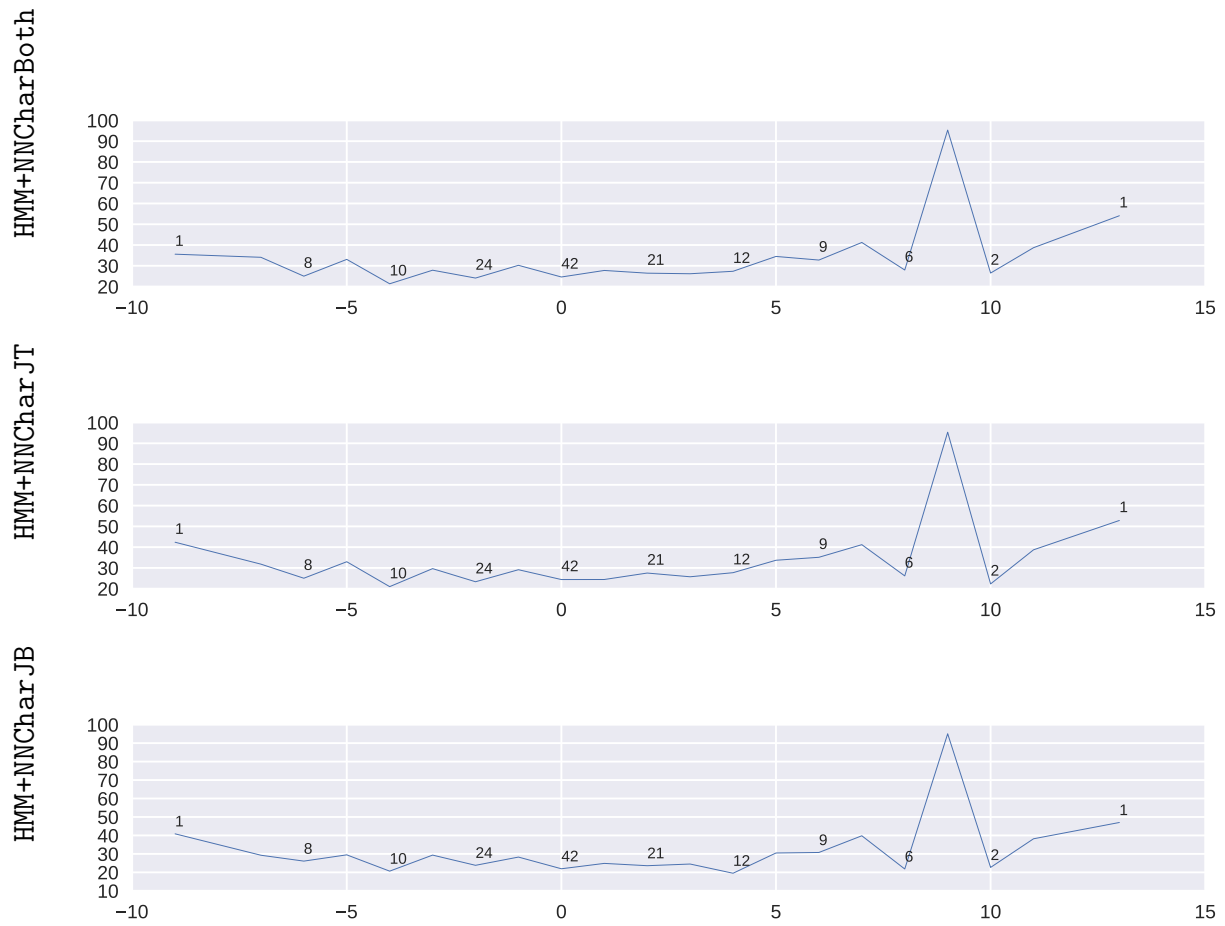
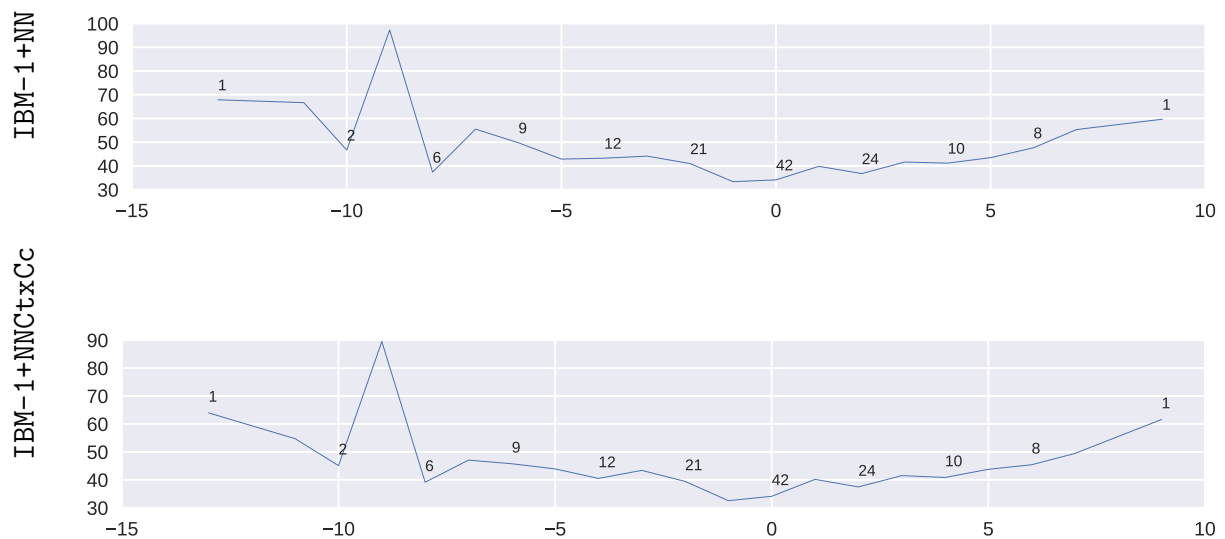
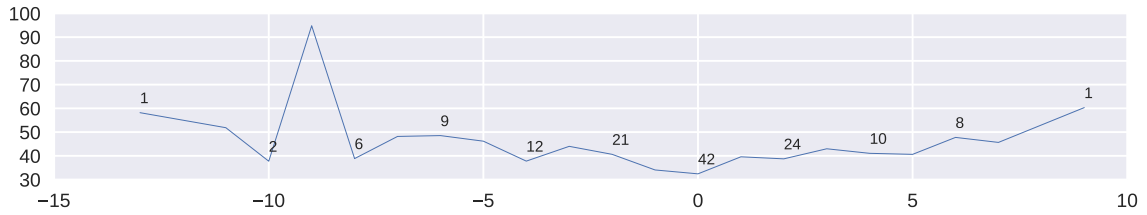


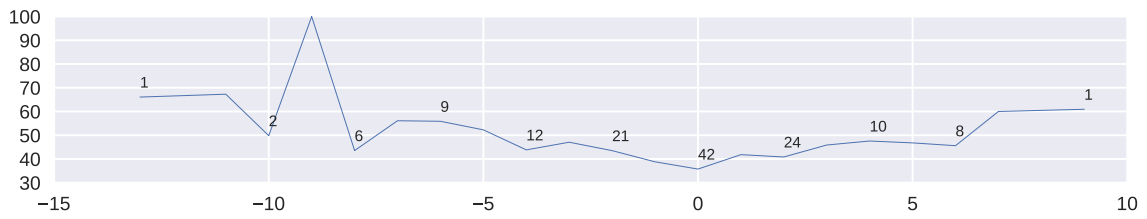
Figure B.67: Neural models for the direction English-Romanian: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



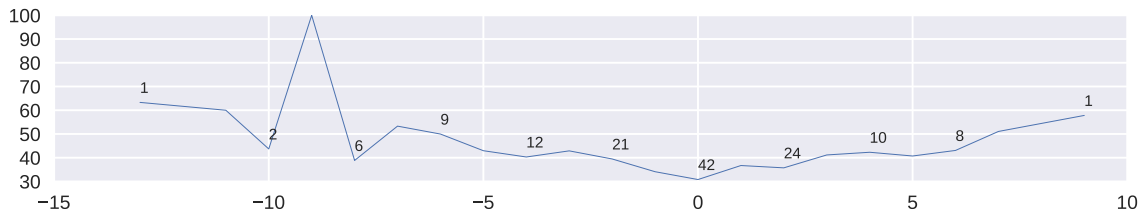
IBM-1+NNctxCNN



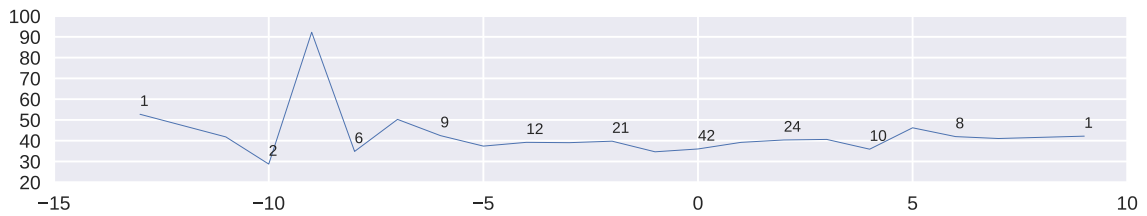
IBM-1+NNChar



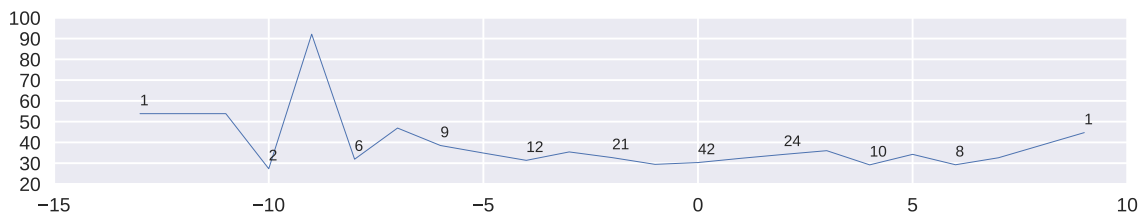
IBM-1+NNCharWord



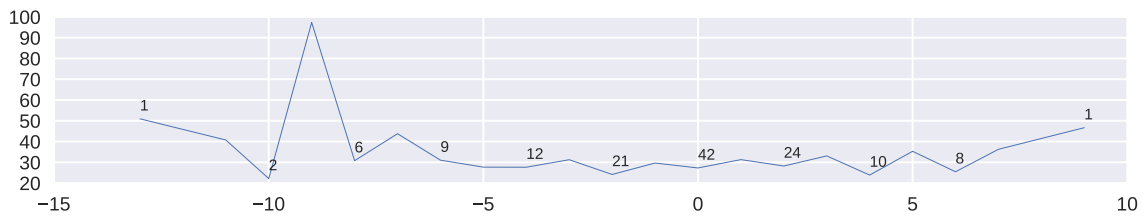
HMM+NN



HMM+NNctxCc



HMM+NNctxCNN



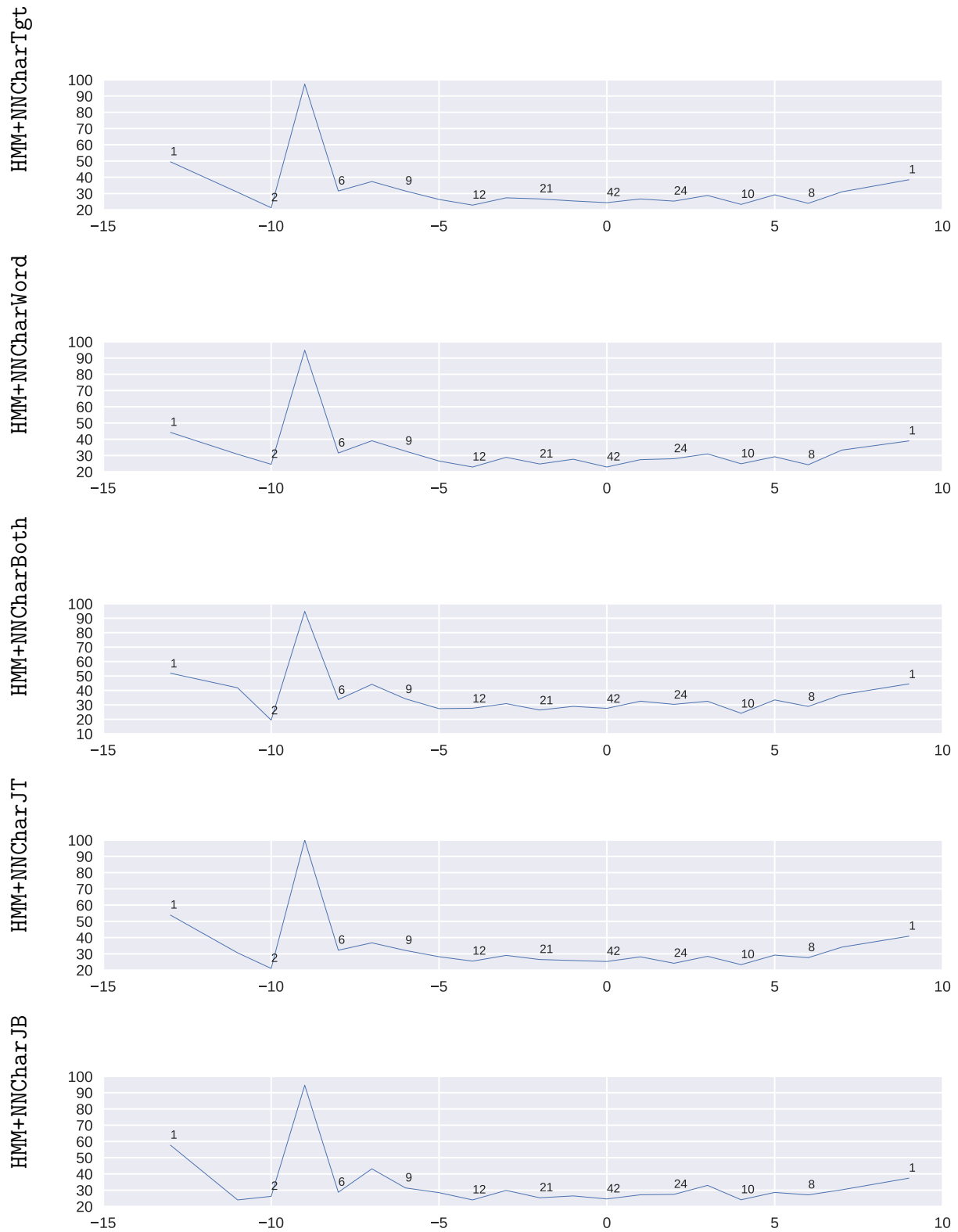
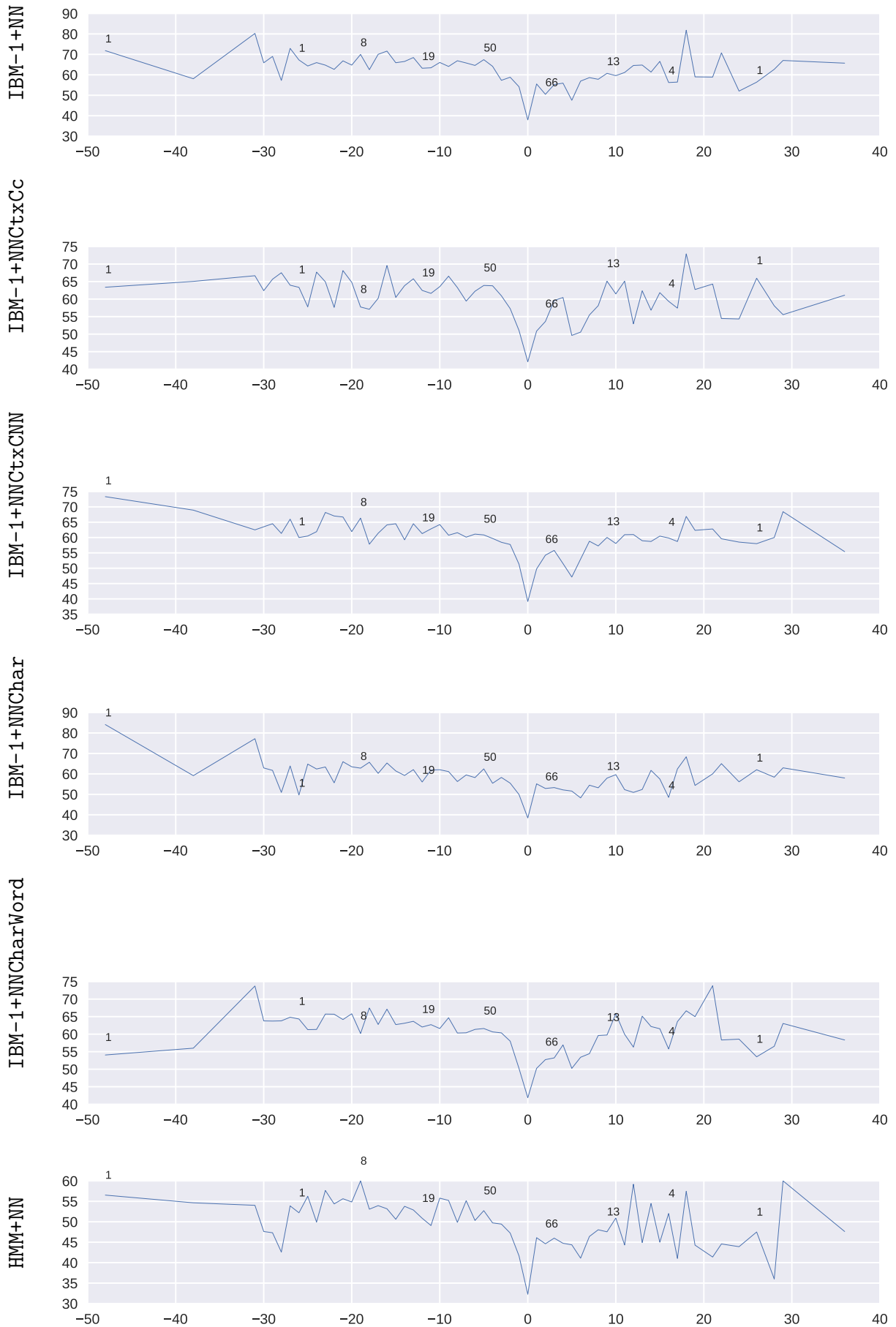
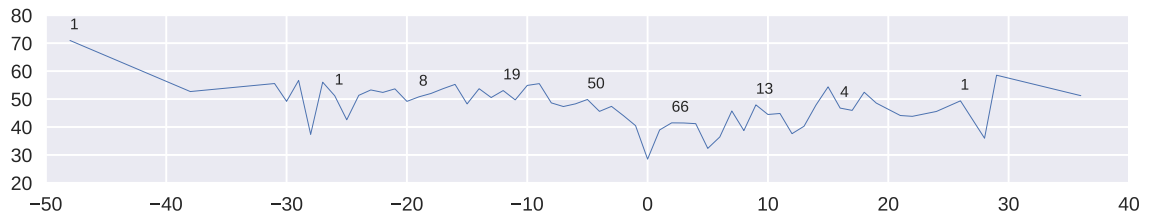


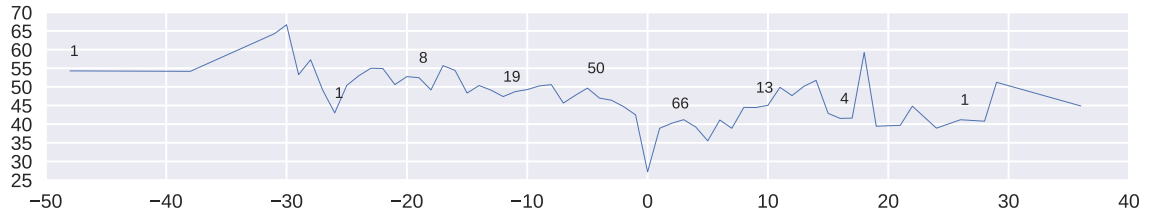
Figure B.68: Neural models for the direction Romanian-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



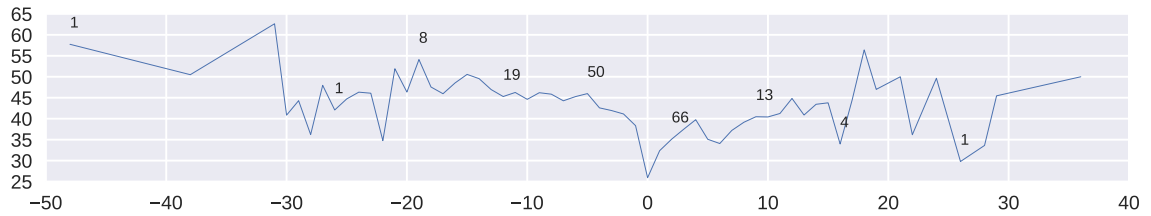
HMM+NNCtxCc



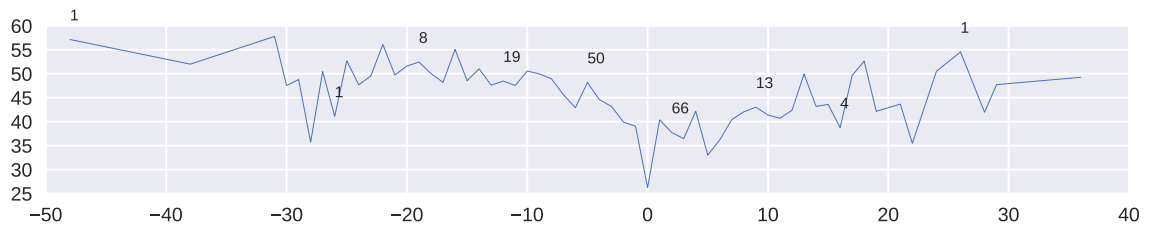
HMM+NNCtxCNN



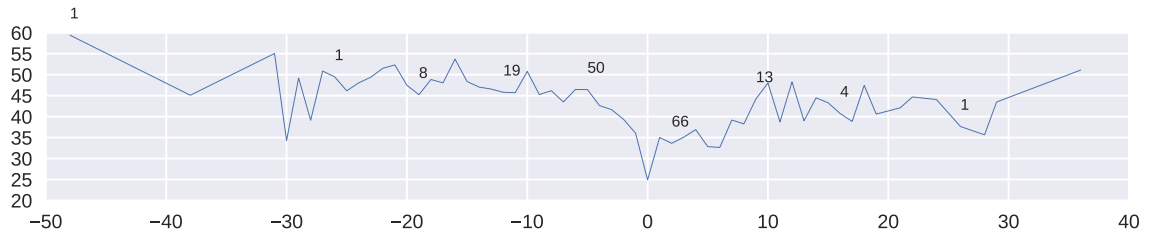
HMM+NNCharTgt



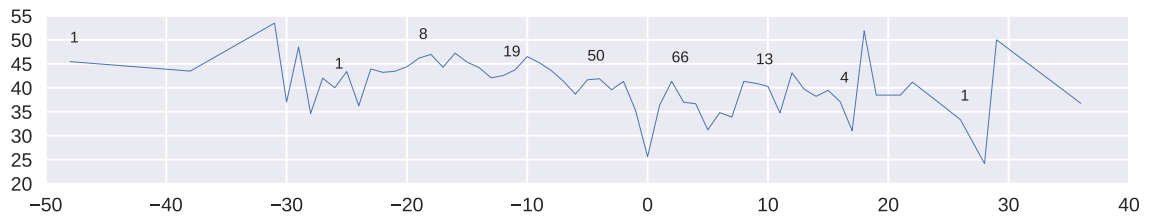
HMM+NNCharWord



HMM+NNCharBoth



HMM+NNCharJT



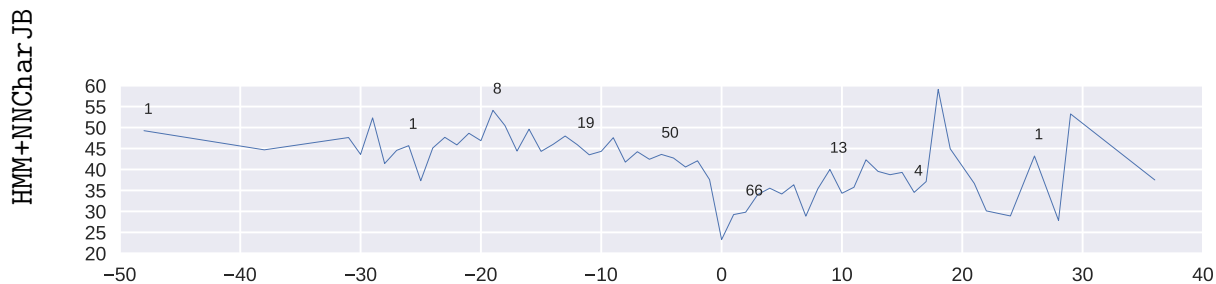
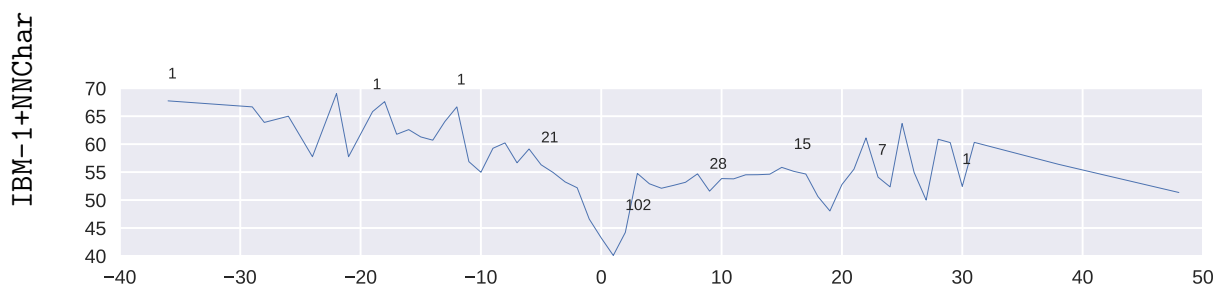
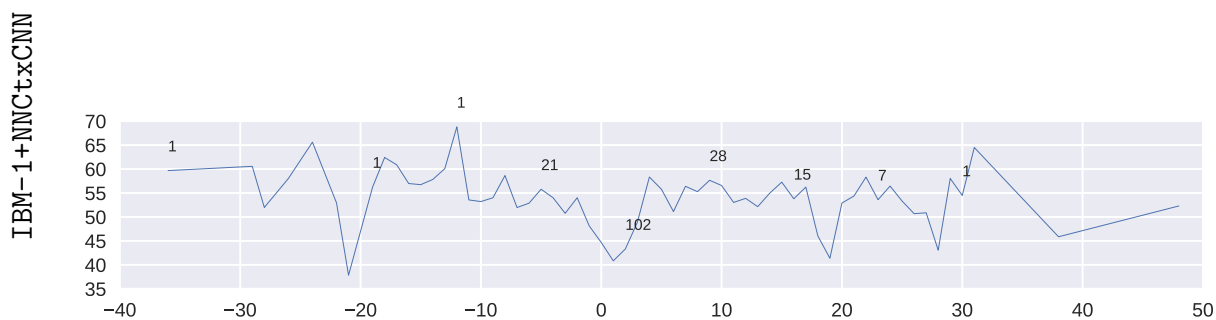
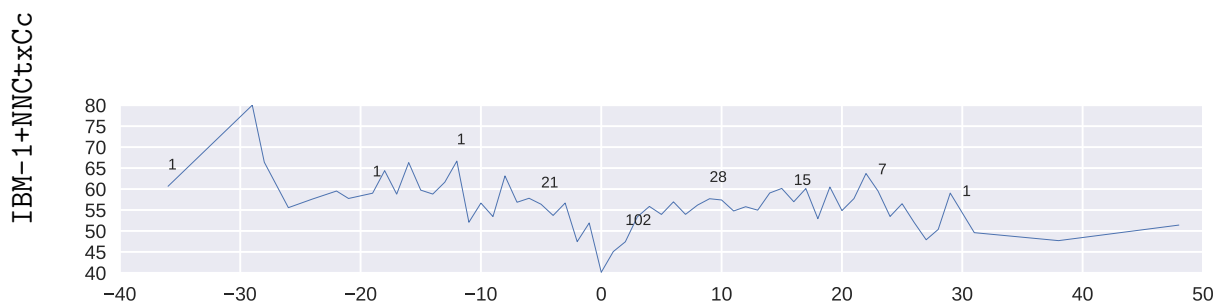
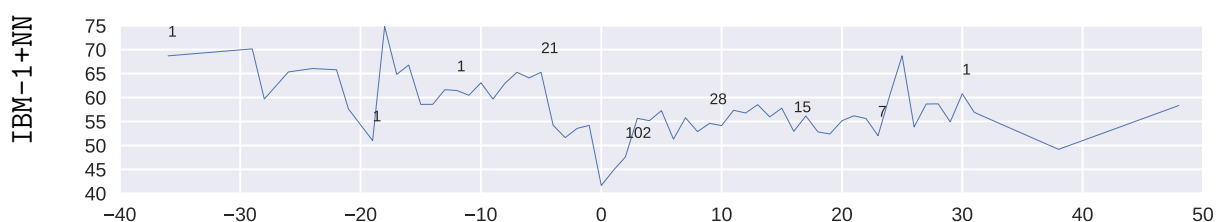
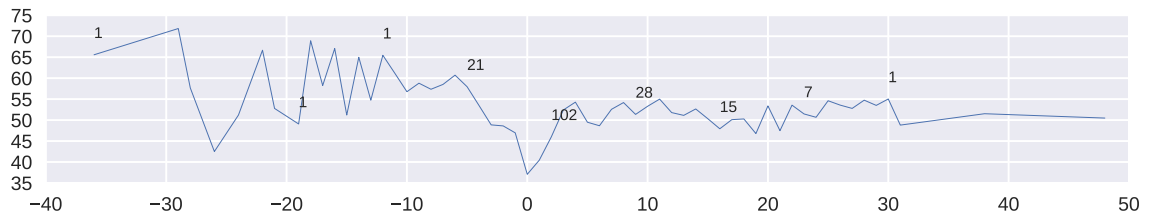


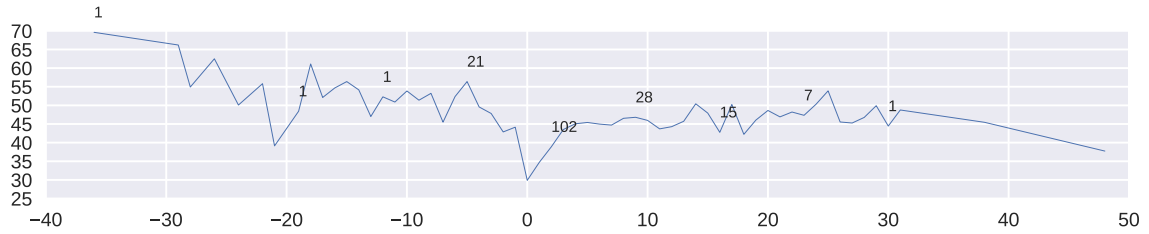
Figure B.69: Neural models for the direction English-Japanese: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



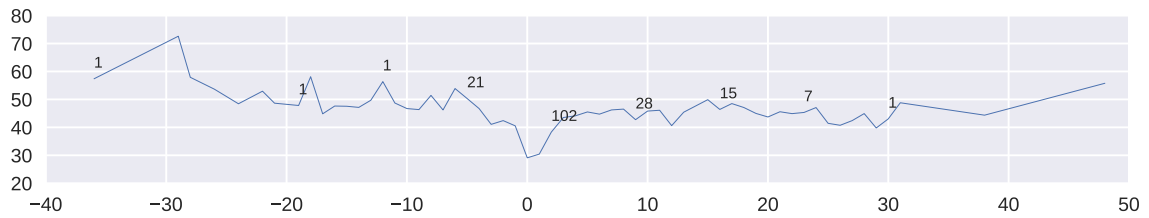
IBM-1+NNCharWord



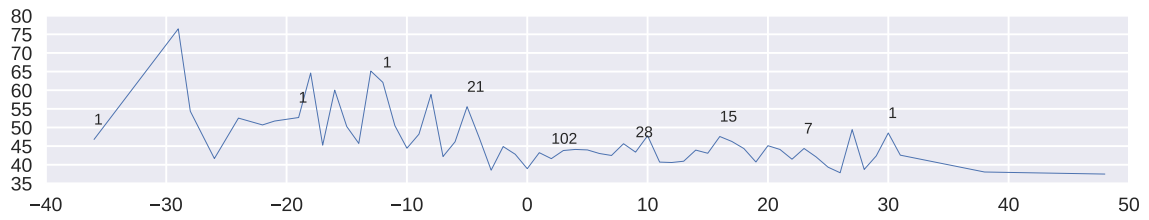
HMM+NN



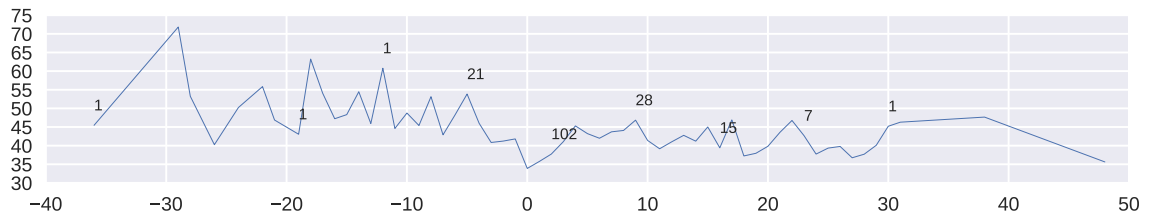
HMM+NNCtxCc



HMM+NNCtxCNN



HMM+NNCharTgt



HMM+NNCharWord

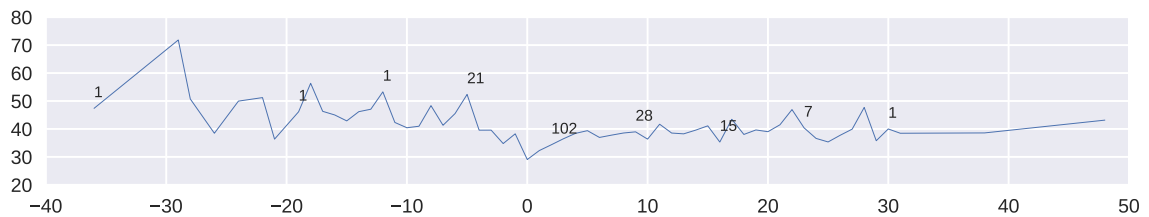
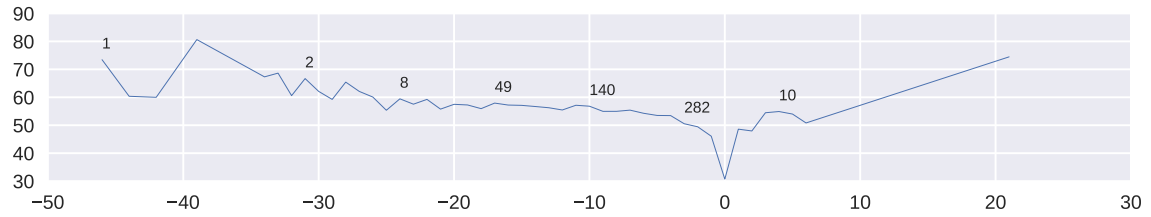




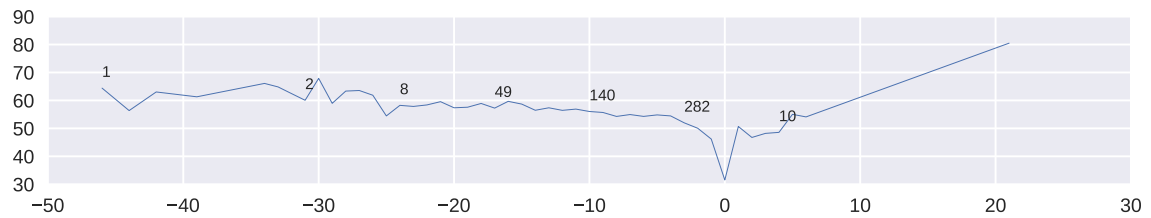
Figure B.70: Neural models for the direction Japanese-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



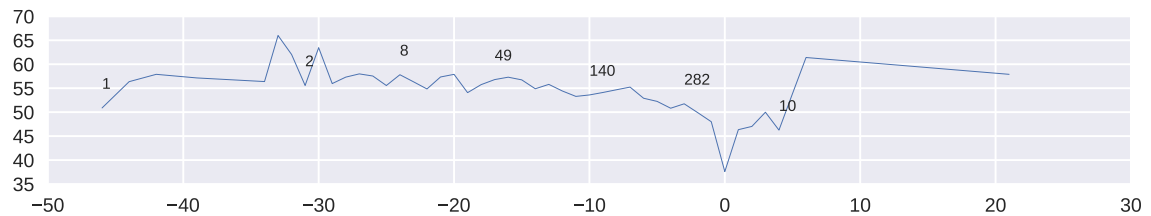
IBM-1+NNctxCNN



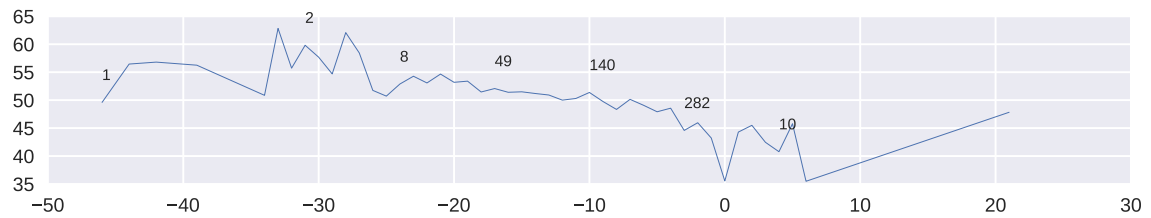
IBM-1+NNChar



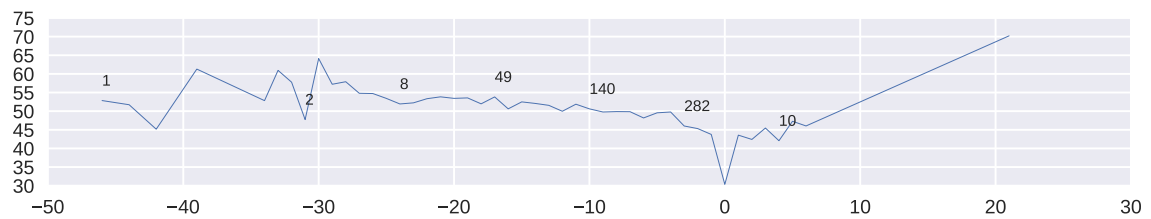
IBM-1+NNCharWord



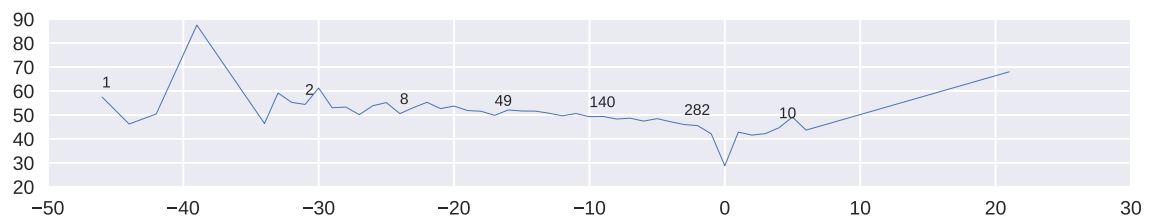
HMM+NN



HMM+NNctxCc



HMM+NNctxCNN



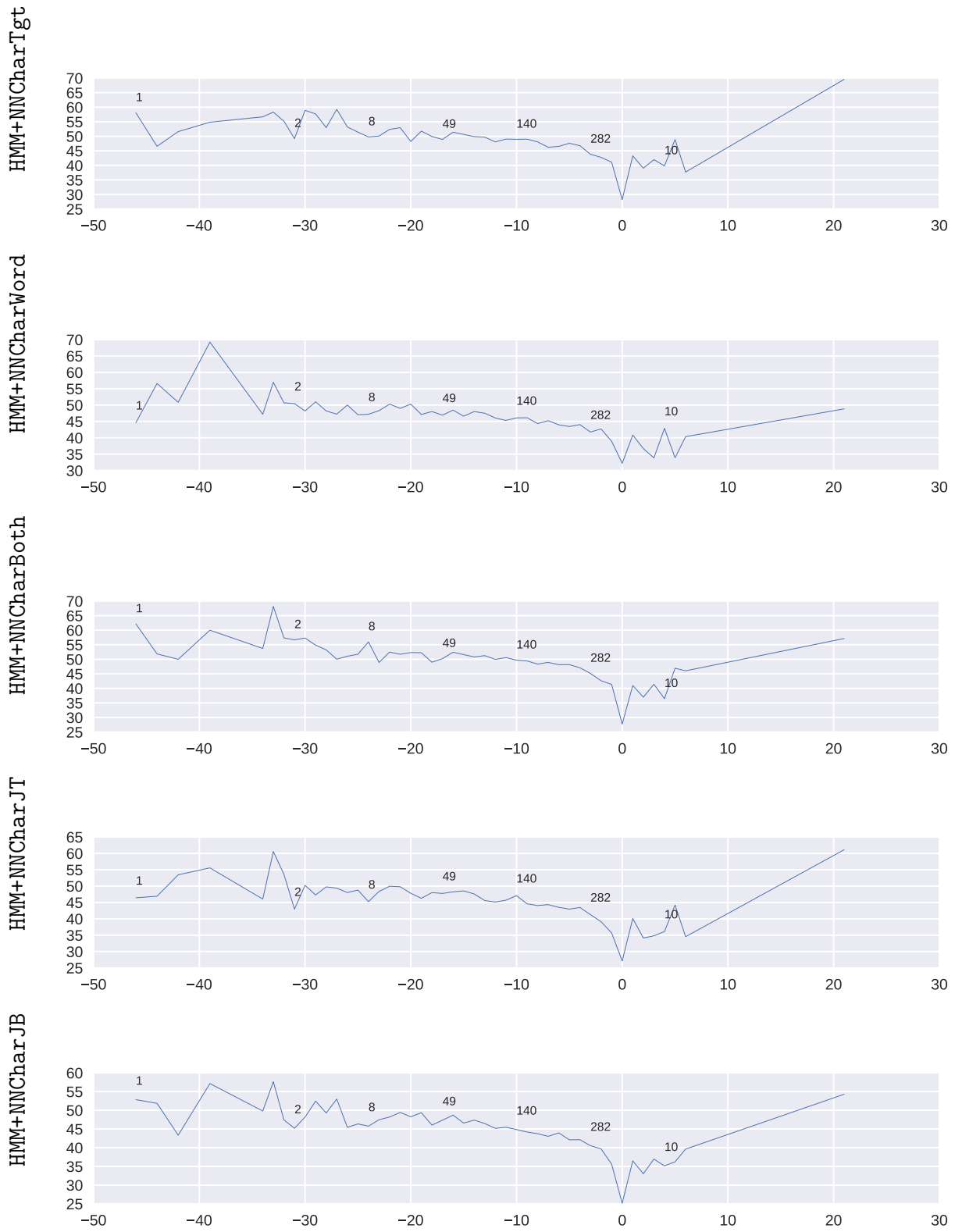
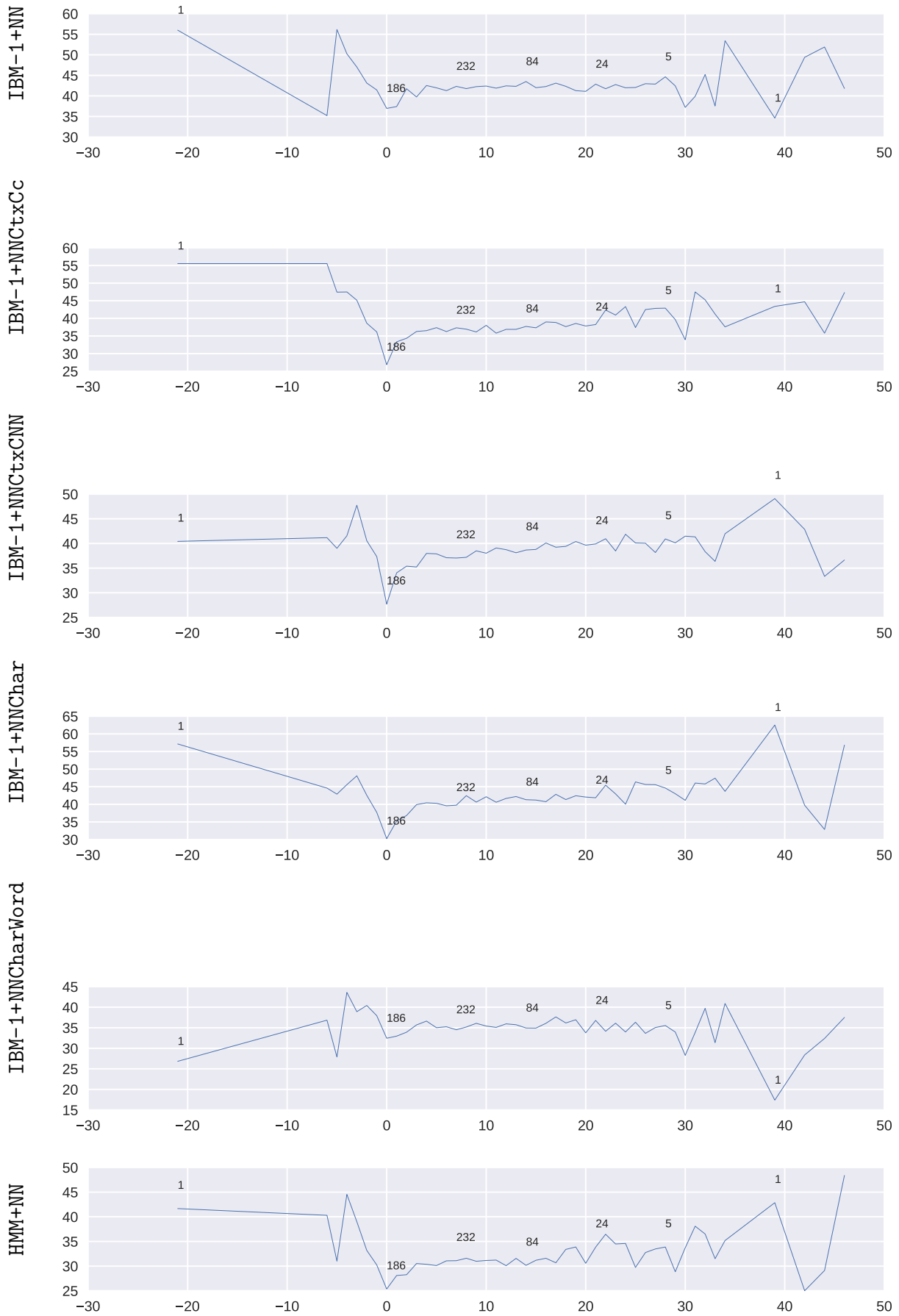
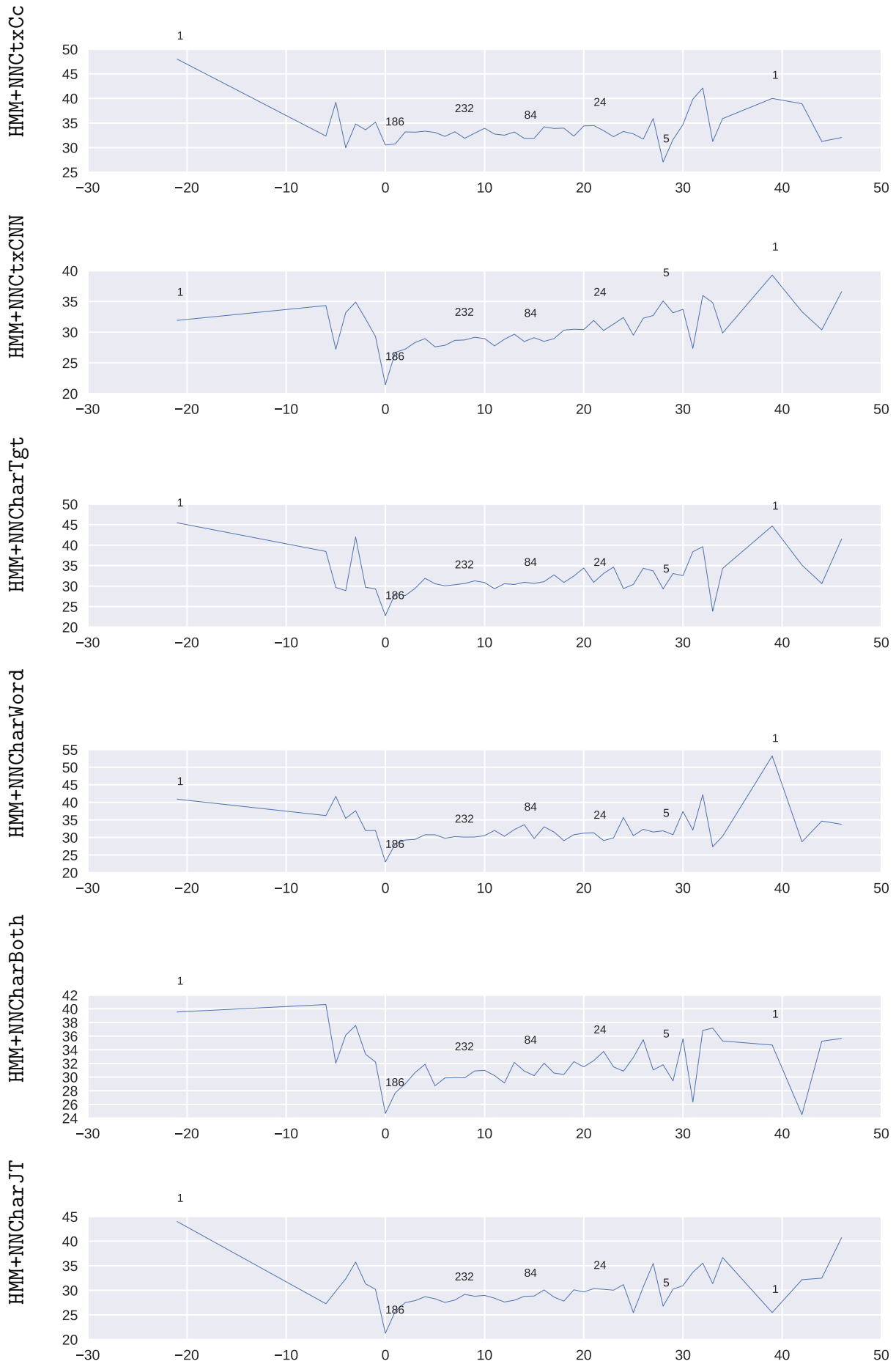


Figure B.71: Neural models for the direction English-Vietnamese: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.





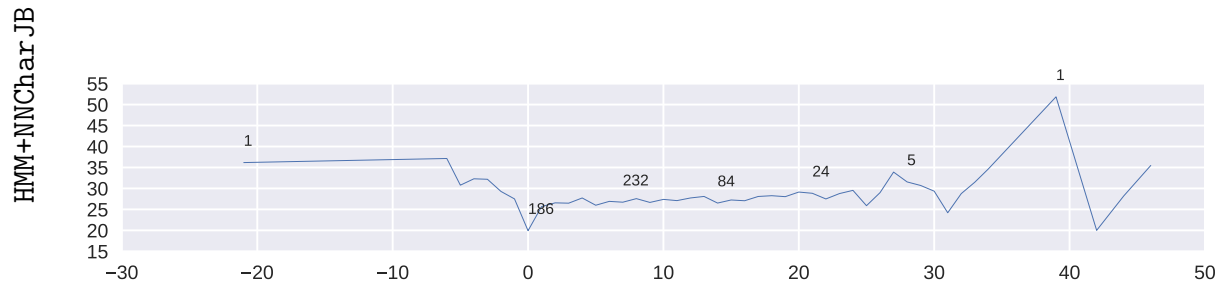


Figure B.72: Neural models for the direction Vietnamese-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.

Appendix C

Appendix: Word alignment by attention

Contents

C.1 Scores	271
C.2 Correct and incorrect alignment links	272
C.3 Unaligned words	274
C.4 Alignment types	278
C.5 AER as a function of sentence length	279
C.6 Symmetrization	285
C.6.1 Intersection	285
C.6.2 Grow-diag-final	286

C.1 Scores

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
GA	37.48	36.77	64.06	25.79	89.19	36.96	39.46	64.06	28.51	89.33
GT	35.63	37.2	66.85	25.77	89.39	34.88	40.18	67.35	28.63	89.6
UA	54.31	28.54	49.72	20.01	87.78	49.31	33	53.58	23.84	88.19
UT	51.19	33.09	48.63	25.08	87.63	47.43	33.98	54.57	24.67	88.31

Table C.1: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) of English-French

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
GA	46.33	52.48	52.78	52.18	95.86	46.82	51.94	53.47	50.49	95.9
GT	44.95	53.57	59.13	48.97	96.28	45.17	53.7	54.11	53.29	95.97
UA	63.82	35.47	35.68	35.27	94.37	62.38	36.88	37.97	35.85	94.62
UT	61.17	38.16	35.99	40.61	94.23	60.2	39.1	38.13	40.12	94.52

Table C.2: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) of English-German

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
GA	45.95	48.86	52.01	46.07	96.08	48.35	44.49	50.61	39.7	95.98
GT	43.96	50.05	56.61	44.84	96.36	46.55	45.7	54.5	39.35	96.2
UA	61.55	35.57	37.86	33.53	95.07	62.17	33.06	37.61	29.5	95.15
UT	59	38.44	38.59	38.29	95.02	59.91	36.8	35.38	38.34	94.65

Table C.3: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) of English-Czech

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
GA	50.71	49.31	51.71	47.11	96.39	46.98	53.03	56.39	50.05	96.69
GT	49.1	50.91	50.82	51	96.33	45.1	54.92	58.24	51.95	96.82
UA	63.36	36.65	38.44	35.02	95.48	63.35	36.66	38.98	34.6	95.54
UT	59.43	40.58	35.97	46.54	94.91	59.93	40.08	34.81	47.24	94.73

Table C.4: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) of English-Romanian

C.2 Correct and incorrect alignment links

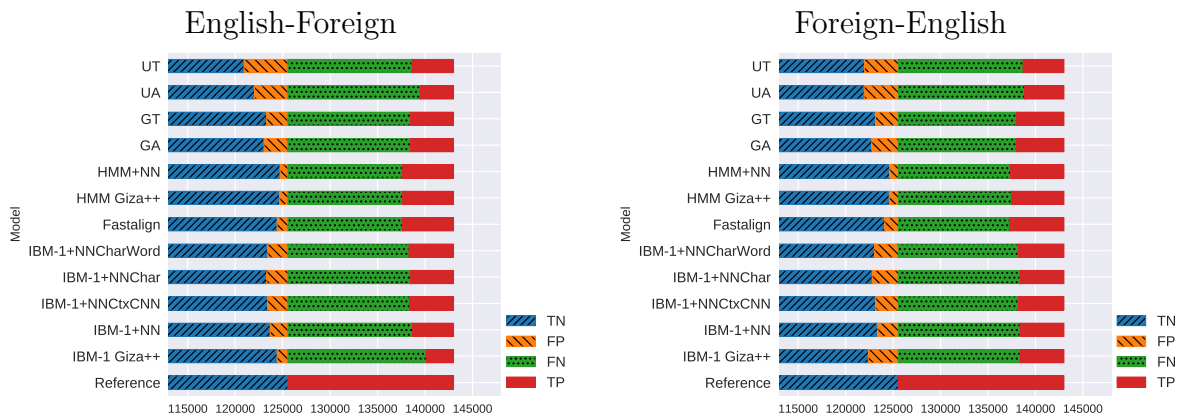


Figure C.1: Results of our attention-based models: Alignment links for English-French

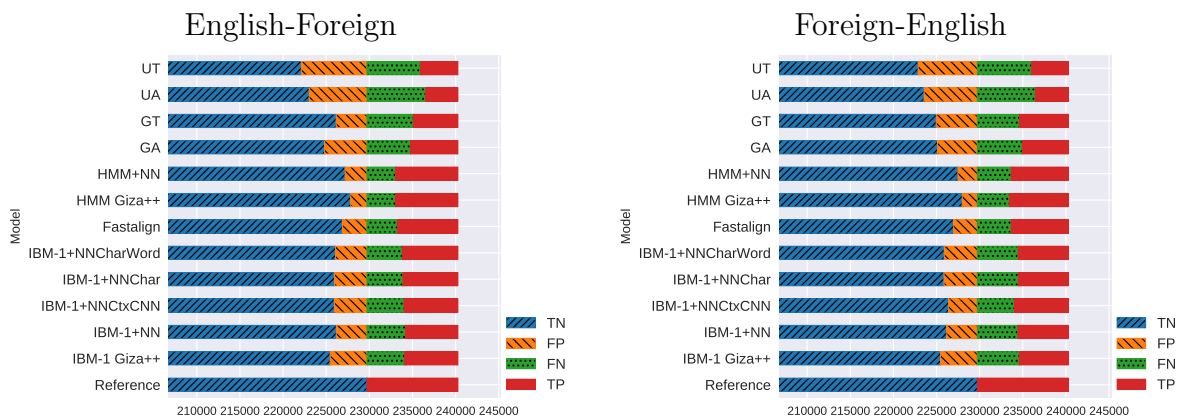


Figure C.2: Results of our attention-based models: Alignment links for English-German

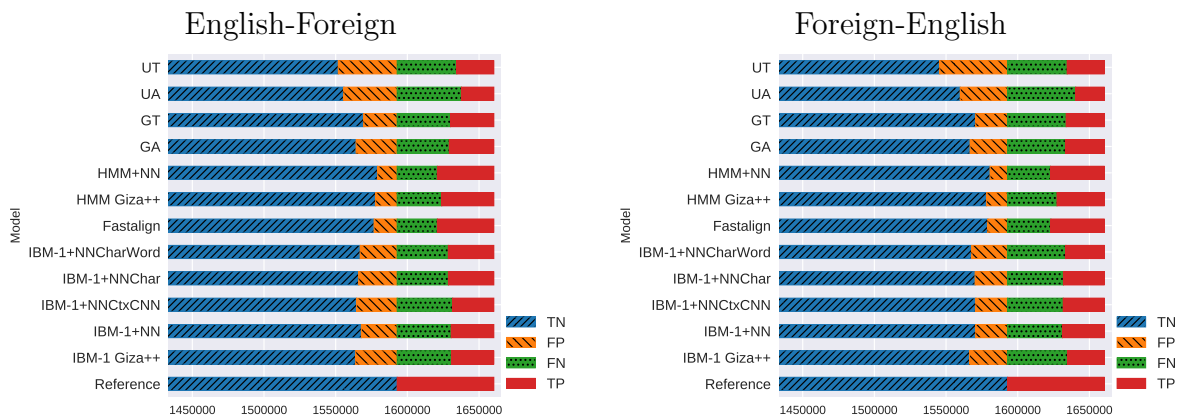


Figure C.3: Results of our attention-based models: Alignment links for English-Czech

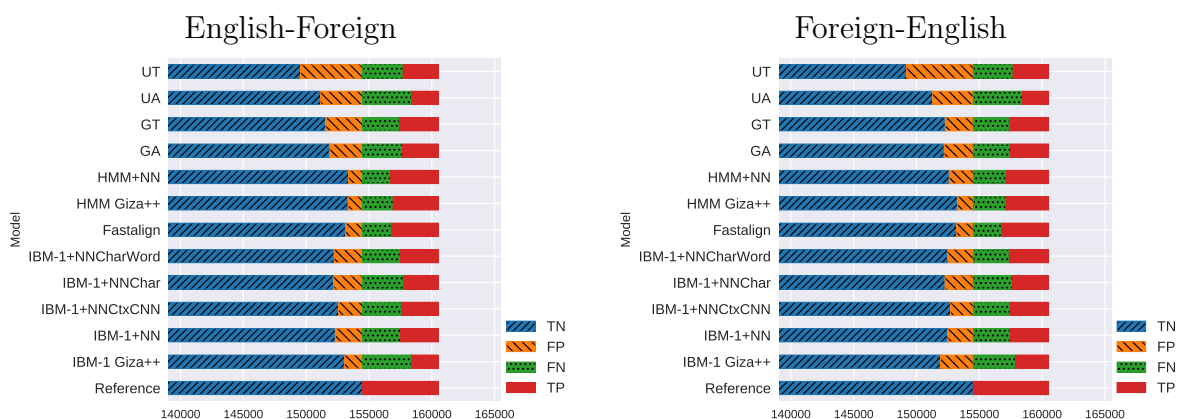


Figure C.4: Results of our attention-based models: Alignment links for English-Romanian

C.3 Unaligned words

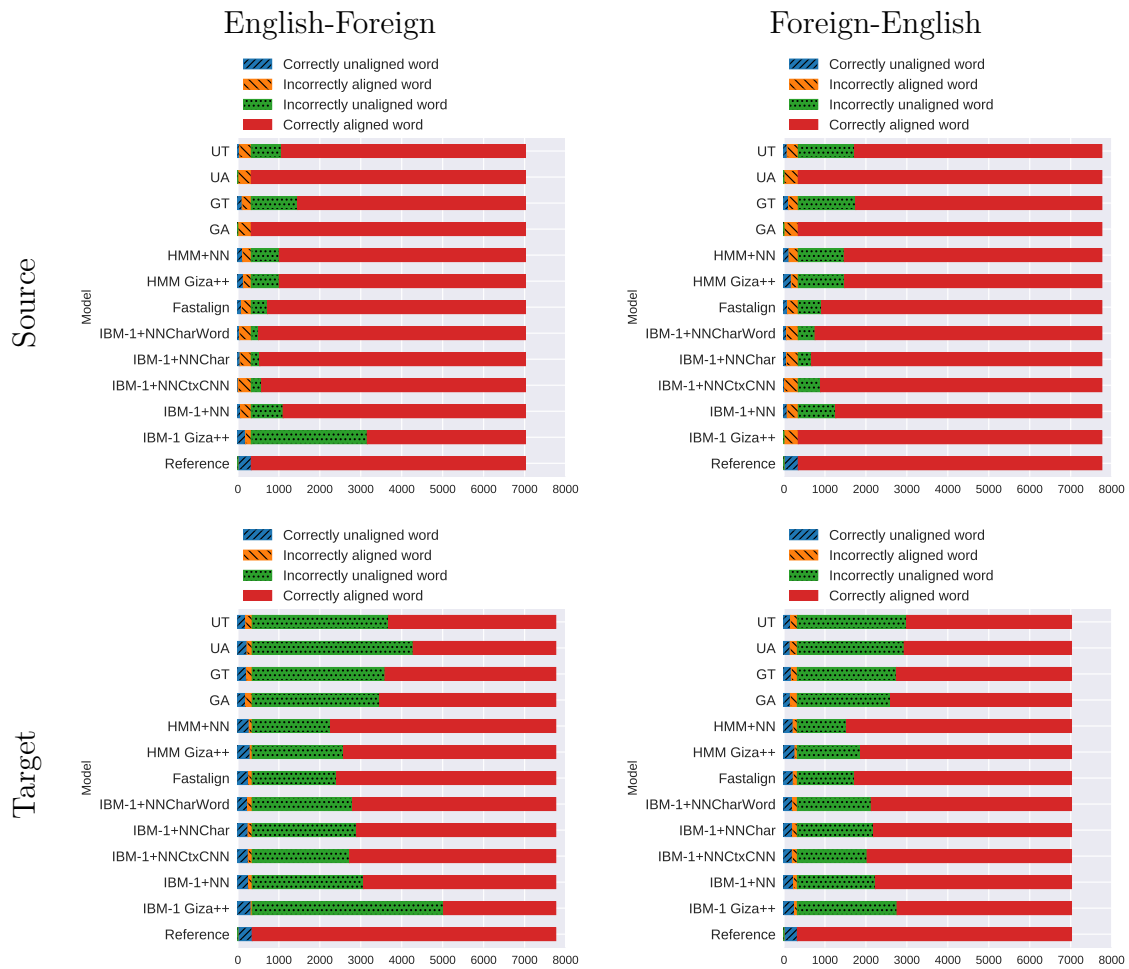


Figure C.5: Results of our attention-based models: Unaligned words for English-French

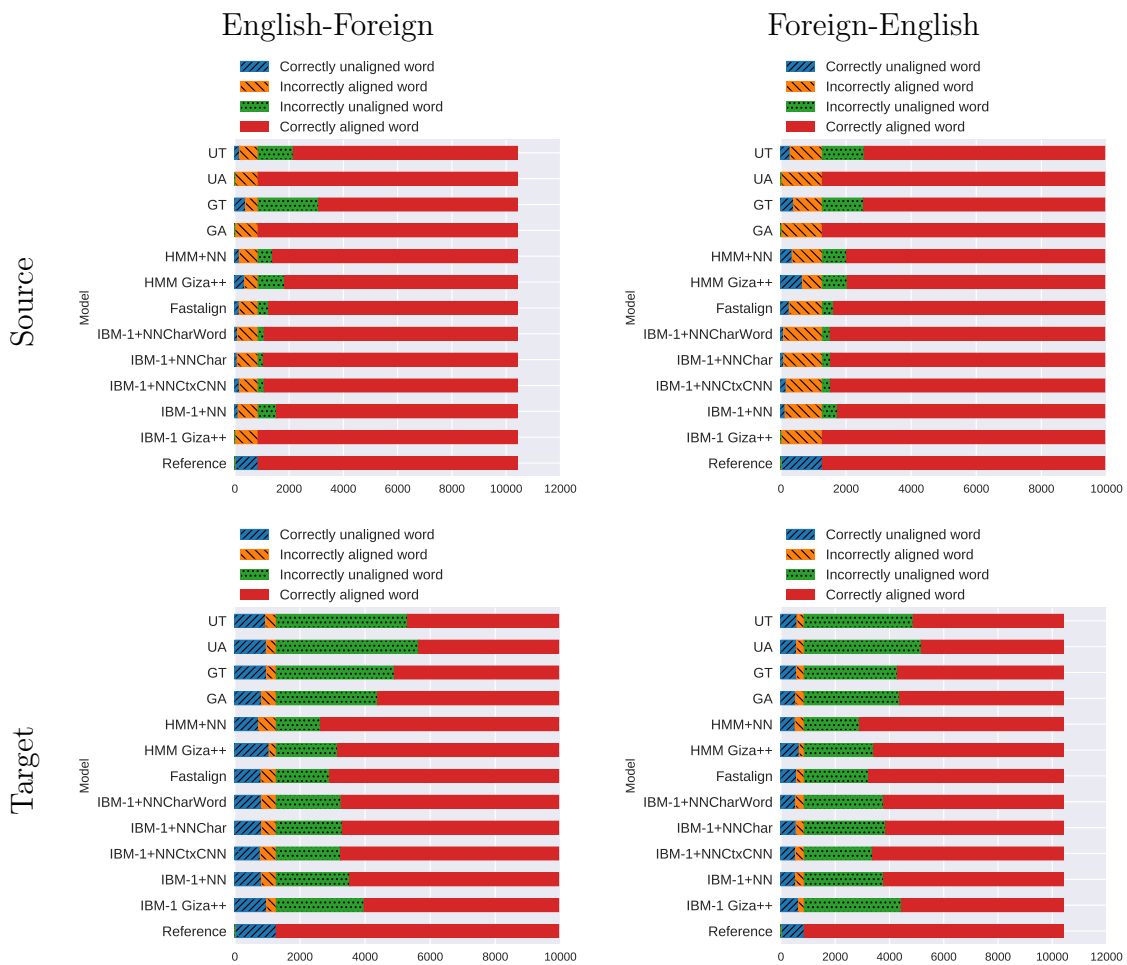


Figure C.6: Results of our attention-based models: Unaligned words for English-German

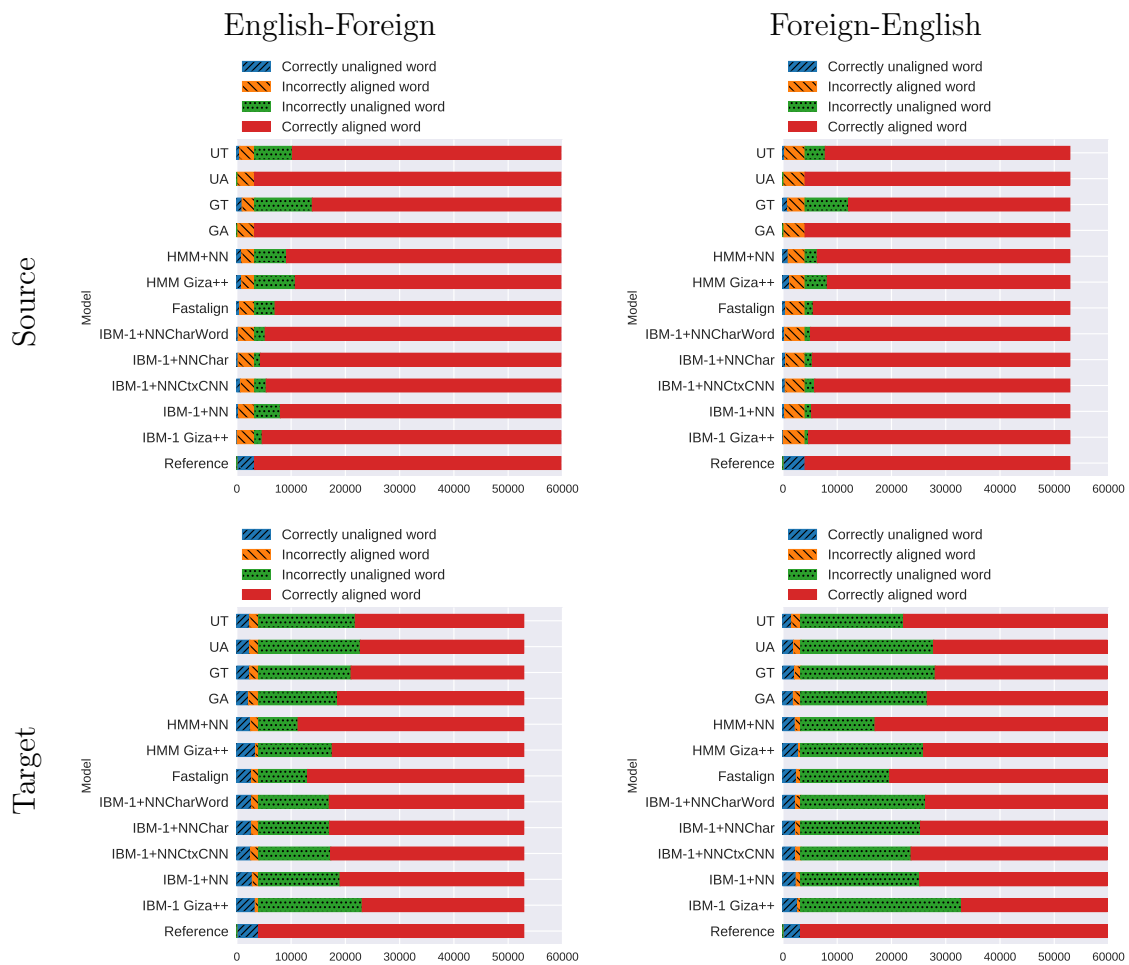


Figure C.7: Results of our attention-based models: Unaligned words for English-Czech

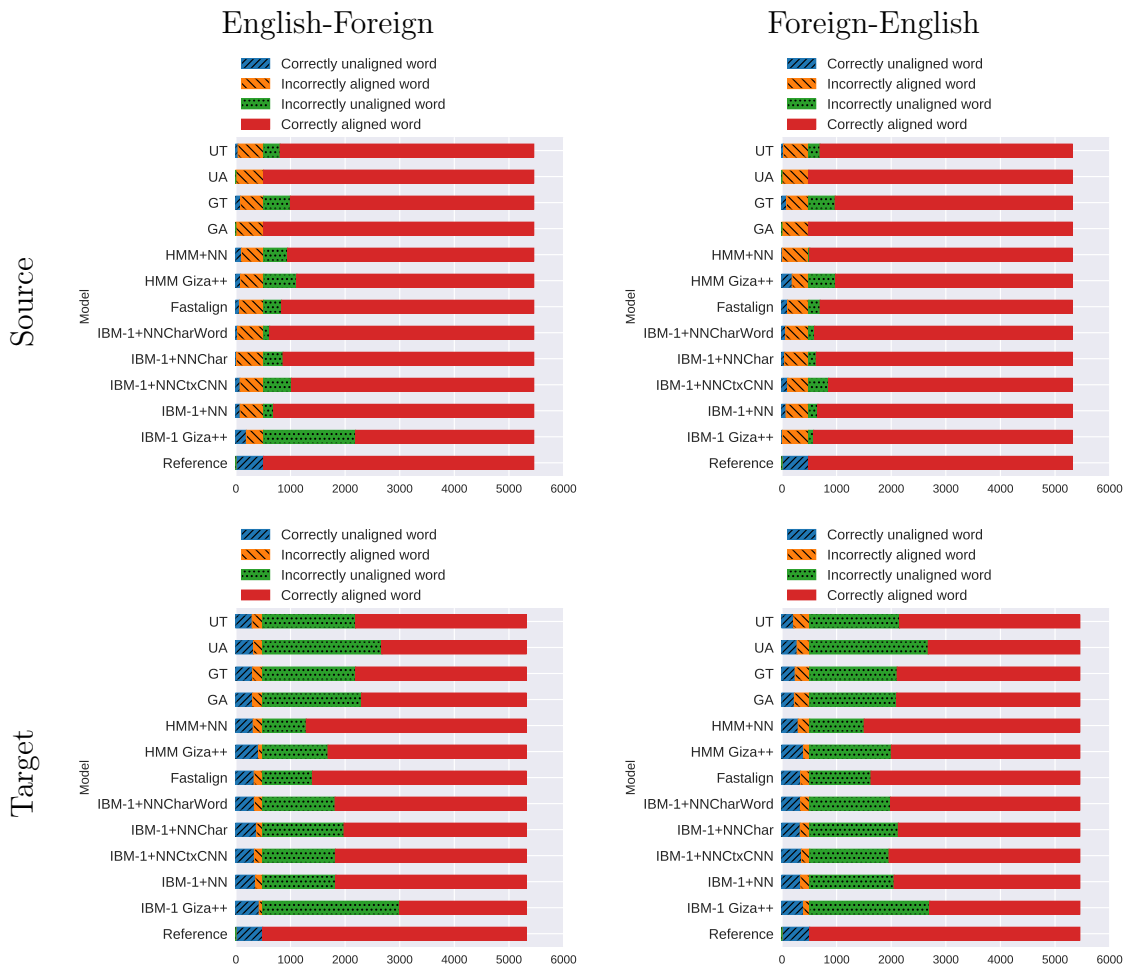


Figure C.8: Results of our attention-based models: Unaligned words for English-Romanian

C.4 Alignment types

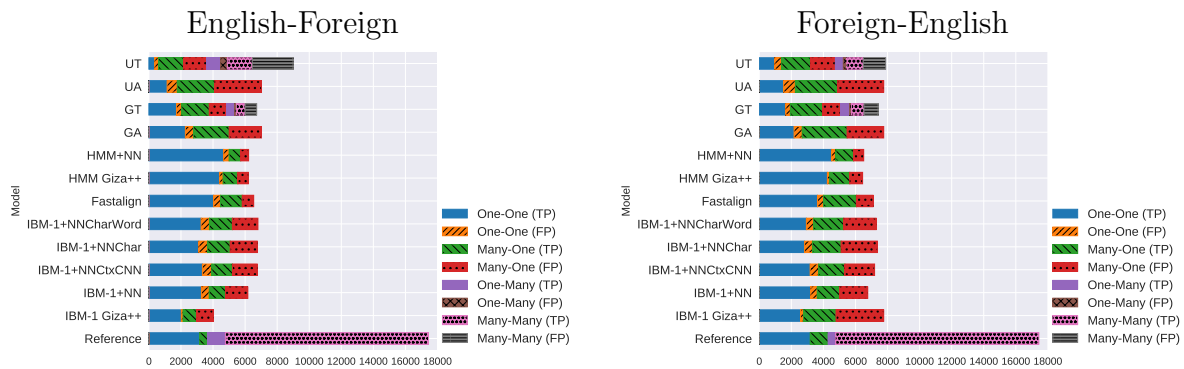


Figure C.9: Results of our attention-based models: Alignment types for English-French

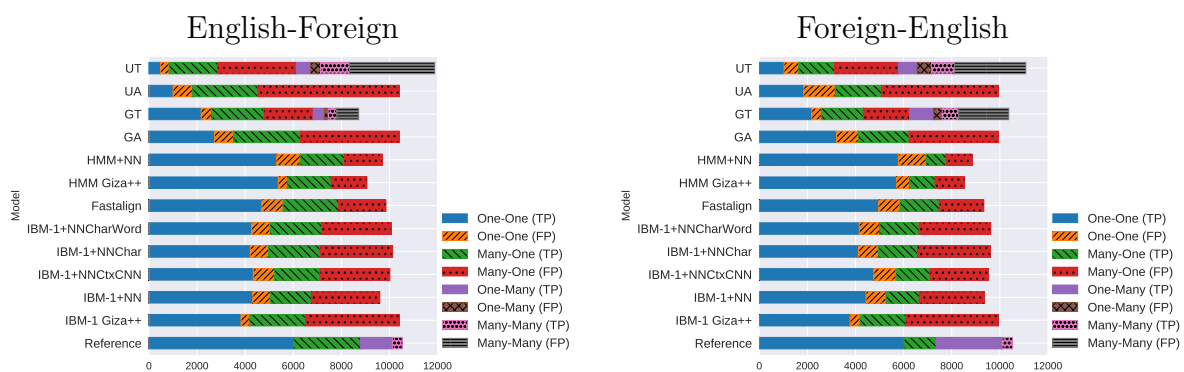


Figure C.10: Results of our attention-based models: Alignment types for English-German

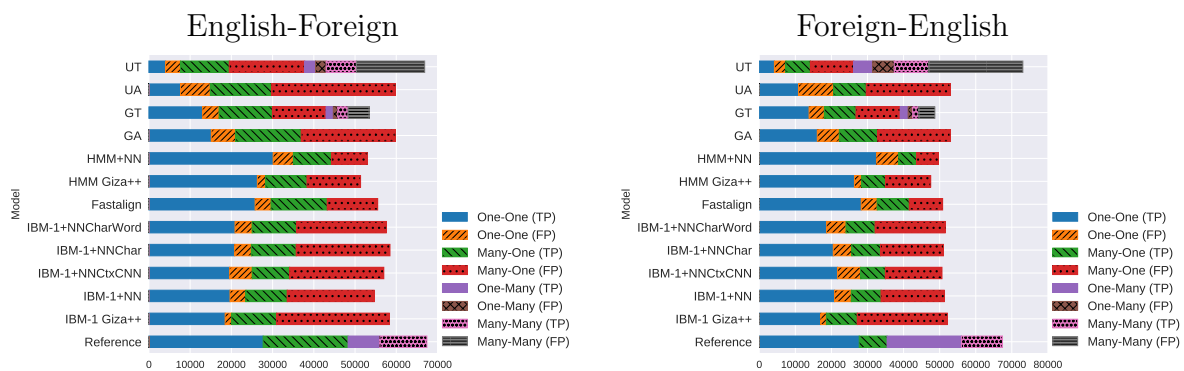


Figure C.11: Results of our attention-based models: Alignment types for English-Czech

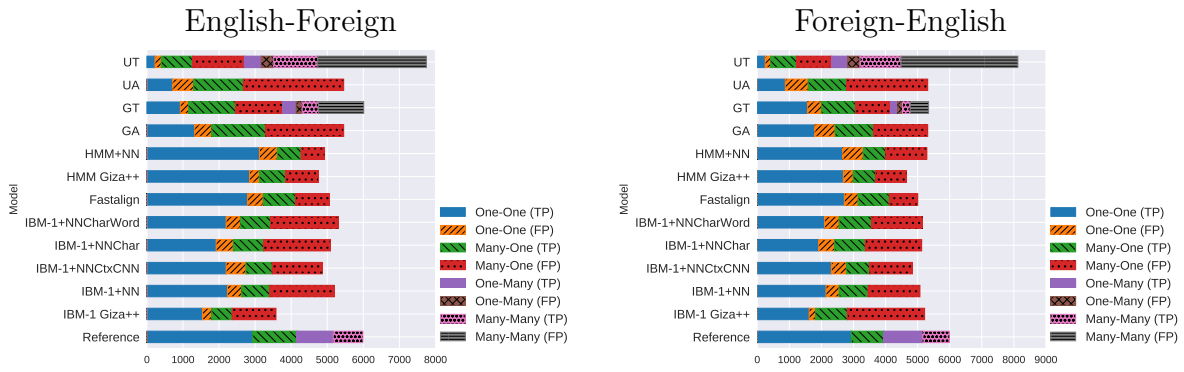


Figure C.12: Results of our attention-based models: Alignment types for English-Romanian

C.5 AER as a function of sentence length

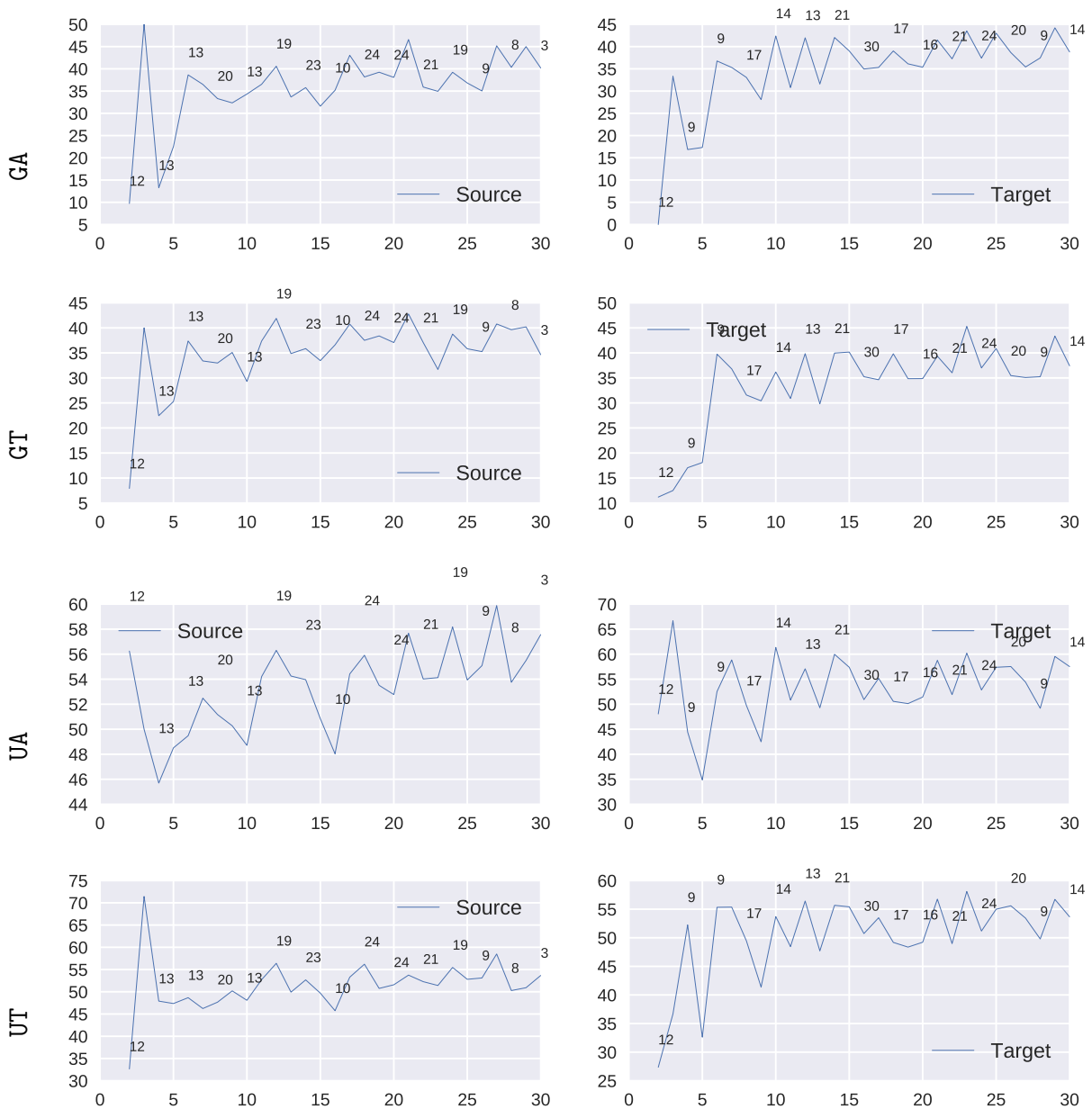


Figure C.13: Attention-based models for the direction English-French: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.

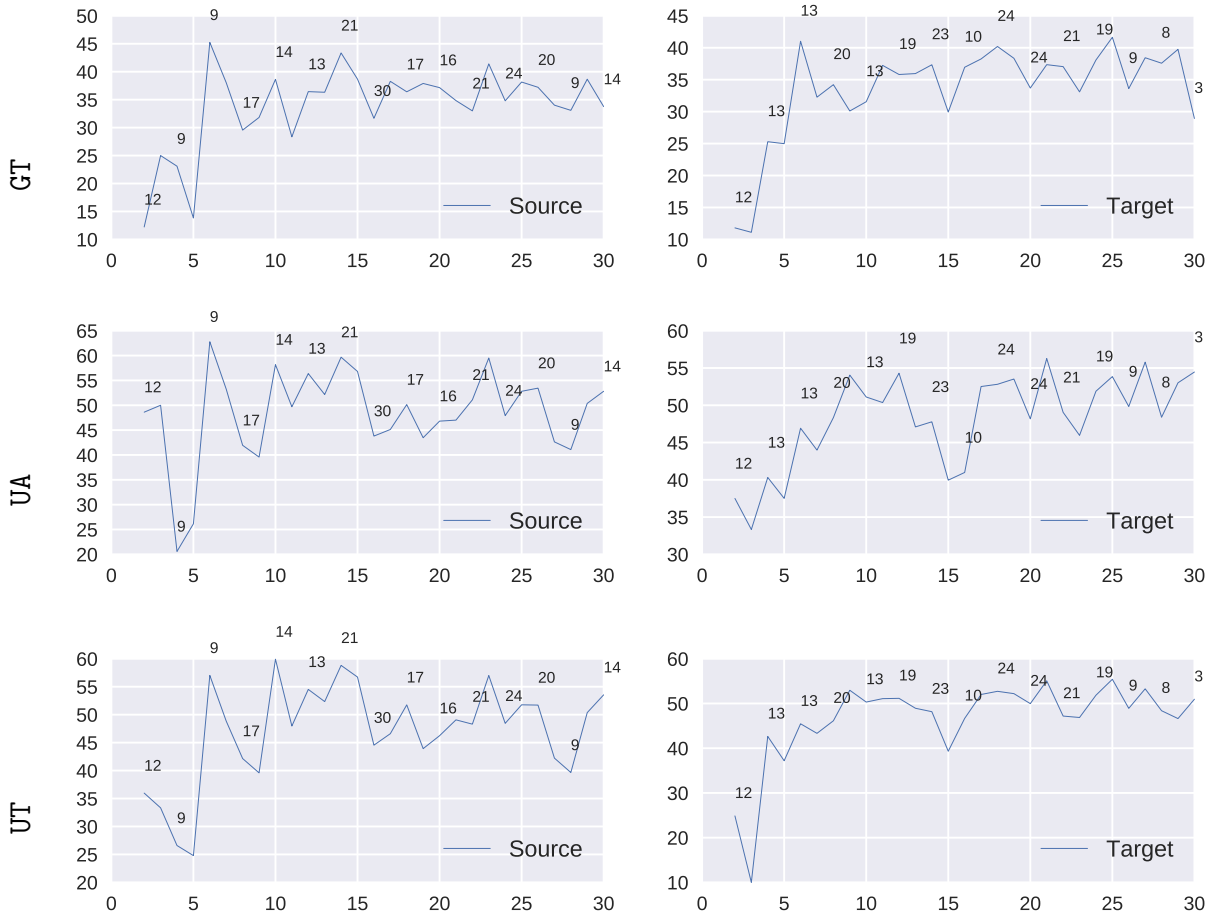


Figure C.14: Attention-based models for the direction French-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.

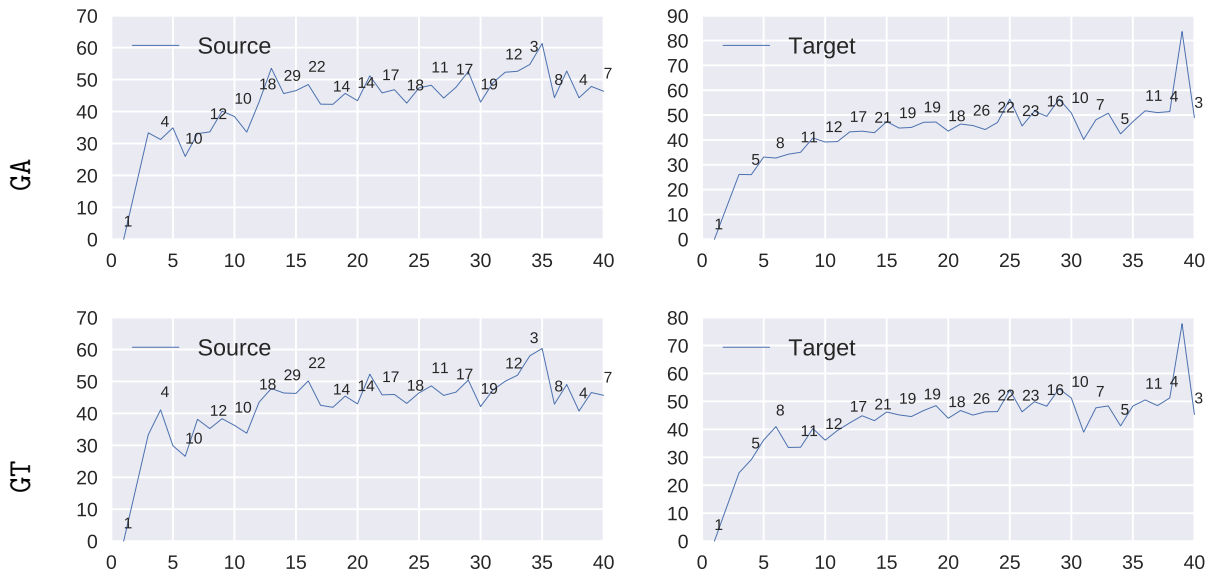
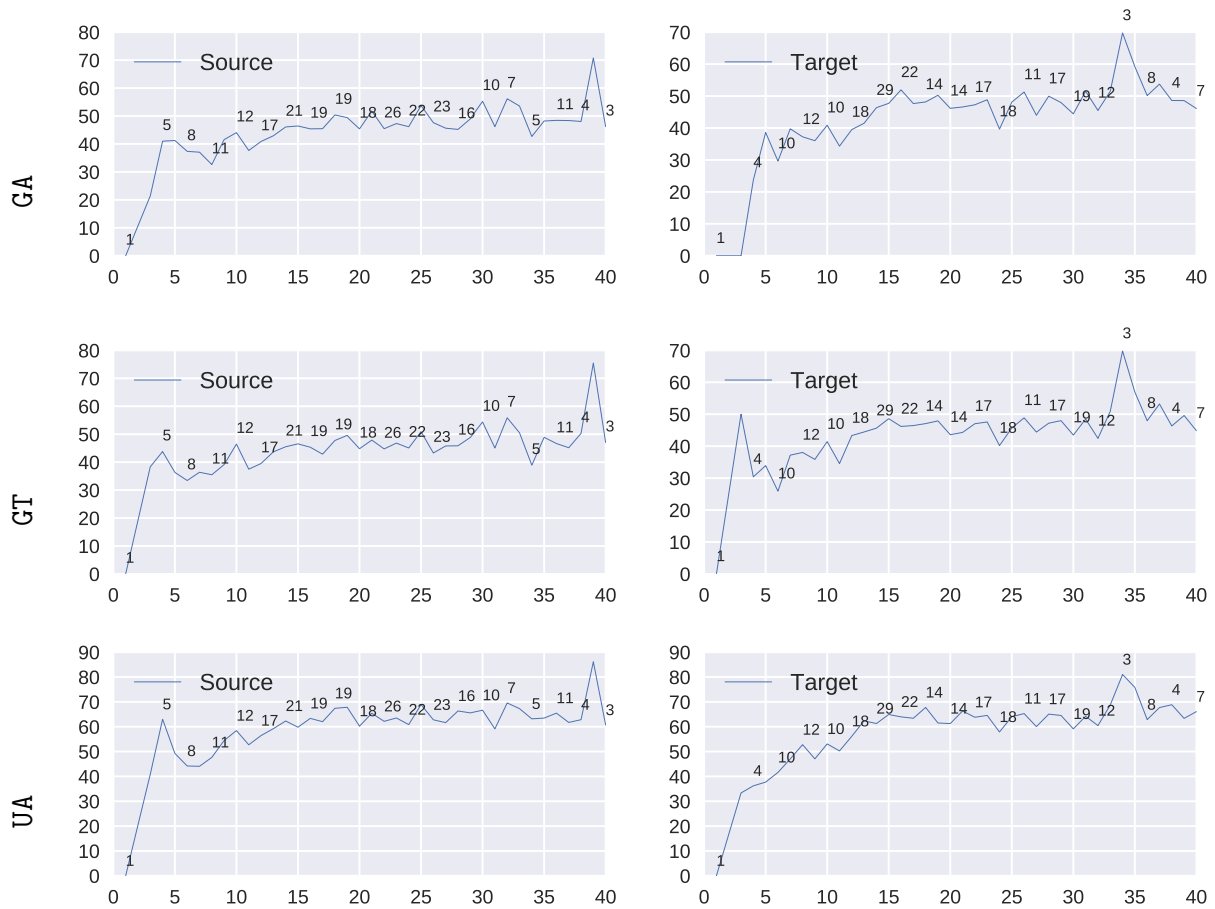




Figure C.15: Attention-based models for the direction English-German: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



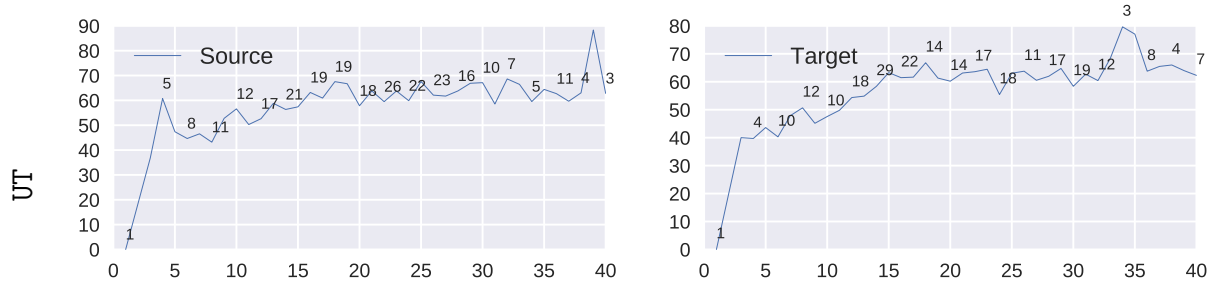


Figure C.16: Attention-based models for the direction German-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.

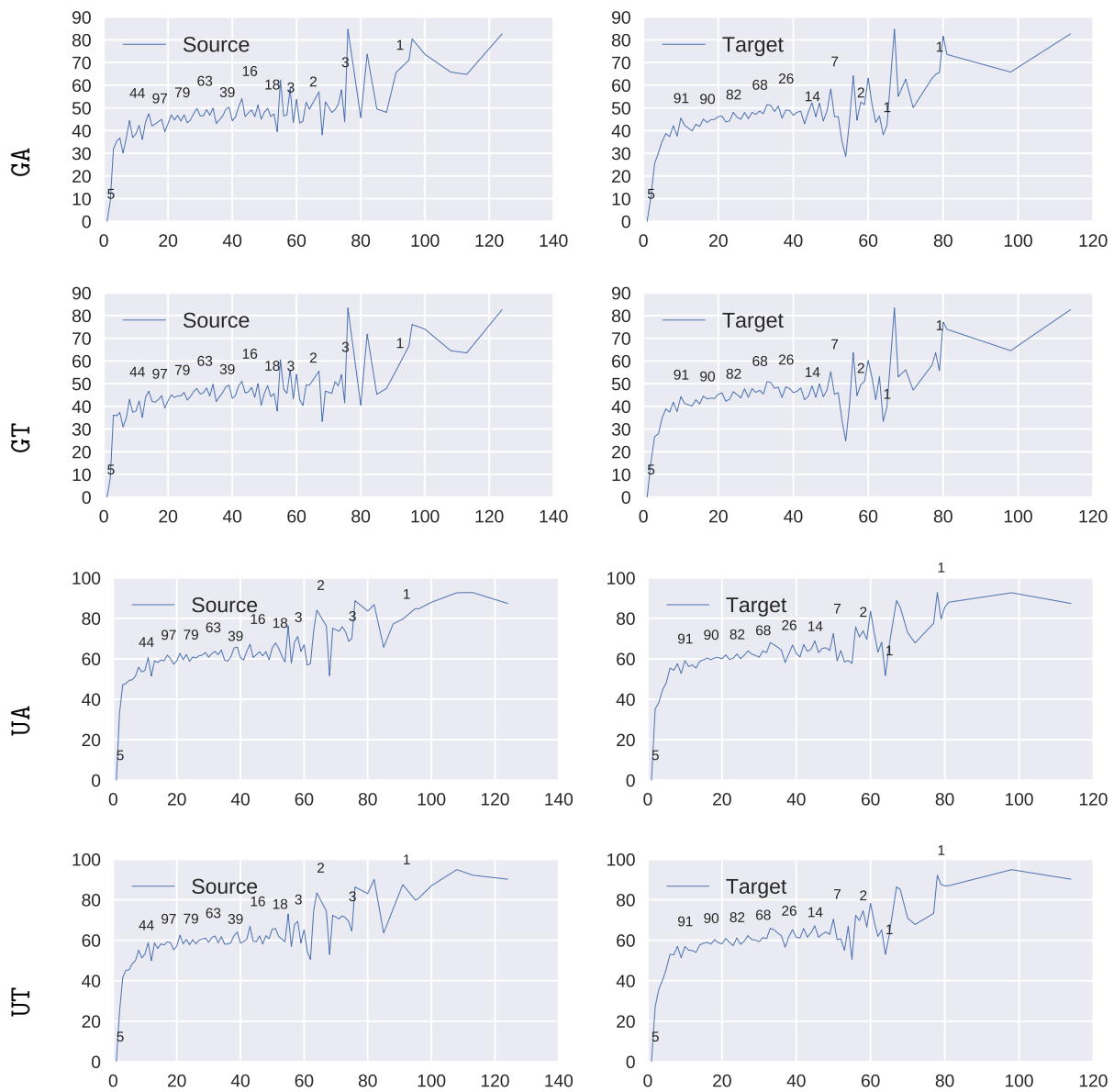


Figure C.17: Attention-based models for the direction English-Czech: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.

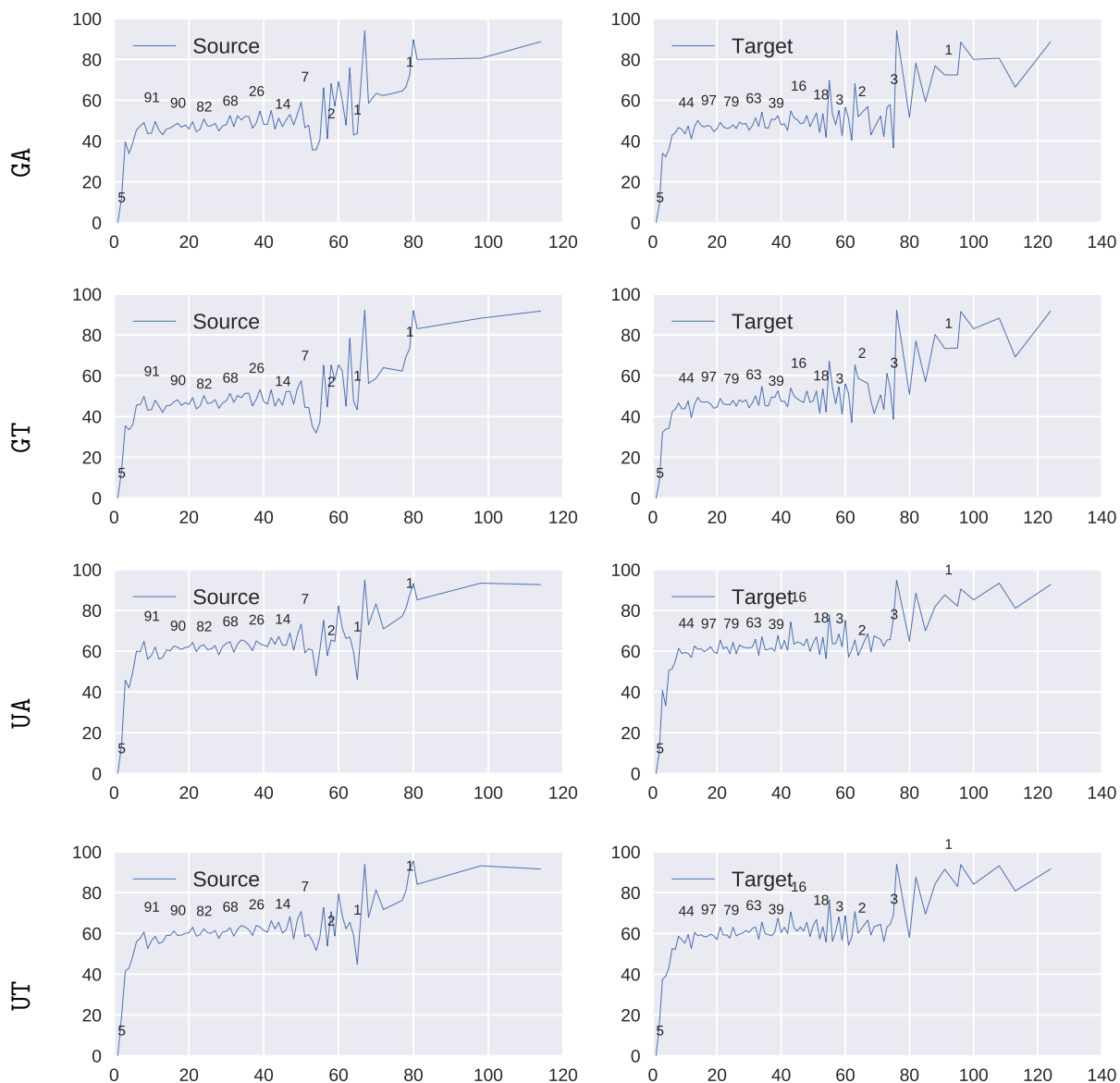
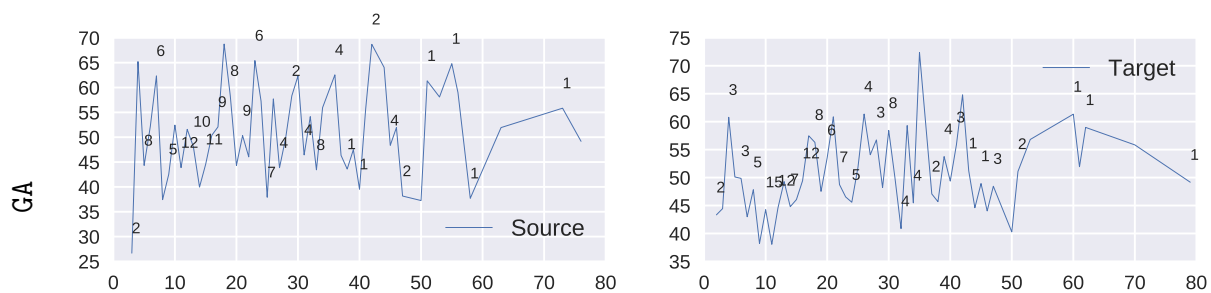


Figure C.18: Attention-based models for the direction Czech-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



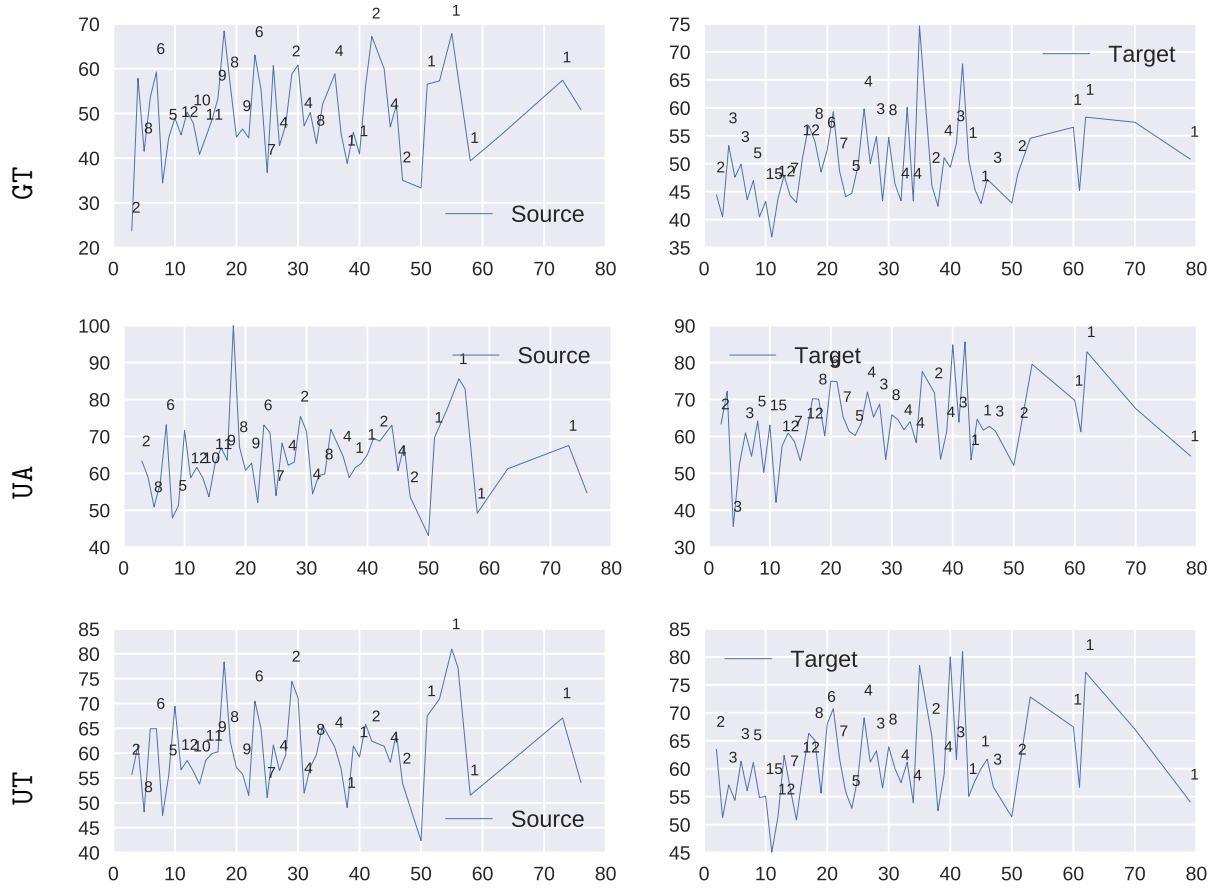


Figure C.19: Attention-based models for the direction English-Romanian: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



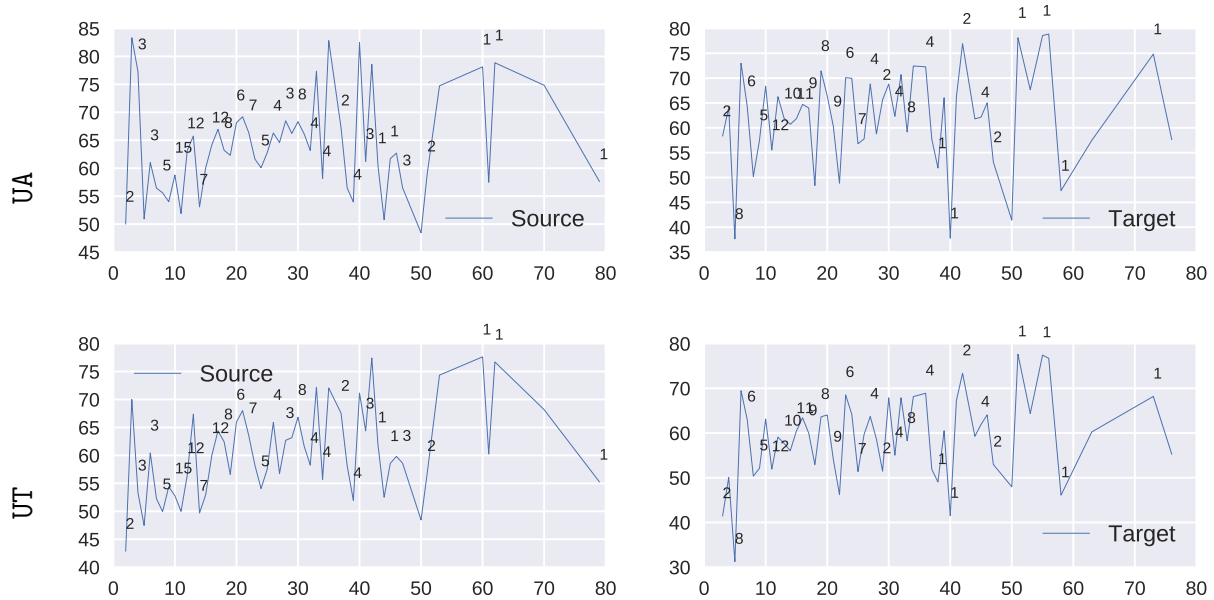


Figure C.20: Attention-based models for the direction Romanian-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.

C.6 Symmetrization

C.6.1 Intersection

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
GA	3115	0.4	0.44	0.36	0.27	0.88	0.16	0.89	0.57	0.82
GT	3353	0.45	0.5	0.34	0.28	0.88	0.17	0.9	0.65	0.91
UA	2256	0.29	0.32	0.54	0.18	0.79	0.1	0.89	0.35	0.82
UT	3182	0.4	0.35	0.47	0.23	0.74	0.14	0.89	0.7	0.91

Table C.13: Intersection alignment for attention-based models: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-French

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
GA	4193	0.42	0.4	0.49	0.49	0.85	0.34	0.97	0.74	0.67
GT	4557	0.44	0.52	0.47	0.51	0.84	0.36	0.97	0.84	0.67
UA	2762	0.28	0.27	0.67	0.31	0.75	0.2	0.96	0.63	0.67
UT	3883	0.35	0.33	0.61	0.37	0.68	0.25	0.96	0.95	0.67

Table C.14: Intersection alignment for attention-based models: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-German

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
GA	24636	0.47	0.41	0.45	0.43	0.79	0.29	0.97	0.66	0.62
GT	25091	0.52	0.47	0.45	0.43	0.79	0.3	0.97	0.66	0.66
UA	17996	0.34	0.3	0.64	0.27	0.65	0.17	0.96	0.52	0.46
UT	27775	0.38	0.42	0.58	0.34	0.59	0.24	0.96	0.78	0.67

Table C.15: Intersection alignment for attention-based models: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Czech

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
GA	2300	0.43	0.42	0.54	0.46	0.84	0.32	0.97	0.71	0.38
GT	2652	0.5	0.44	0.5	0.51	0.82	0.36	0.97	0.84	0.62
UA	1625	0.31	0.3	0.7	0.3	0.71	0.19	0.97	0.56	0.46
UT	3073	0.38	0.4	0.59	0.41	0.61	0.31	0.97	1.09	1.15

Table C.16: Intersection alignment for attention-based models: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Romanian

C.6.2 Grow-diag-final

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
GA	37.48	36.77	36.96	39.46	33.05	43.4	66.54	32.2
GT	35.63	37.2	34.88	40.18	31	41.48	71.05	29.29
UA	54.31	28.54	49.31	33	48.9	35.24	52.24	26.59
UT	51.19	33.09	47.43	33.98	46.76	34.66	54.98	25.31

Table C.17: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-French

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
GA	46.33	52.48	46.82	51.94	42.29	56.6	54.78	58.54
GT	44.95	53.57	45.17	53.7	40.59	57.97	61.45	54.86
UA	63.82	35.47	62.38	36.88	60.58	38.81	36.37	41.59
UT	61.17	38.16	60.2	39.1	59.18	40.07	38.62	41.63

Table C.18: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-German

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
GA	45.95	48.86	48.35	44.49	42.87	52.26	53.56	51.02
GT	43.96	50.05	46.55	45.7	41.66	52.24	58.53	47.17
UA	61.55	35.57	62.17	33.06	58.36	39.05	38.17	39.96
UT	59	38.44	59.91	36.8	57.78	39.38	39.08	39.68

Table C.19: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Czech

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
GA	50.71	49.31	46.98	53.03	43.74	56.27	54.64	58
GT	49.1	50.91	45.1	54.92	41.42	58.59	58.55	58.63
UA	63.36	36.65	63.35	36.66	59.58	40.43	38	43.19
UT	59.43	40.58	59.93	40.08	57.29	42.72	38.87	47.41

Table C.20: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Romanian

Appendix D

Appendix: Generative latent neural alignment models

Contents

D.1 Scores	289
D.2 Correct and incorrect alignment links	293
D.3 Unaligned words	295
D.4 Alignment types	299
D.4.1 Correct/incorrect links as a function of alignment types	301
D.5 Jump widths	303
D.5.1 Jump width statistics	303
D.5.2 Correct and incorrect jumps	309
D.6 Rare words	333
D.7 Unknown words	337
D.8 Part-of-Speech	340
D.8.1 Alignment links for content and function words	340
D.8.2 Unaligned words for content and function words	346
D.9 Symmetrization	352
D.9.1 Intersection	352
D.9.2 Grow-diag-final (GDF)	356
D.10 Training with monolingual data	360
D.11 AER as a function of sentence length	360
D.12 AER as a function of sentence length difference	383

D.1 Scores

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
IBM-1 Giza++	40.1	26.7	71.55	16.41	89.01	33.9	36.49	59.24	26.37	88.81
IBM-1+NN	27.96	36.42	69.66	24.65	89.5	27.21	38.08	68.12	26.43	89.52
IBM-1+NNChar	28.76	37.5	67.13	26.01	89.42	31.4	37.21	62.64	26.47	89.11
IBM-1+BPE	22.98	34.56	83.07	21.81	89.92	28.69	33.12	78.79	20.97	89.67
IBM-1+BPE+B	25.71	37.27	70.64	25.31	89.61	24.05	40.42	70.84	28.28	89.83
IBM-1+BPE+VAE	33.42	29.21	71.75	18.34	89.16	34.36	30.96	67.44	20.09	89.07
IBM-1+BPE+VAE+SP	22.12	34.73	78.48	22.3	89.78	23.89	34.97	74.9	22.81	89.65
IBM-1+BPE+VAE+SP+AC	22.87	34.44	77.78	22.12	89.73	23.61	34.35	77.06	22.1	89.7
Fastalign	15.19	44.98	82.5	30.92	90.78	16.23	46.32	80.08	32.58	90.79
HMM Giza++	11.99	45.18	86.12	30.62	90.94	11.97	45.98	85.2	31.49	90.98
HMM+NN	11.84	45.57	86.68	30.91	91	11.15	46.86	86.17	32.18	91.1
HMM+NNCharTgt	9.17	47.22	89.53	32.07	91.26	9.56	47.87	88.15	32.86	91.27
HMM+NNCharJB	8.47	44.38	91.8	29.26	91.06	7.74	46.26	91.42	30.96	91.23
HMM+BPE	10.91	41.1	92.18	26.44	90.76	10.97	42.98	91.15	28.12	90.9
HMM+BPE+B	9.84	41.74	93.17	26.9	90.84	10.48	43.2	91.57	28.27	90.94
HMM+BPE+VAE	18.92	34.11	88.66	21.11	90.05	12.94	37.53	90.35	23.68	90.39
HMM+BPE+VAE+SP	12.94	35.86	93.37	22.19	90.32	12.27	36.09	93.7	22.35	90.35
HMM+BPE+VAE+SP+AC	11.47	38.73	91.18	24.59	90.51	10.84	39.2	91.24	24.96	90.56
IBM-4 Giza++	10	44.43	90.61	29.43	91.02	9.64	45.43	89.58	30.43	91.08

Table D.1: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-French

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
IBM-1 Giza++	39.03	58.76	59.1	58.43	96.4	42.66	55.39	57.02	53.84	96.2
IBM-1+NN	37.64	60.07	62.98	57.41	96.65	39.22	58.53	62.22	55.25	96.57
IBM-1+NNChar	36.22	61.55	62.76	60.39	96.69	40.88	56.99	59.75	54.48	96.4
IBM-1+BPE	32.91	64.95	71.99	59.17	97.2	35.54	62.35	63.21	61.51	96.74
IBM-1+BPE+B	31.36	66.52	73.38	60.83	97.32	34.46	63.34	64.35	62.36	96.84
IBM-1+BPE+VAE	38.92	58.4	66.56	52.02	96.75	40.87	56.51	66.84	48.94	96.7
IBM-1+BPE+VAE+SP	24.41	73.38	79.9	67.83	97.84	38.72	58.67	68.3	51.42	96.82
IBM-1+BPE+VAE+SP+AC	24.3	73.6	78.96	68.92	97.83	29.37	68.64	75.55	62.88	97.48
Fastalign	28.98	68.75	71.11	66.54	97.35	31.28	66.47	70.73	62.69	97.23
HMM Giza++	23.92	73.3	79.23	68.2	97.82	26.33	71.04	79.47	64.23	97.7
HMM+NN	26.78	70.95	73.94	68.2	97.55	29.44	68.21	74.69	62.76	97.44
HMM+NNCharTgt	26.04	71.57	75.99	67.64	97.64	28.11	69.48	75.59	64.29	97.52
HMM+NNCharJB	23.69	73.38	82.38	66.15	97.9	24.9	72.16	83.36	63.61	97.85
HMM+BPE	23.66	74.28	80.69	68.81	97.91	22.75	74.94	79.81	70.64	97.93
HMM+BPE+B	19.61	78.25	85.82	71.92	98.25	20.38	77.38	84.28	71.52	98.17
HMM+BPE+VAE	23.96	72.98	82.54	65.4	97.88	26.5	70.63	78.56	64.15	97.66
HMM+BPE+VAE+SP	23.73	73.73	84.45	65.42	97.96	26.4	71.13	82.89	62.29	97.78
HMM+BPE+VAE+SP+AC	19.13	78.38	89.92	69.46	98.32	20.58	76.98	90.72	66.86	98.25
IBM-4 Giza++	21.46	75.48	85.79	67.39	98.08	23.31	73.63	86.56	64.06	97.99

Table D.2: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-German

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
IBM-1 Giza++	45.09	46.75	50.4	43.59	95.97	48.47	42.88	49.17	38.02	95.89
IBM-1+NN	42.29	48.64	54.32	44.04	96.22	40.97	49.08	56.81	43.2	96.36
IBM-1+NNChar	40.85	50.35	54.2	47.01	96.24	42.35	47.99	55.68	42.17	96.29
IBM-1+BPE	34.3	56.91	69.38	48.24	97.03	35.54	54.2	69.05	44.6	96.94
IBM-1+BPE+B	30.76	59.86	72.25	51.1	97.22	32.71	56.45	71.4	46.68	97.08
IBM-1+BPE+VAE	41.92	47.3	59.51	39.25	96.45	45.3	44.9	56.14	37.41	96.27
IBM-1+BPE+VAE+SP	39.61	49.2	61.7	40.91	96.57	42.63	47.6	55.41	41.72	96.27
IBM-1+BPE+VAE+SP+AC	35.41	56.47	61.84	51.96	96.75	33.83	56.95	64.02	51.28	96.85
Fastalign	25.75	64.09	70.98	58.42	97.34	25.3	62.86	73.13	55.13	97.36
HMM Giza++	27.86	61.22	70.81	53.92	97.23	30.38	57.28	69.26	48.83	97.04
HMM+NN	23.5	65.45	74.39	58.43	97.5	24.06	64.03	75.48	55.6	97.46
HMM+NNCharTgt	16.74	69.36	83.82	59.15	97.88	24.61	63.94	73.42	56.63	97.41
HMM+NNCharJB	16.38	69.35	83.84	59.13	97.88	23.88	62.72	78.56	52.19	97.48
HMM+BPE	16.68	71.66	84.98	61.95	98.01	19.87	67.02	82.23	56.56	97.74
HMM+BPE+B	16.24	71.87	85.27	62.1	98.03	18.74	67.93	82.85	57.56	97.79
HMM+BPE+VAE	23.71	62.57	77.34	52.54	97.45	29.01	59.05	70.32	50.89	97.13
HMM+BPE+VAE+SP	20.31	69.01	82.62	59.25	97.84	20.12	67.94	84.46	56.83	97.82
HMM+BPE+VAE+SP+AC	17.56	72.53	84.91	63.3	98.05	20.1	68.03	84.41	56.98	97.83
IBM-4 Giza++	20.92	65.7	79.48	56	97.63	26.5	59.81	75.58	49.48	97.3

Table D.3: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Czech

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
IBM-1 Giza++	56.02	43.99	58.8	35.14	96.66	53.52	46.49	49.92	43.5	96.26
IBM-1+NN	46.4	53.62	57.71	50.07	96.77	44.9	55.11	60.08	50.9	96.91
IBM-1+NNChar	50.16	49.85	54.28	46.09	96.54	48.28	51.73	56.08	48.01	96.66
IBM-1+BPE	43.77	56.25	79.61	43.49	97.48	43.96	56.05	70.8	46.39	97.29
IBM-1+BPE+B	43.47	56.54	63.95	50.67	97.09	40.42	59.59	64.19	55.61	97.19
IBM-1+BPE+VAE	56.39	43.63	56.66	35.47	96.58	55.7	44.32	53.94	37.61	96.47
IBM-1+BPE+VAE+SP	49.3	50.71	60.69	43.55	96.84	51.49	48.52	57.82	41.8	96.69
IBM-1+BPE+VAE+SP+AC	49.12	50.89	61.31	43.5	96.87	49.2	50.81	59	44.62	96.78
Fastalign	33.36	66.65	72.77	61.49	97.7	32.91	67.1	73.7	61.59	97.75
HMM Giza++	33.36	66.65	75.28	59.8	97.77	36.38	63.64	72.9	56.46	97.59
HMM+NN	30.69	69.33	76.93	63.09	97.92	40.12	59.89	63.85	56.4	97.18
HMM+NNCharTgt	27.54	72.47	78.13	67.59	98.08	28.01	72.01	82.79	63.71	98.15
HMM+NNCharJB	26.85	73.17	83.55	65.08	98.22	29.76	70.26	80.7	62.21	98.04
HMM+BPE	34.23	65.79	84.07	54.04	97.9	30.75	69.27	82.44	59.74	98.02
HMM+BPE+B	34.41	65.61	84.04	53.81	97.9	29.34	70.68	82.91	61.59	98.09
HMM+BPE+VAE	50.29	49.73	75.24	37.14	97.2	38.64	61.38	79.8	49.87	97.66
HMM+BPE+VAE+SP	37.52	62.5	87.99	48.46	97.83	38.04	61.98	88.3	47.75	97.81
HMM+BPE+VAE+SP+AC	35.55	64.47	81.66	53.26	97.81	38.87	61.15	78.09	50.25	97.62
IBM-4 Giza++	31.04	68.98	79.28	61.04	97.95	32.3	67.72	80.97	58.2	97.93

Table D.4: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Romanian

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
IBM-1 Giza++	63.12	36.88	40.24	34.04	97.22	61.55	38.45	40.13	36.9	97.18
IBM-1+NN	62.64	37.36	48.83	30.25	97.58	56.9	43.1	49.95	37.91	97.61
IBM-1+NNChar	57.96	42.04	57.54	33.12	97.82	54.91	45.09	47.76	42.7	97.52
IBM-1+BPE	57.41	42.59	48.09	38.21	97.54	57.39	42.61	44.26	41.08	97.36
IBM-1+BPE+B	56.51	43.49	48.67	39.3	97.56	57.27	42.73	44.39	41.18	97.37
IBM-1+BPE+VAE	58.66	41.34	49.2	35.65	97.59	59.95	40.05	44.25	36.57	97.39
IBM-1+BPE+VAE+SP	55.54	44.43	47.58	41.67	97.51	55.1	44.9	47.39	42.66	97.5
IBM-1+BPE+VAE+SP+AC	54.81	45.18	48.2	42.51	97.54	54.23	45.76	46.7	44.85	97.46
Fastalign	50.67	49.33	55.79	44.21	97.83	49.37	50.63	54.59	47.2	97.8
HMM Giza++	57.01	42.99	52.22	36.53	97.69	54.41	45.59	52.45	40.32	97.7
HMM+NN	49.68	50.32	70.46	39.13	98.16	47.09	52.91	56.81	49.51	97.9
HMM+NNCharTgt	43.92	56.08	65.13	49.25	98.16	43.6	56.4	67.58	48.39	98.21
HMM+NNCharJB	40.92	59.08	84.77	45.35	98.5	37.71	62.29	67.77	57.62	98.33
HMM+BPE	41.22	58.69	69.82	50.61	98.3	41.01	58.97	60.26	57.74	98.08
HMM+BPE+B	38.33	61.65	72.6	53.56	98.41	38.93	61.04	62.67	59.5	98.19
HMM+BPE+VAE	49.27	50.73	62.15	42.86	98.01	53.78	46.22	46.61	45.83	97.45
HMM+BPE+VAE+SP	43.67	56.32	65.78	49.24	98.18	48.99	51.01	52.15	49.91	97.71
HMM+BPE+VAE+SP+AC	40.86	59.12	70.07	51.14	98.31	45.24	54.76	56.57	53.06	97.91
IBM-4 Giza++	52.52	47.48	61.27	38.76	97.95	49.23	50.77	61.46	43.24	98

Table D.5: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Japanese

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
IBM-1 Giza++	69.43	30.58	43.42	23.59	96.51	46.45	53.56	50.79	56.64	96.8
IBM-1+NN	58.87	41.13	57.25	32.1	97.01	42.25	57.75	57.33	58.19	97.23
IBM-1+NNChar	55.06	44.94	62.54	35.07	97.2	41.15	58.86	63.9	54.56	97.51
IBM-1+BPE	57.17	42.83	47.14	39.25	96.59	37.59	62.38	64.37	60.51	97.62
IBM-1+BPE+B	55.85	44.15	49.18	40.06	96.7	37.72	62.26	64.34	60.31	97.62
IBM-1+BPE+VAE	56.47	43.53	49.74	38.7	96.73	53.56	46.44	50.54	42.96	96.77
IBM-1+BPE+VAE+SP	51.34	48.61	54.75	43.71	96.99	41.38	58.63	58.37	58.88	97.29
IBM-1+BPE+VAE+SP+AC	50.84	49.09	55.3	44.13	97.02	38.77	61.23	61.17	61.3	97.47
Fastalign	48.89	51.11	56.87	46.42	97.11	32.82	67.19	64.5	70.11	97.77
HMM Giza++	57.86	42.14	49.18	36.86	96.7	37.57	62.43	60.28	64.74	97.46
HMM+NN	49.27	50.73	63.52	42.24	97.33	31.45	68.56	67.75	69.39	97.93
HMM+NNCharTgt	47.52	52.49	67.71	42.86	97.47	30.94	69.06	75.69	63.51	98.15
HMM+NNCharJB	43.28	56.73	84.49	42.7	97.88	27.59	72.42	72.7	72.14	98.21
HMM+BPE	48.34	51.66	60.41	45.13	97.25	28.16	71.85	74.15	69.69	98.22
HMM+BPE+B	47.03	52.97	62.75	45.83	97.35	27.76	72.24	74.13	70.45	98.24
HMM+BPE+VAE	48.97	51.04	63.3	42.75	97.33	39.2	60.8	65.17	56.99	97.61
HMM+BPE+VAE+SP	45.87	54.13	64.85	46.45	97.43	33.78	66.22	69.36	63.35	97.89
HMM+BPE+VAE+SP+AC	43.94	56.05	68.4	47.48	97.57	32.59	67.41	71.33	63.9	97.99
IBM-4 Giza++	51.91	48.1	58.38	40.89	97.12	33.19	66.82	67.17	66.47	97.85

Table D.6: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Vietnamese

D.2 Correct and incorrect alignment links

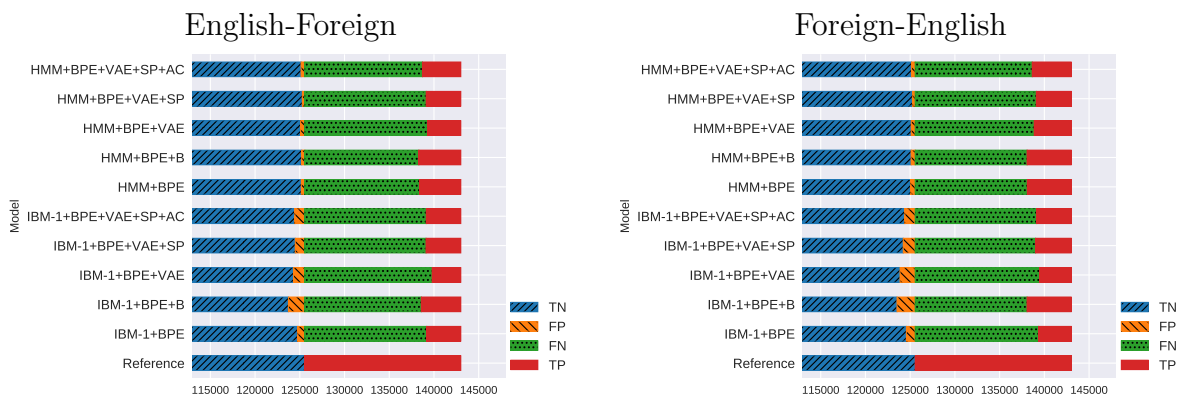


Figure D.1: Results of our variational models: Alignment links for English-French

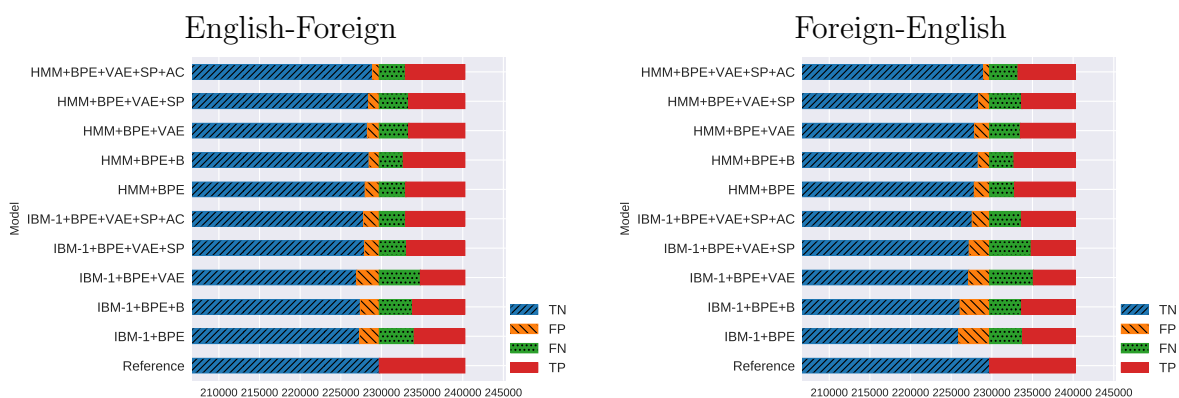


Figure D.2: Results of our variational models: Alignment links for English-German

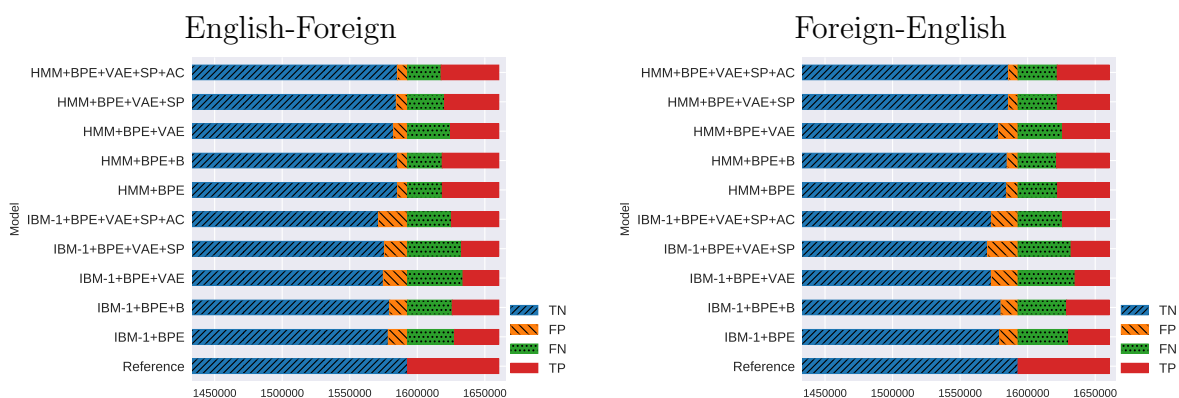


Figure D.3: Results of our variational models: Alignment links for English-Czech

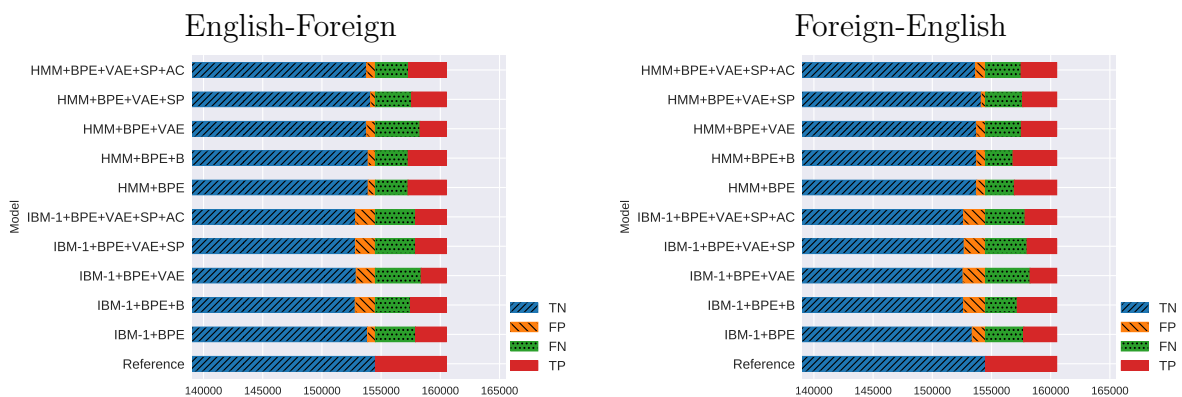


Figure D.4: Results of our variational models: Alignment links for English-Romanian

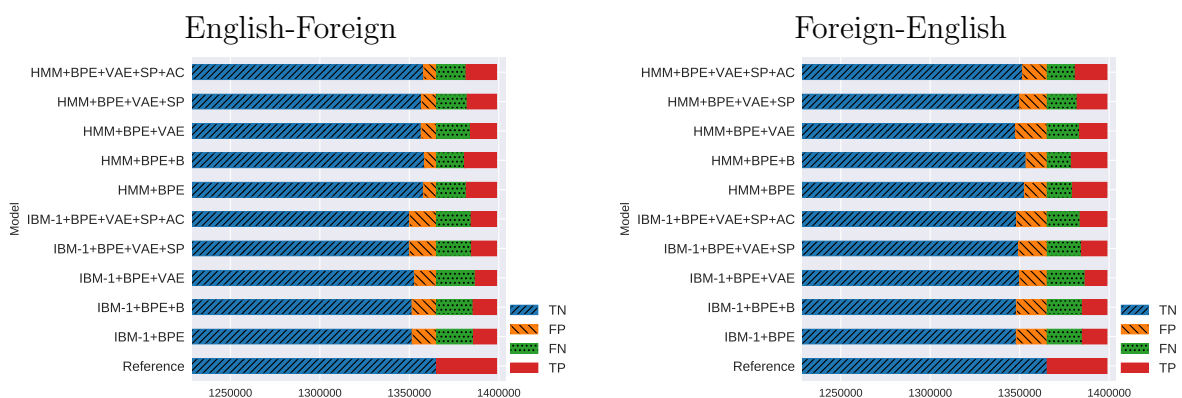


Figure D.5: Results of our variational models: Alignment links for English-Japanese

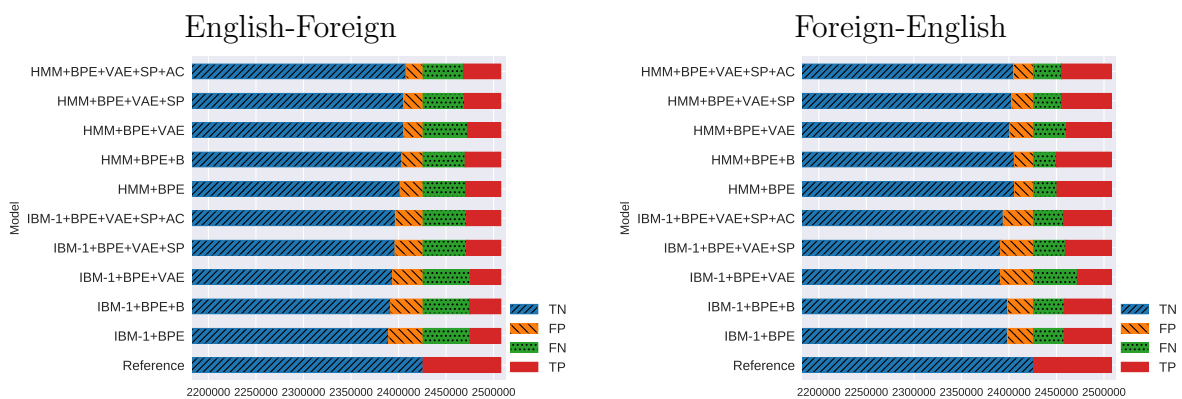


Figure D.6: Results of our variational models: Alignment links for English-Vietnamese

D.3 Unaligned words

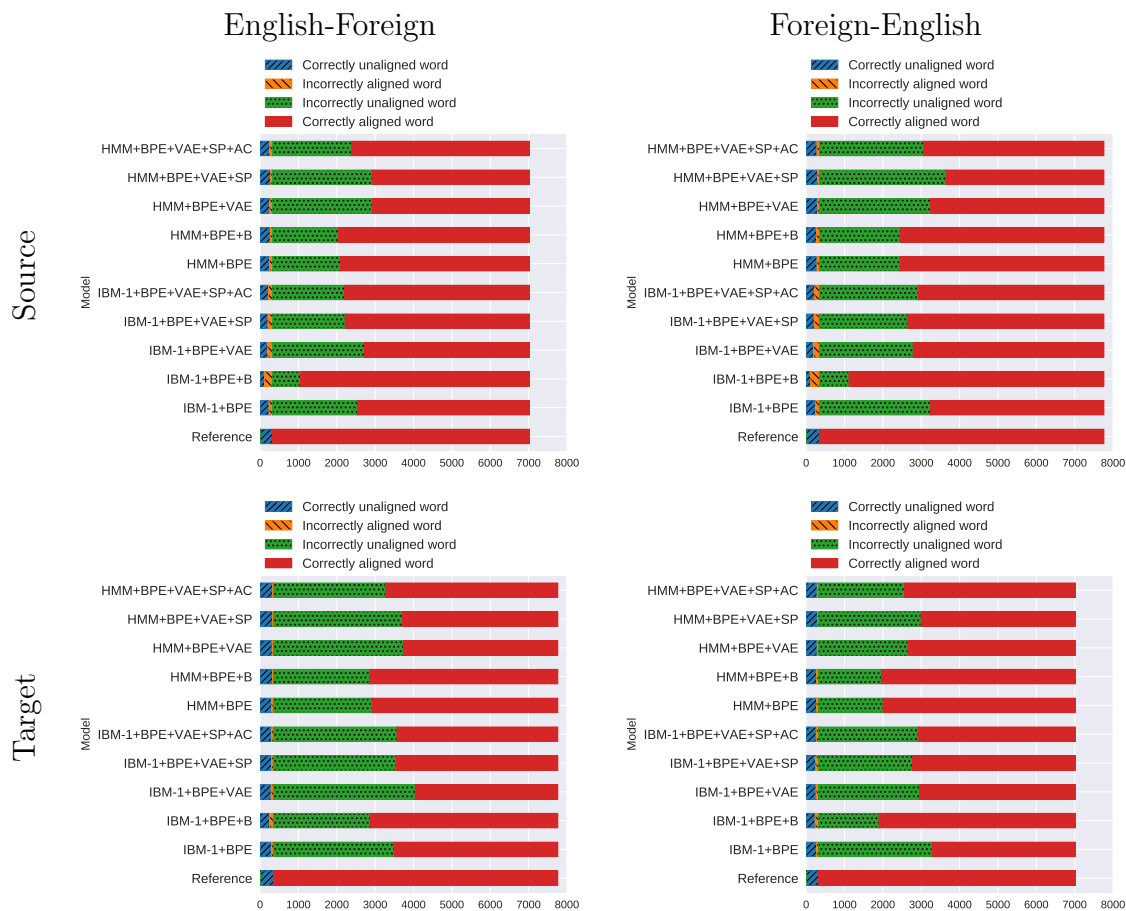


Figure D.7: Results of our variational models: Unaligned words for English-French

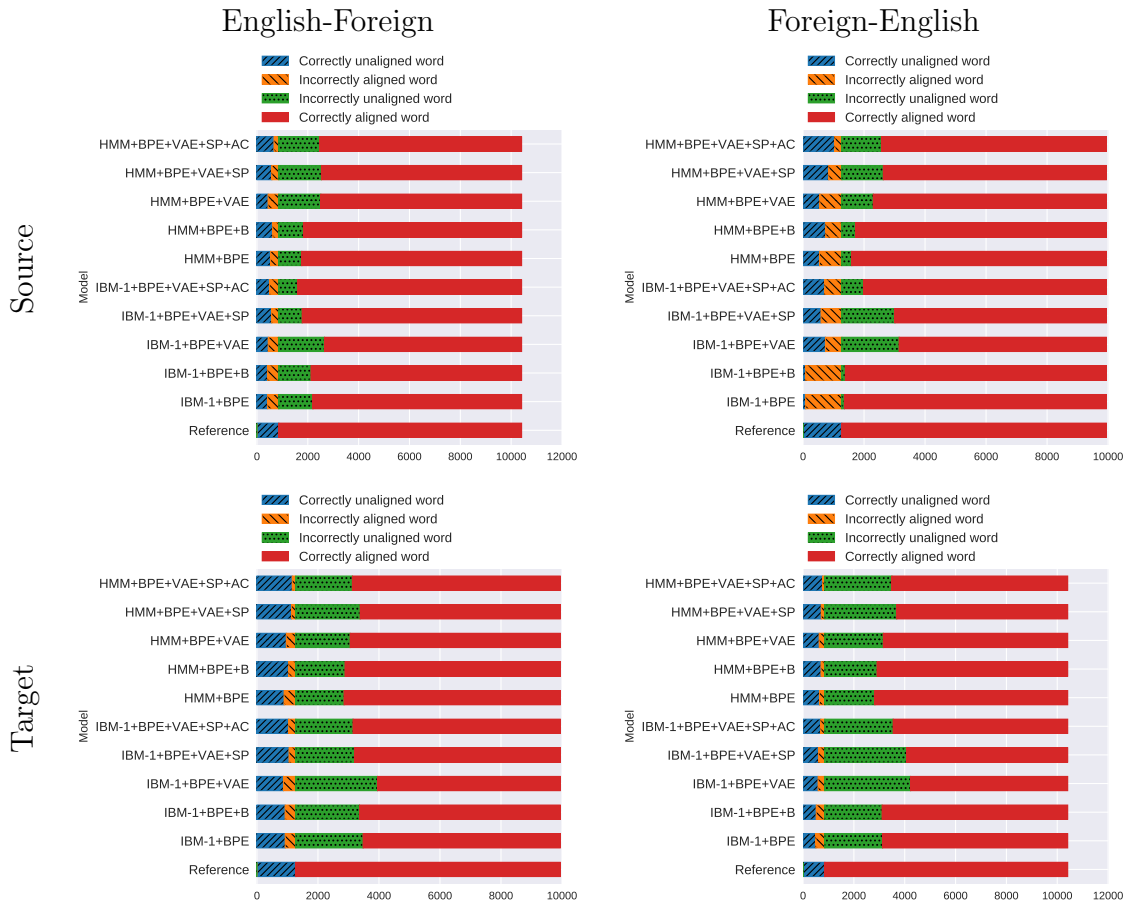


Figure D.8: Results of our variational models: Unaligned words for English-German

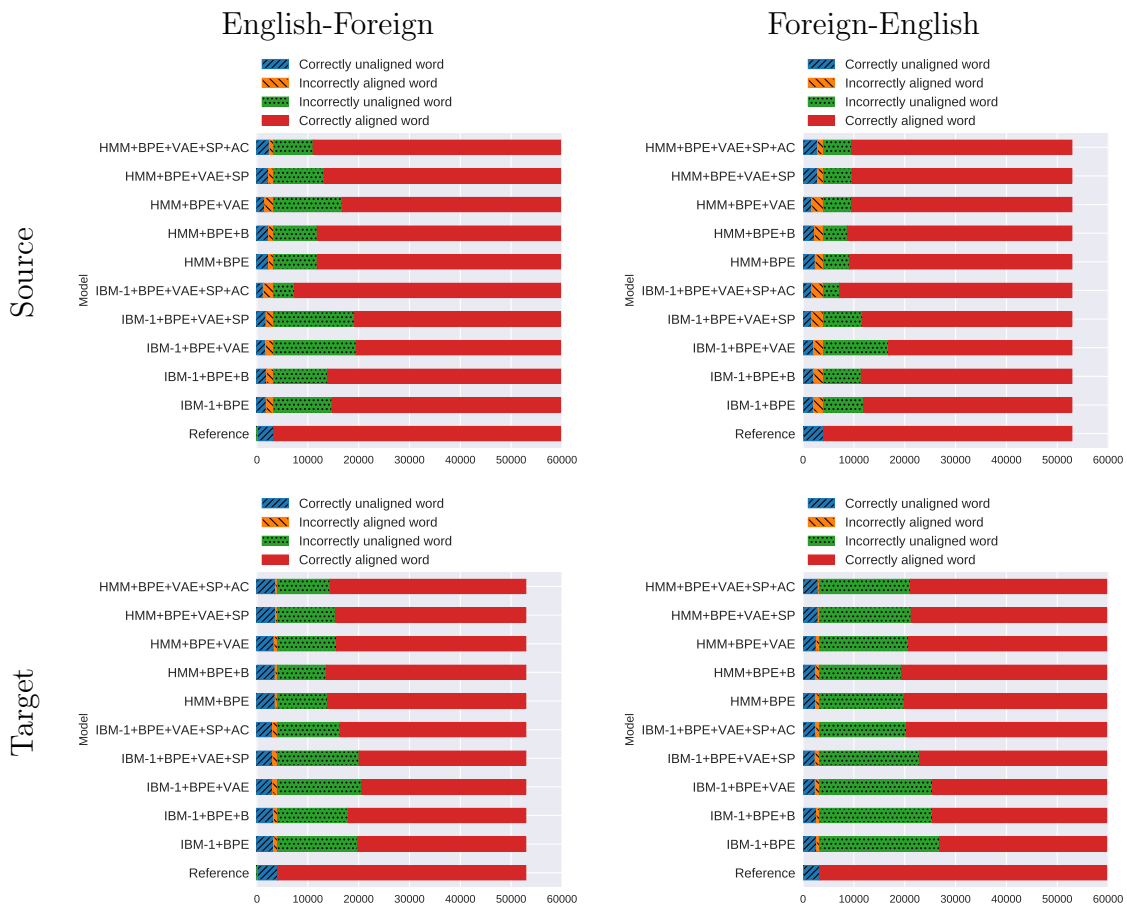


Figure D.9: Results of our variational models: Unaligned words for English-Czech

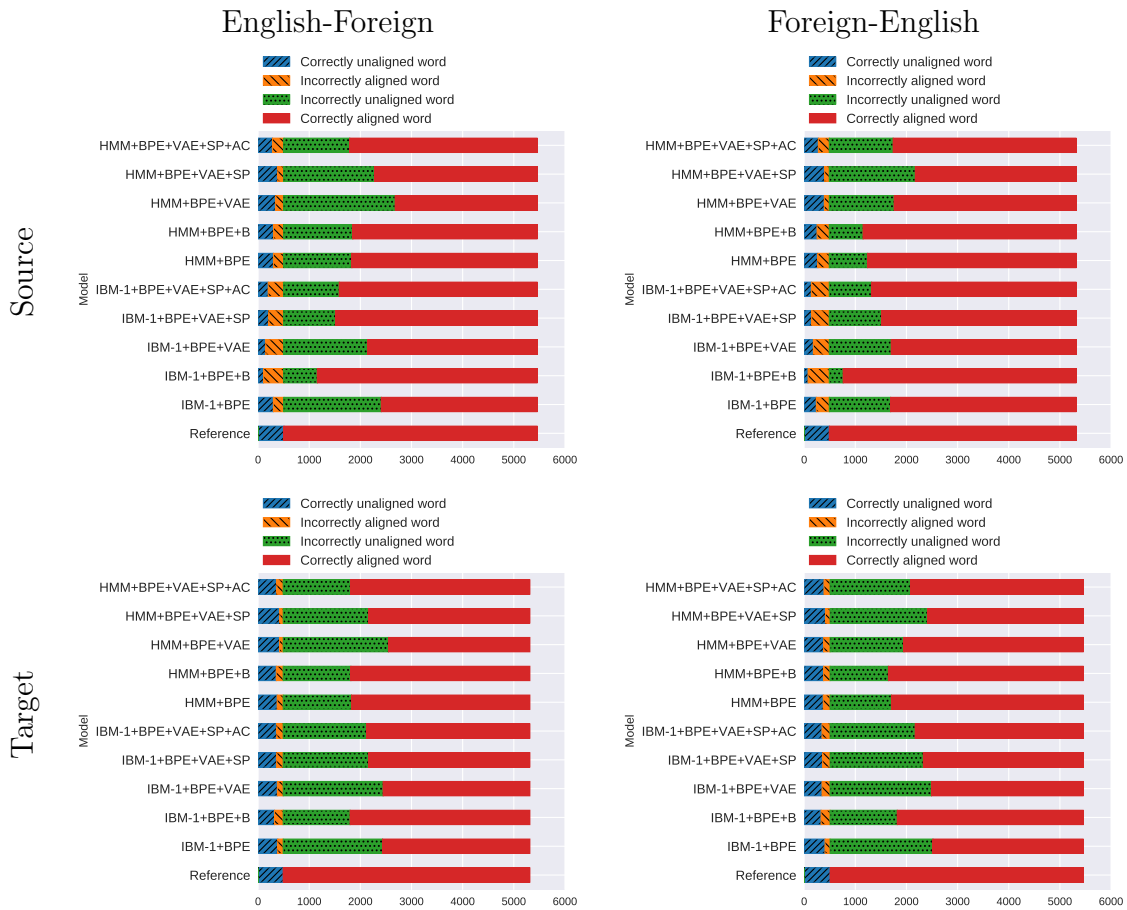


Figure D.10: Results of our variational models: Unaligned words for English-Romanian

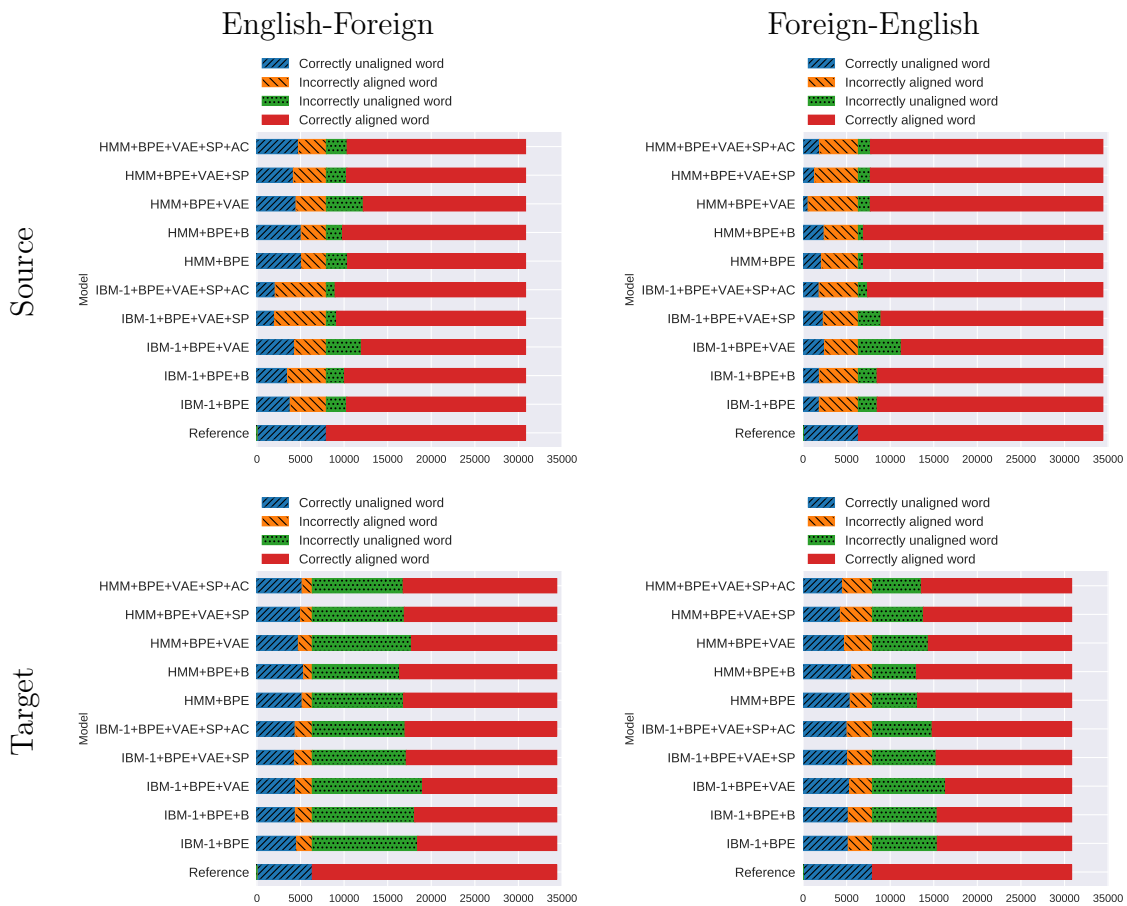


Figure D.11: Results of our variational models: Unaligned words for English-Japanese

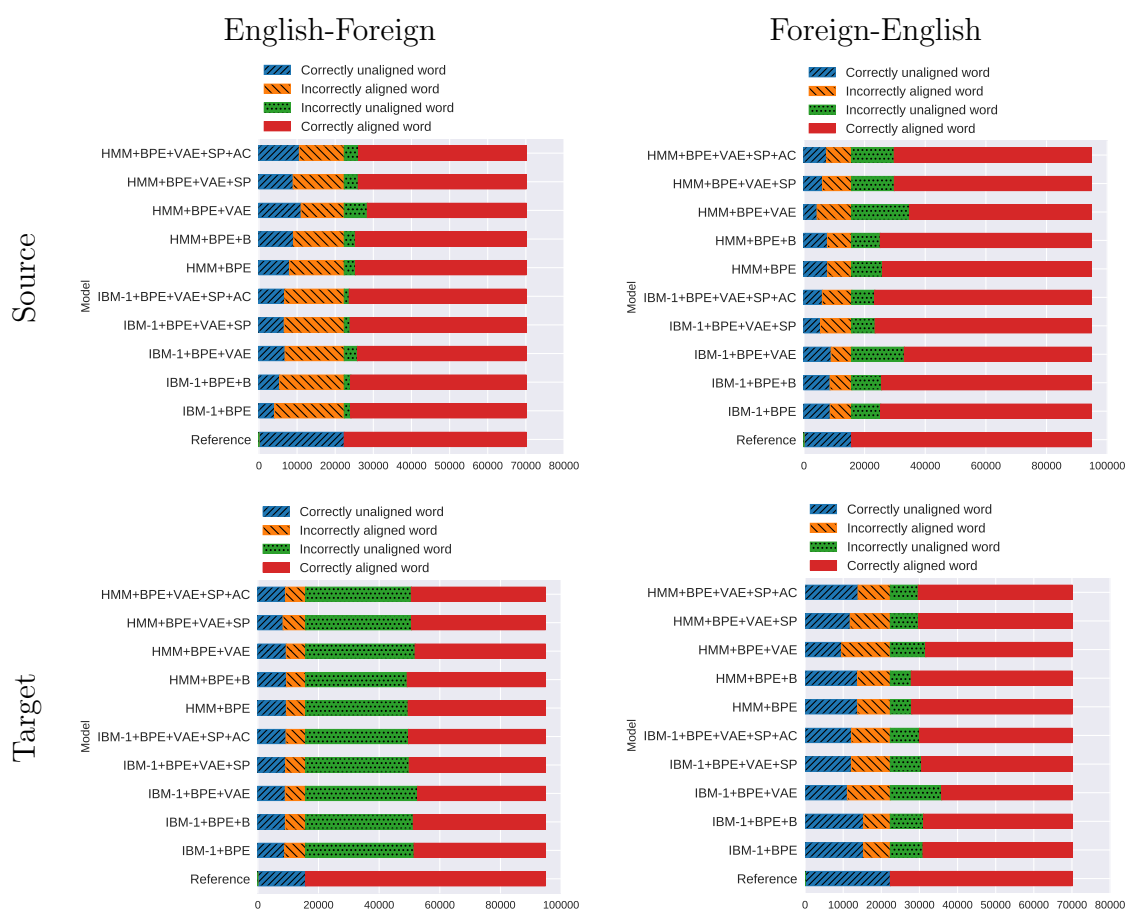


Figure D.12: Results of our variational models: Unaligned words for English-Vietnamese

D.4 Alignment types

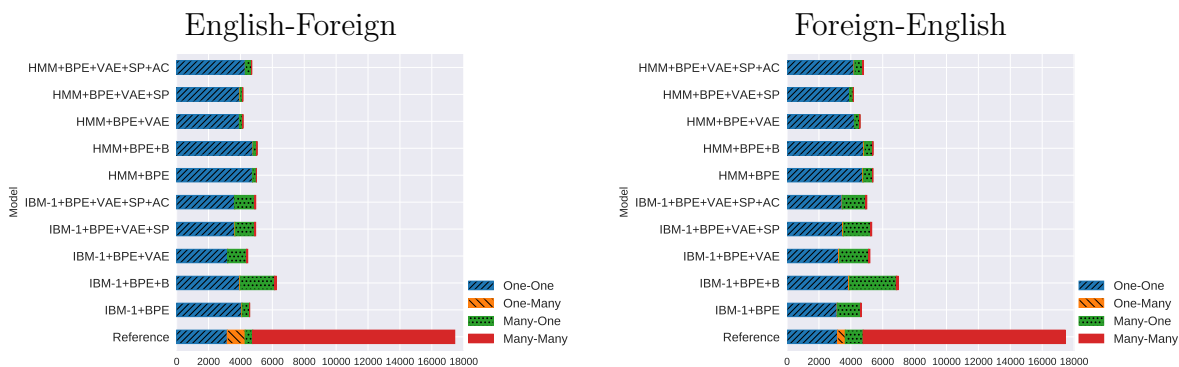


Figure D.13: Results of our variational models: Alignment types of English-French



Figure D.14: Results of our variational models: Alignment types of English-German

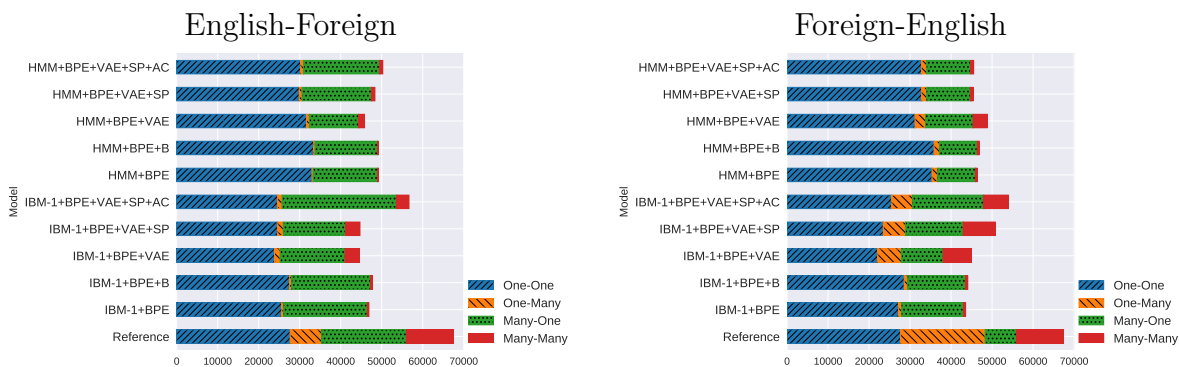


Figure D.15: Results of our variational models: Alignment types of English-Czech

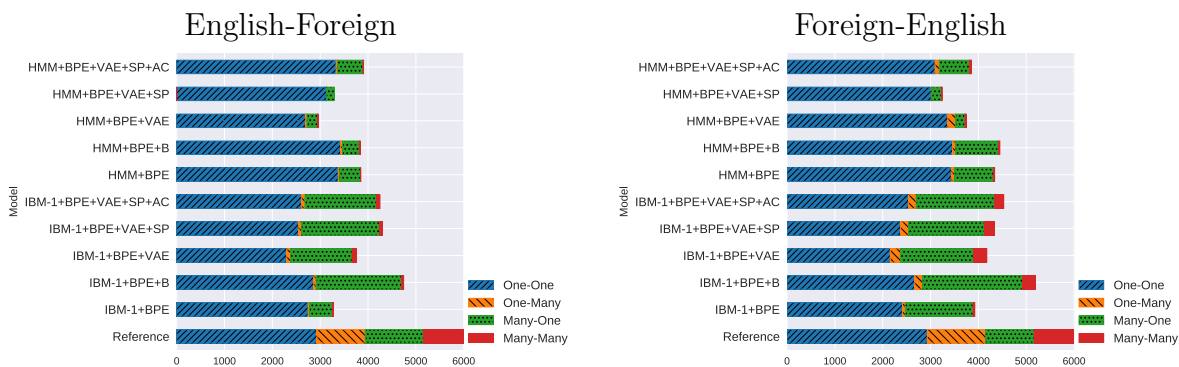


Figure D.16: Results of our variational models: Alignment types of English-Romanian

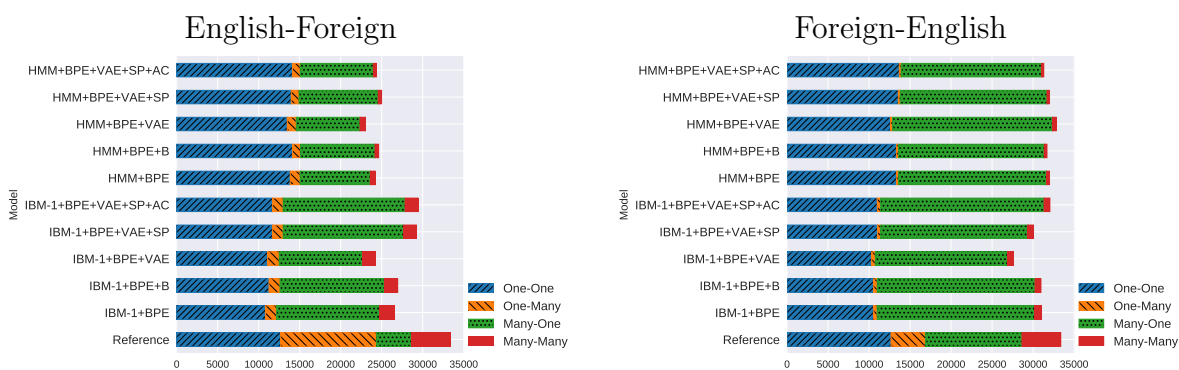


Figure D.17: Results of our variational models: Alignment types of English-Japanese

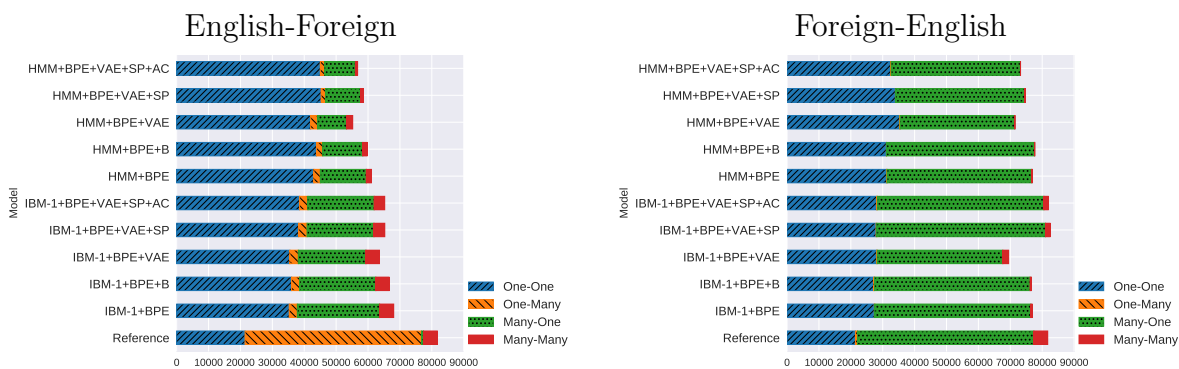


Figure D.18: Results of our variational models: Alignment types of English-Vietnamese

D.4.1 Correct/incorrect links as a function of alignment types

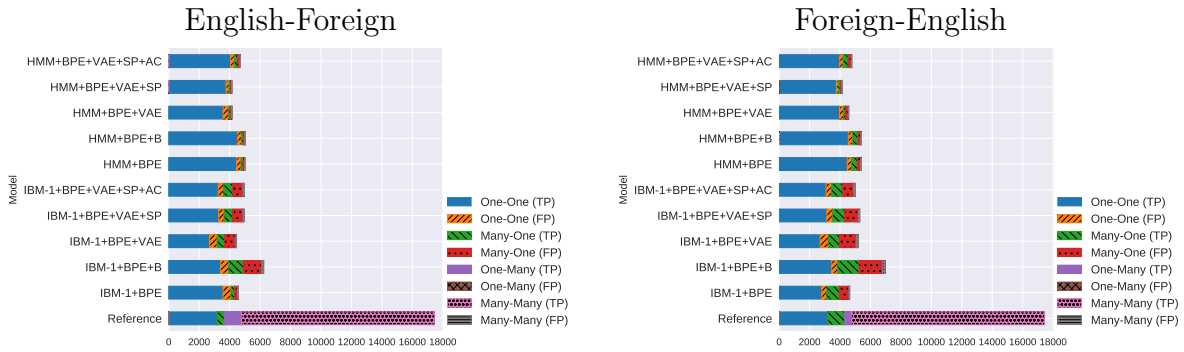


Figure D.19: Results of our variational models: Alignment types for English-French

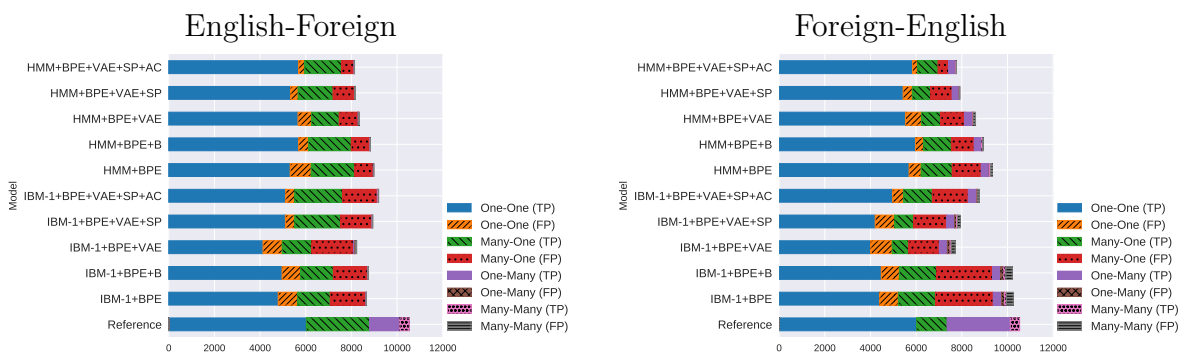


Figure D.20: Results of our variational models: Alignment types for English-German

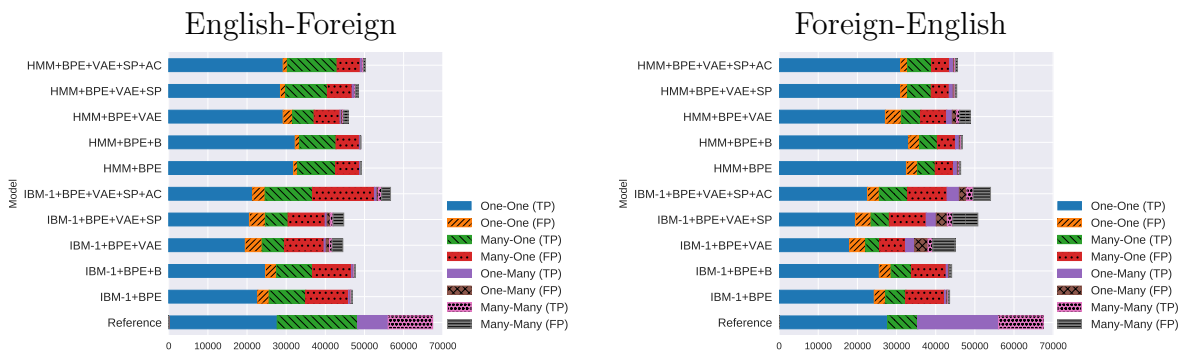


Figure D.21: Results of our variational models: Alignment types for English-Czech

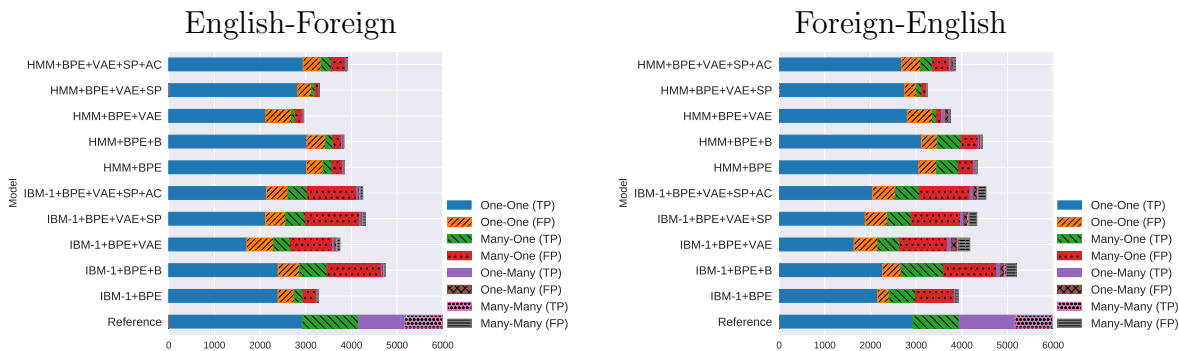


Figure D.22: Results of our variational models: Alignment types for English-Romanian

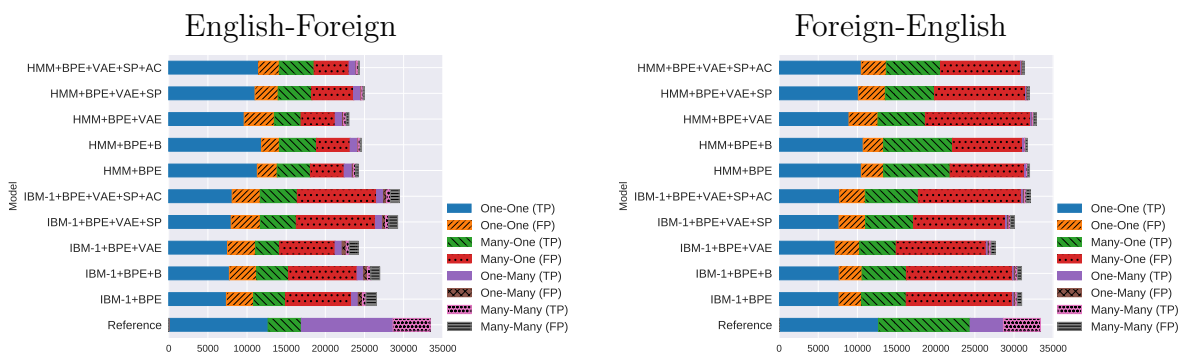


Figure D.23: Results of our variational models: Alignment types for English-Japanese

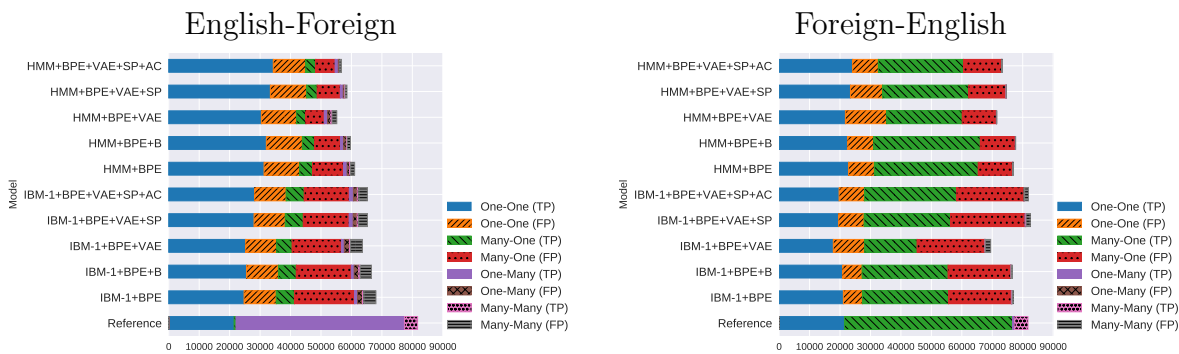


Figure D.24: Results of our variational models: Alignment types for English-Vietnamese

D.5 Jump widths

D.5.1 Jump width statistics

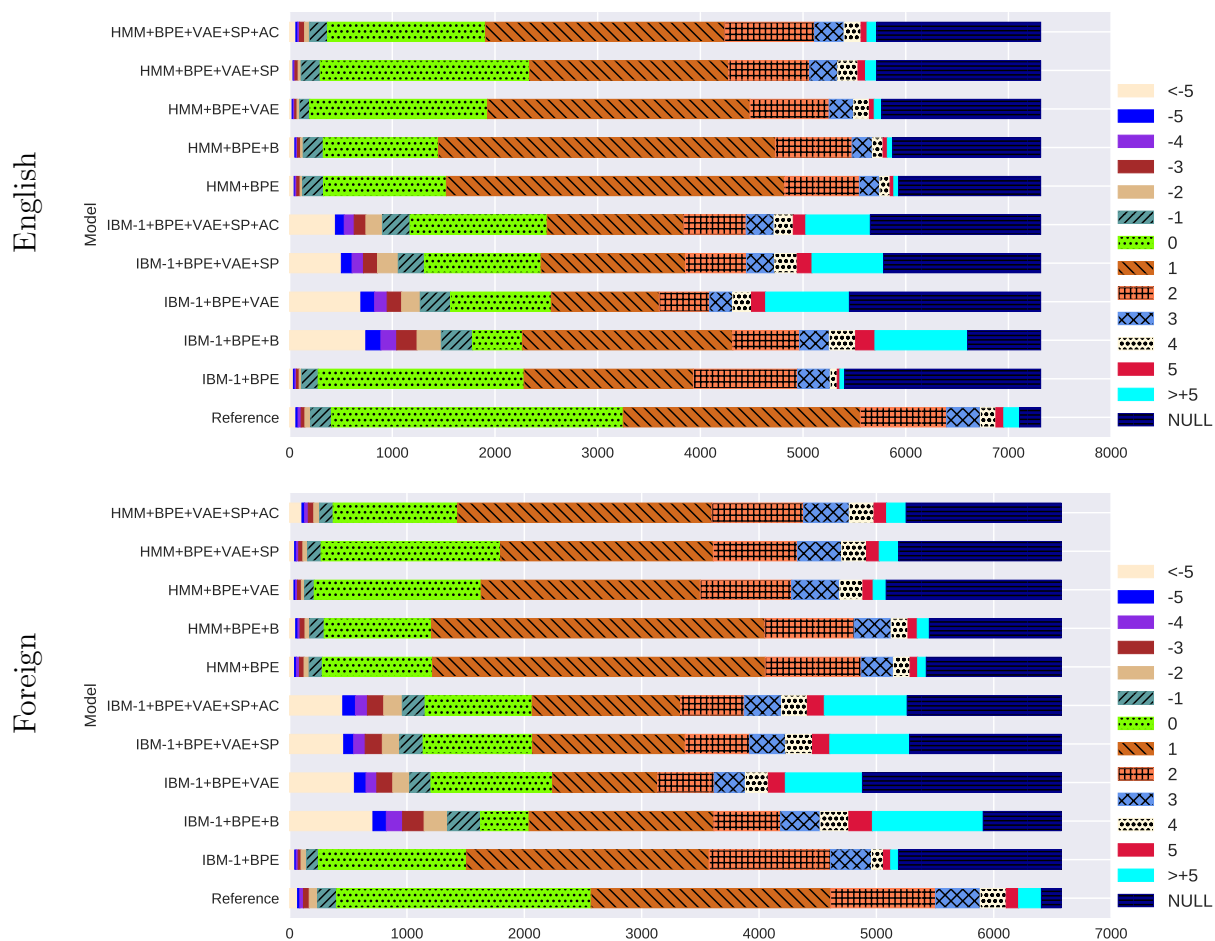


Figure D.25: Variational models: Jump widths for English words in the direction French-English and jump widths for foreign words in the direction English-French

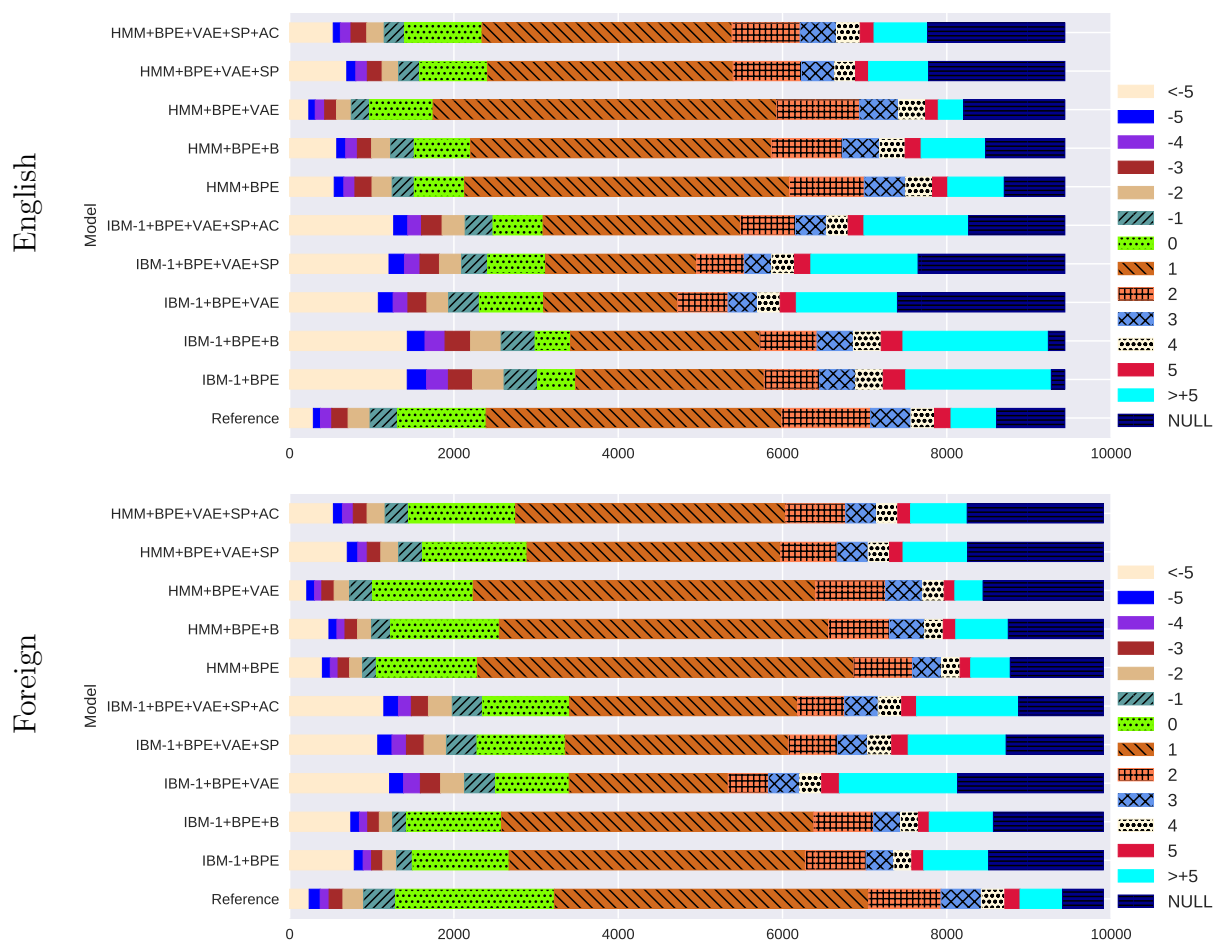


Figure D.26: Variational models: Jump widths for English words in the direction German-English and jump widths for foreign words in the direction English-German

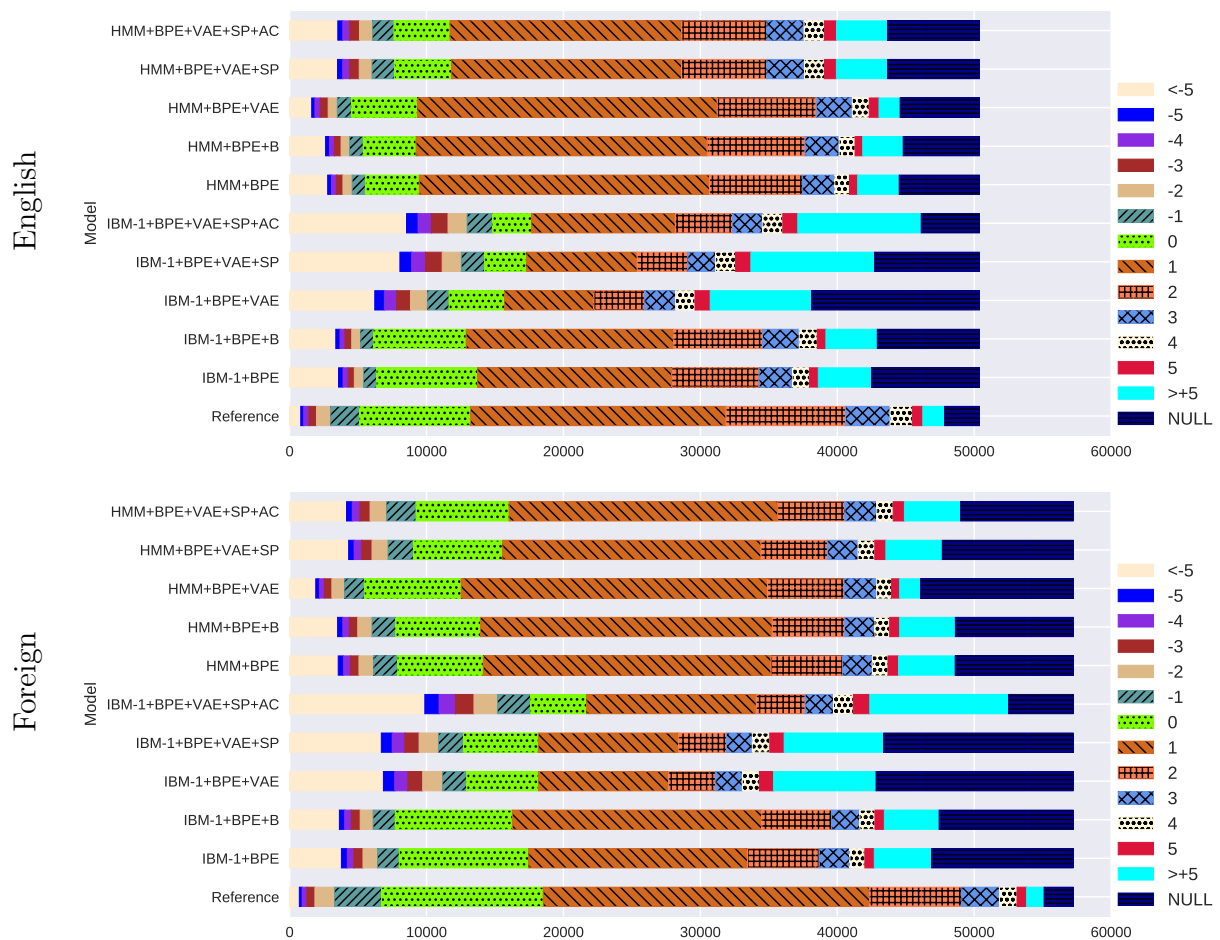


Figure D.27: Variational models: Jump widths for English words in the direction Czech-English and jump widths for foreign words in the direction English-Czech

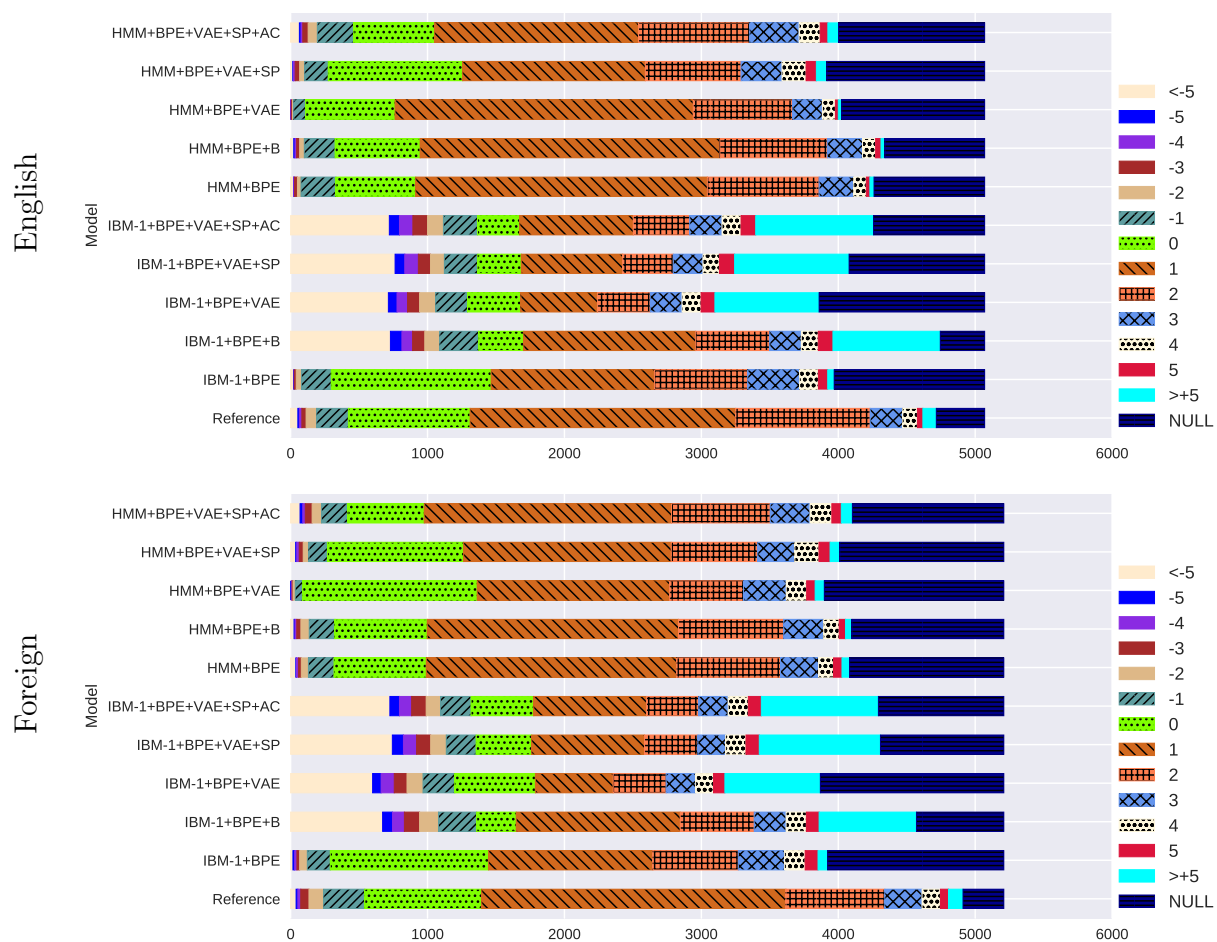


Figure D.28: Variational models: Jump widths for English words in the direction Romanian-English and jump widths for foreign words in the direction English-Romanian

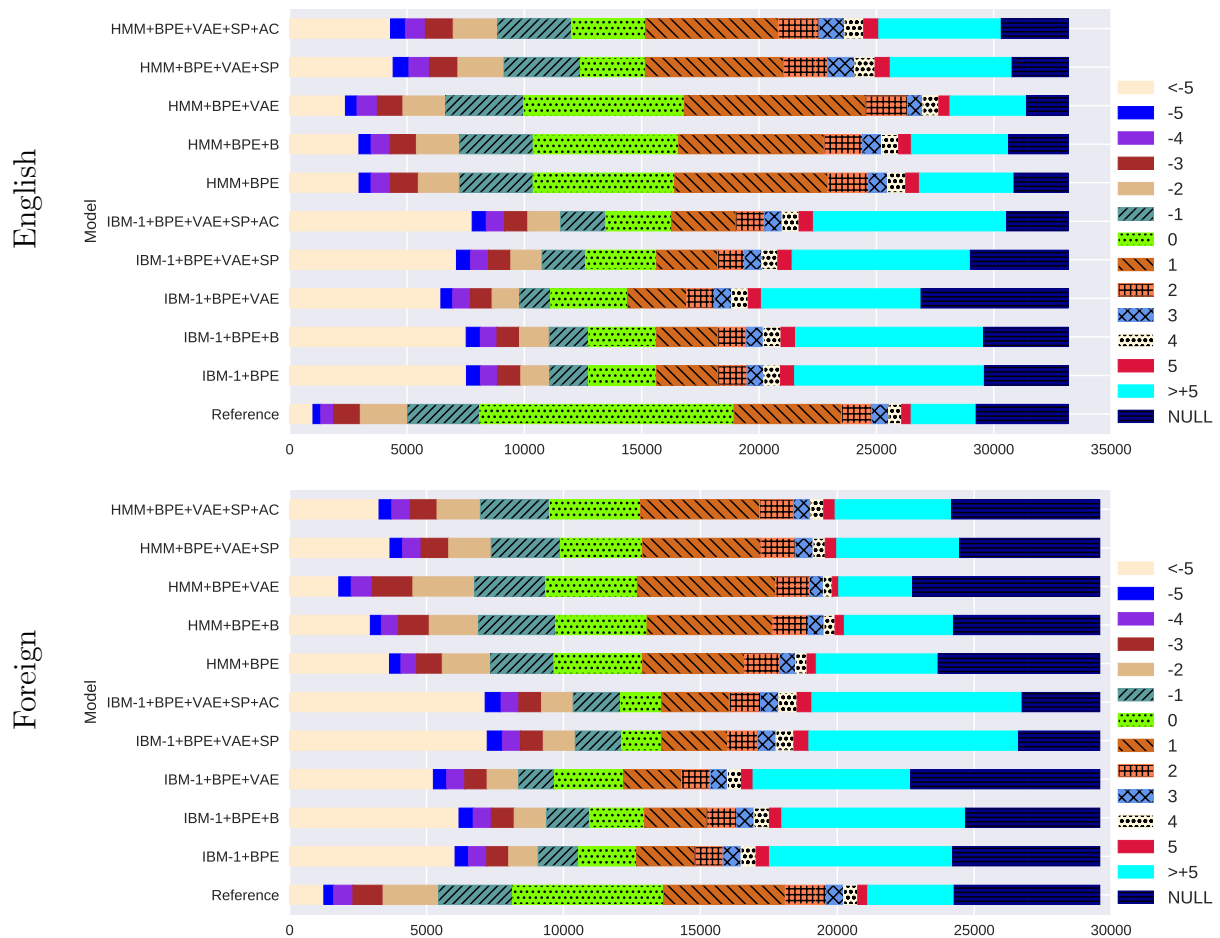


Figure D.29: Variational models: Jump widths for English words in the direction Japanese-English and jump widths for foreign words in the direction English-Japanese

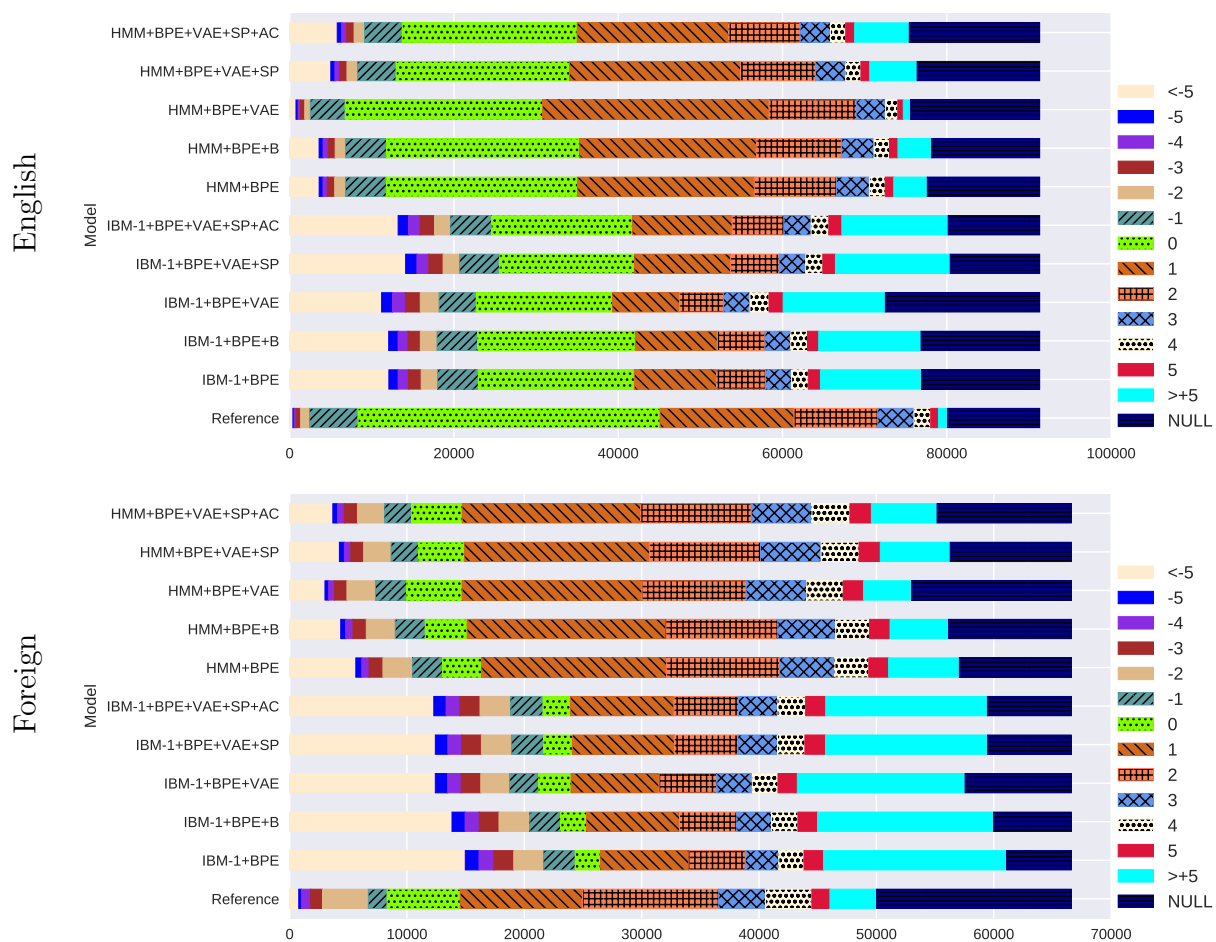
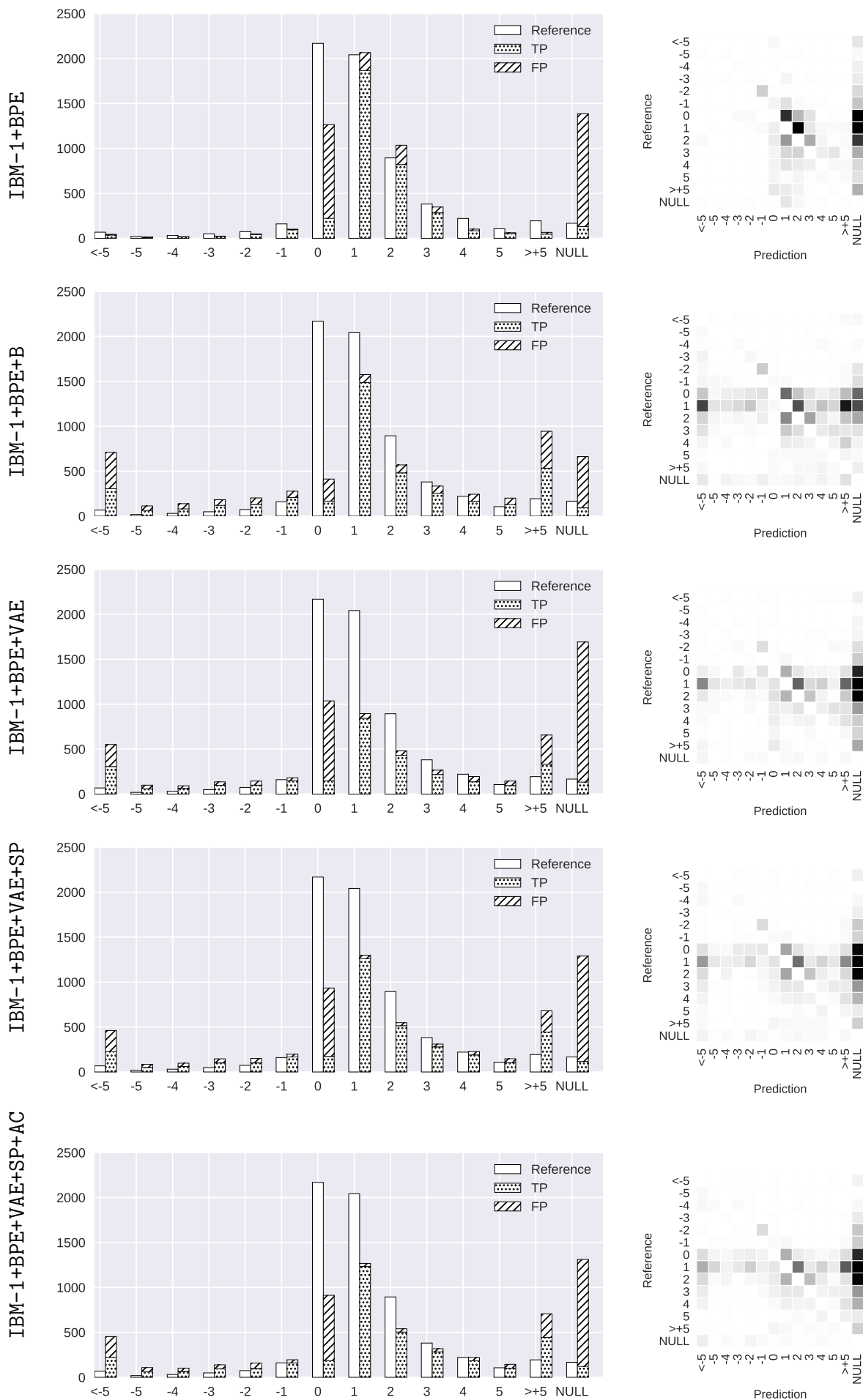


Figure D.30: Variational models: Jump widths for English words in the direction Vietnamese-English and jump widths for foreign words in the direction English-Vietnamese

D.5.2 Correct and incorrect jumps



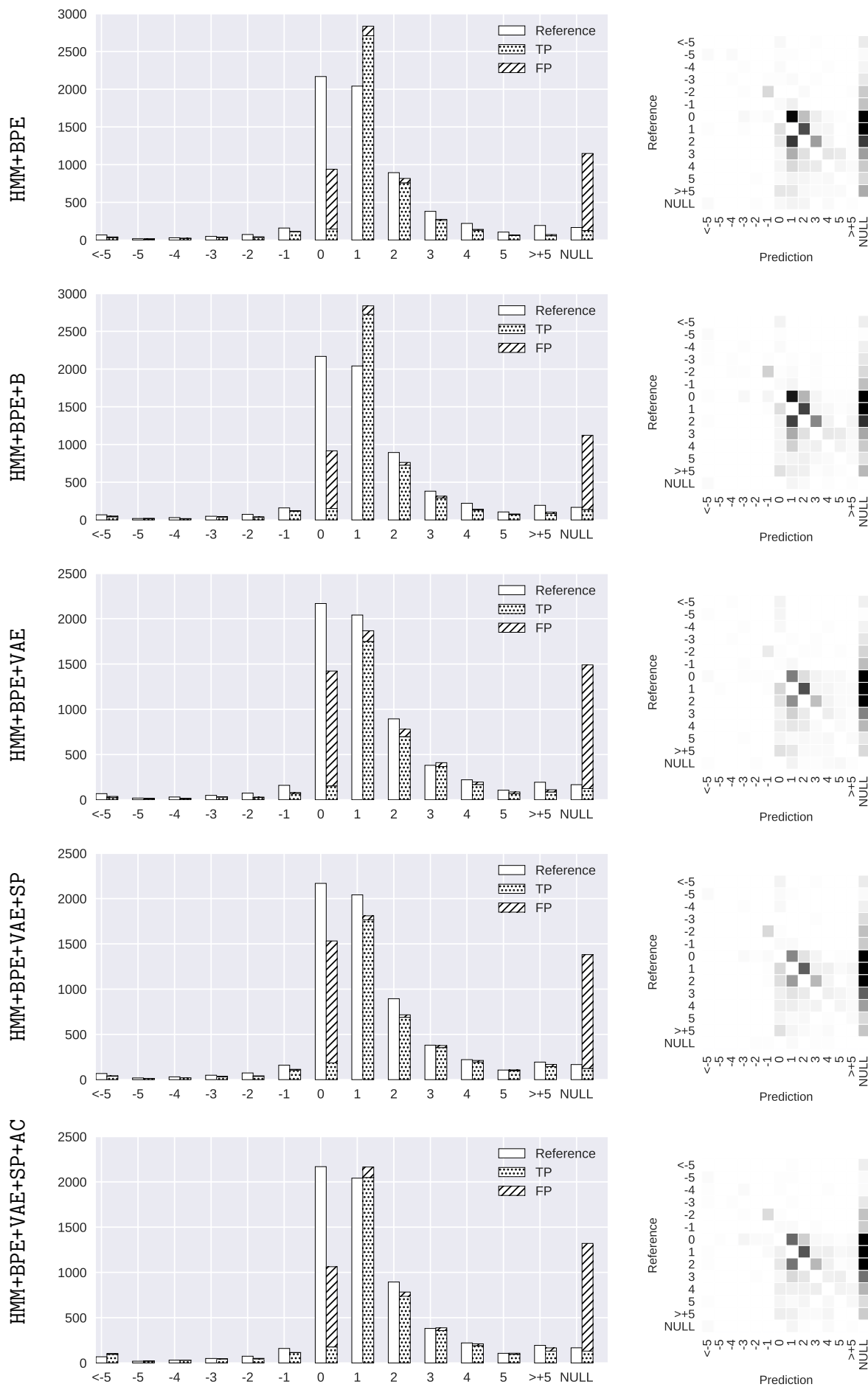
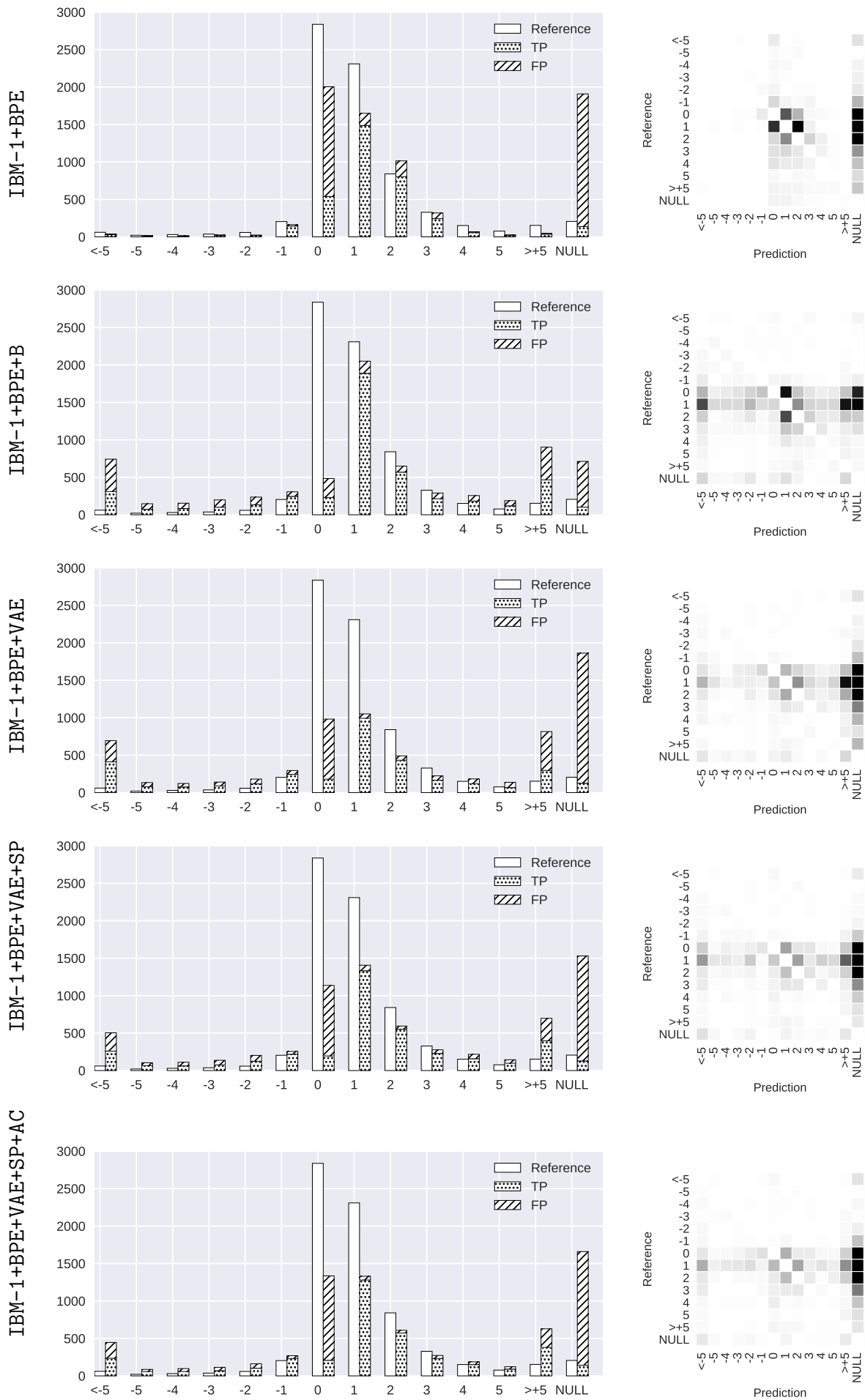


Figure D.31: Variational models for the direction English-French: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the right

graph: The darker the cell, the greater the number of confusions.



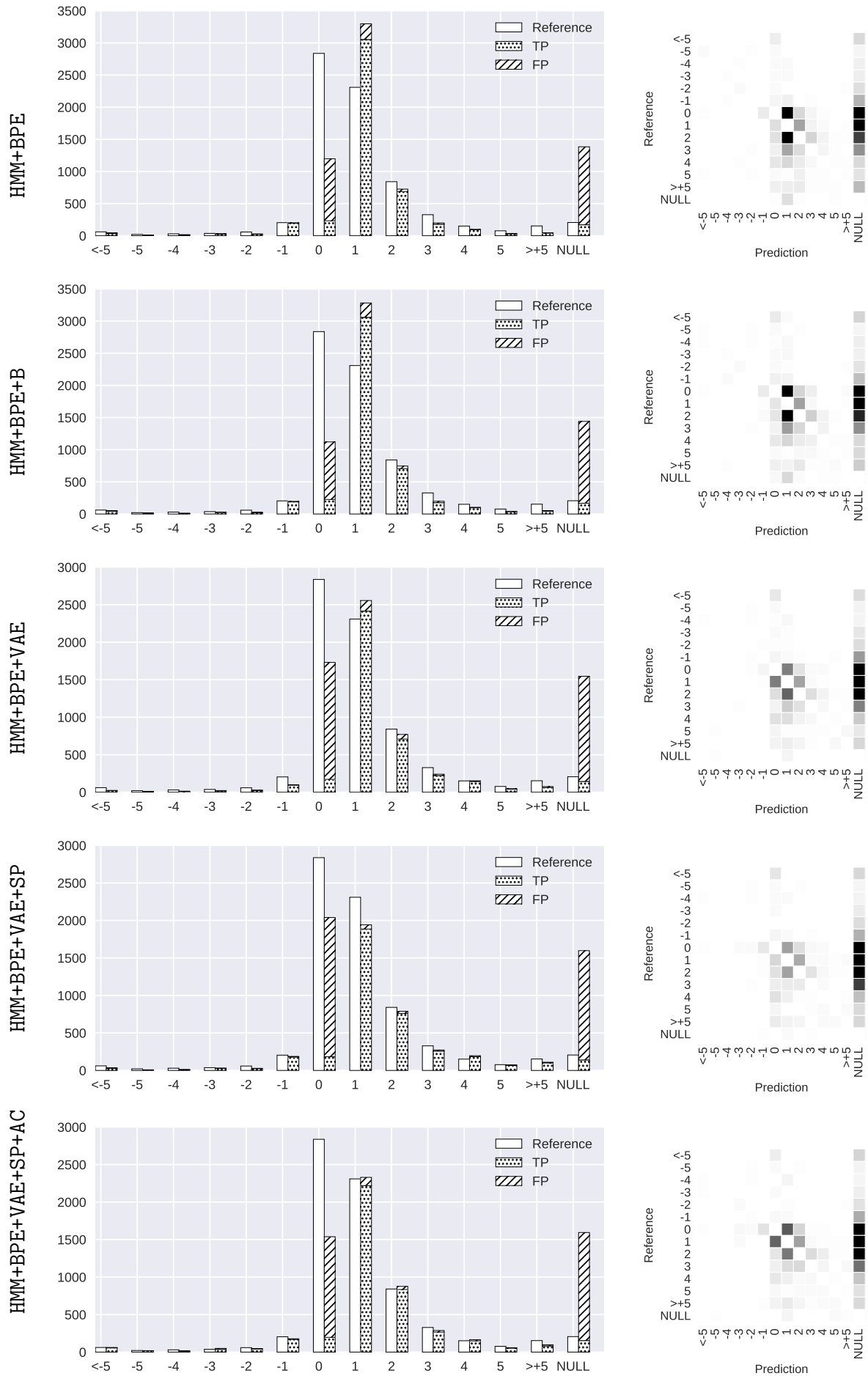
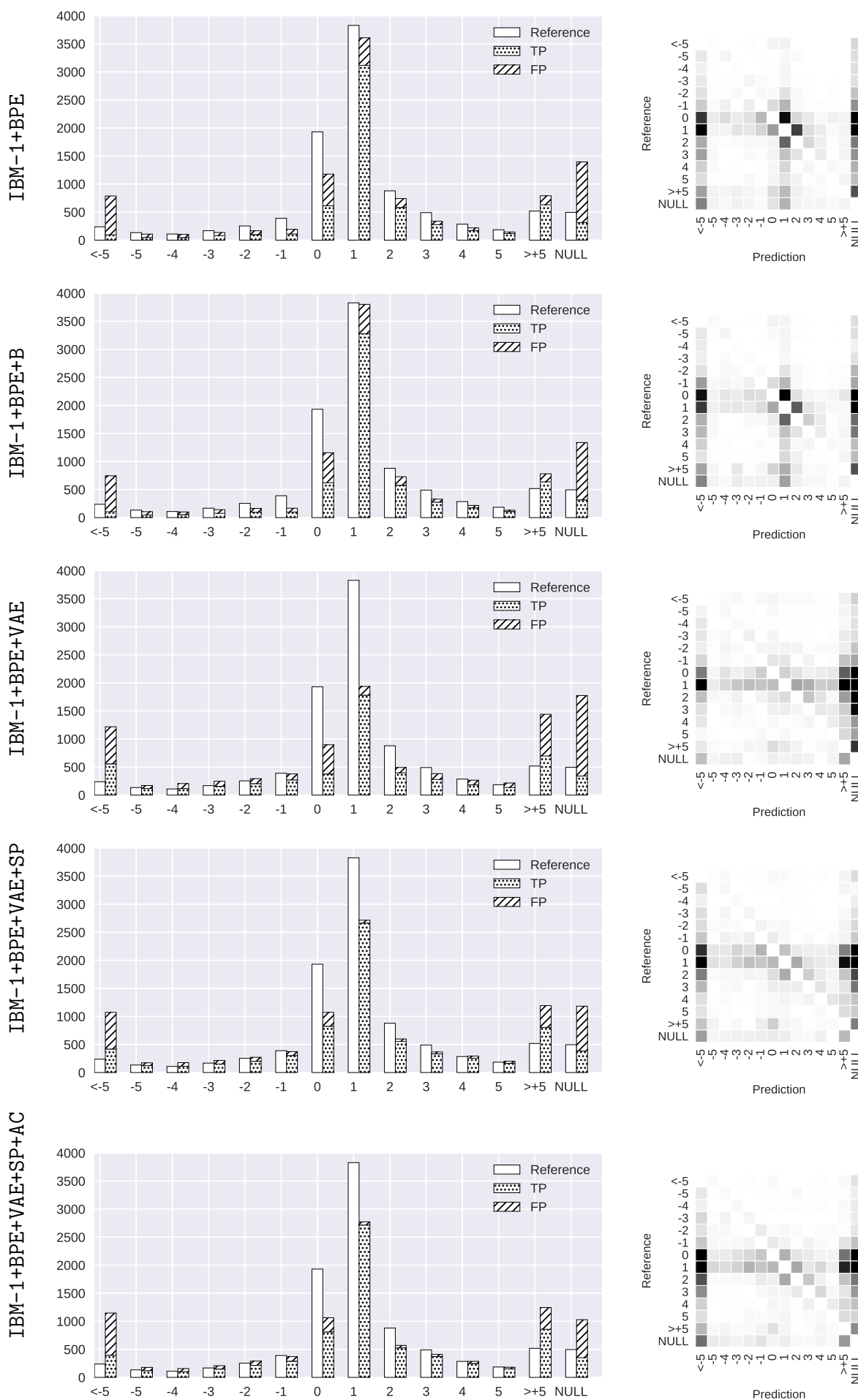


Figure D.32: Variational models for the direction French-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the right

graph: The darker the cell, the greater the number of confusions.



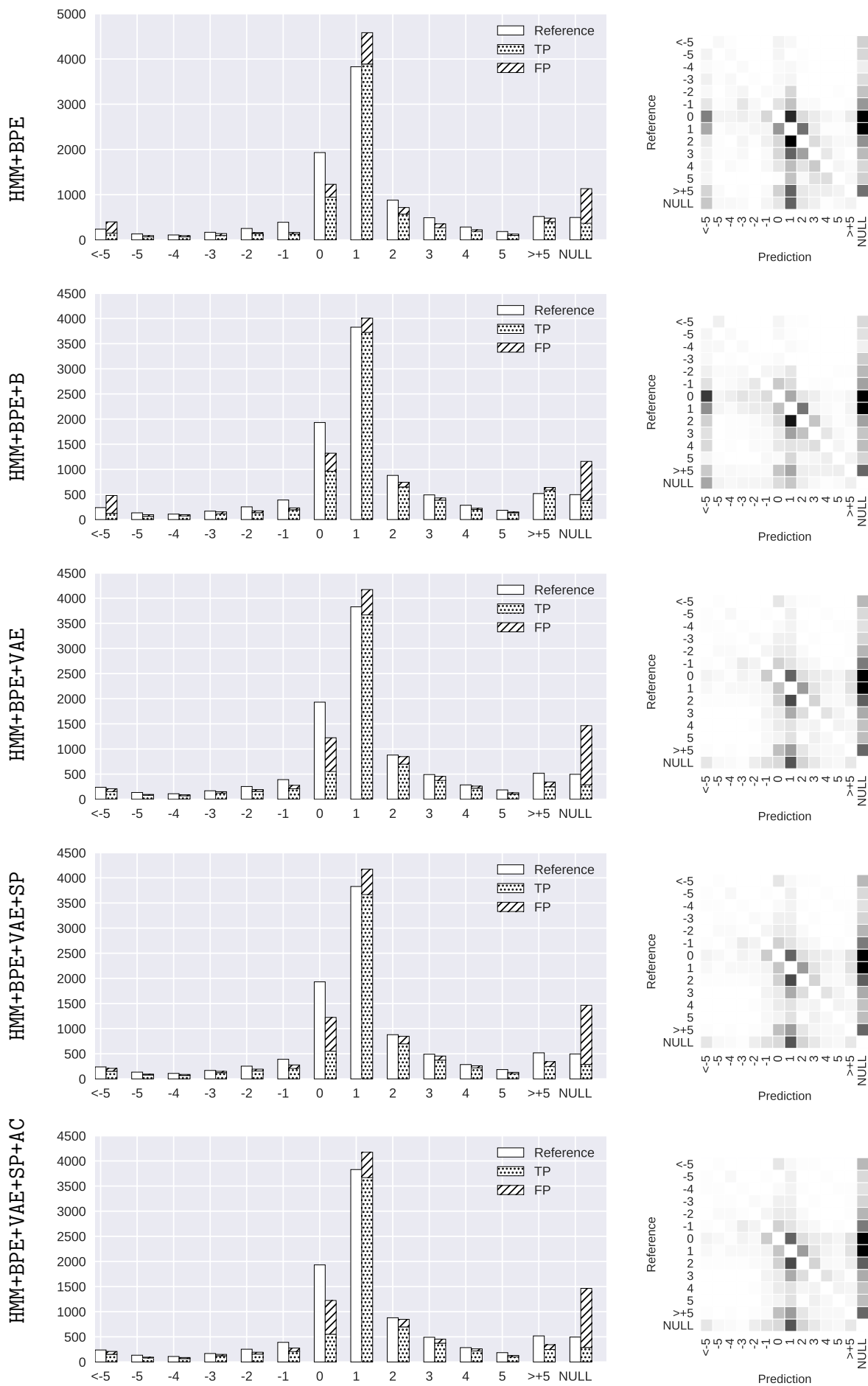
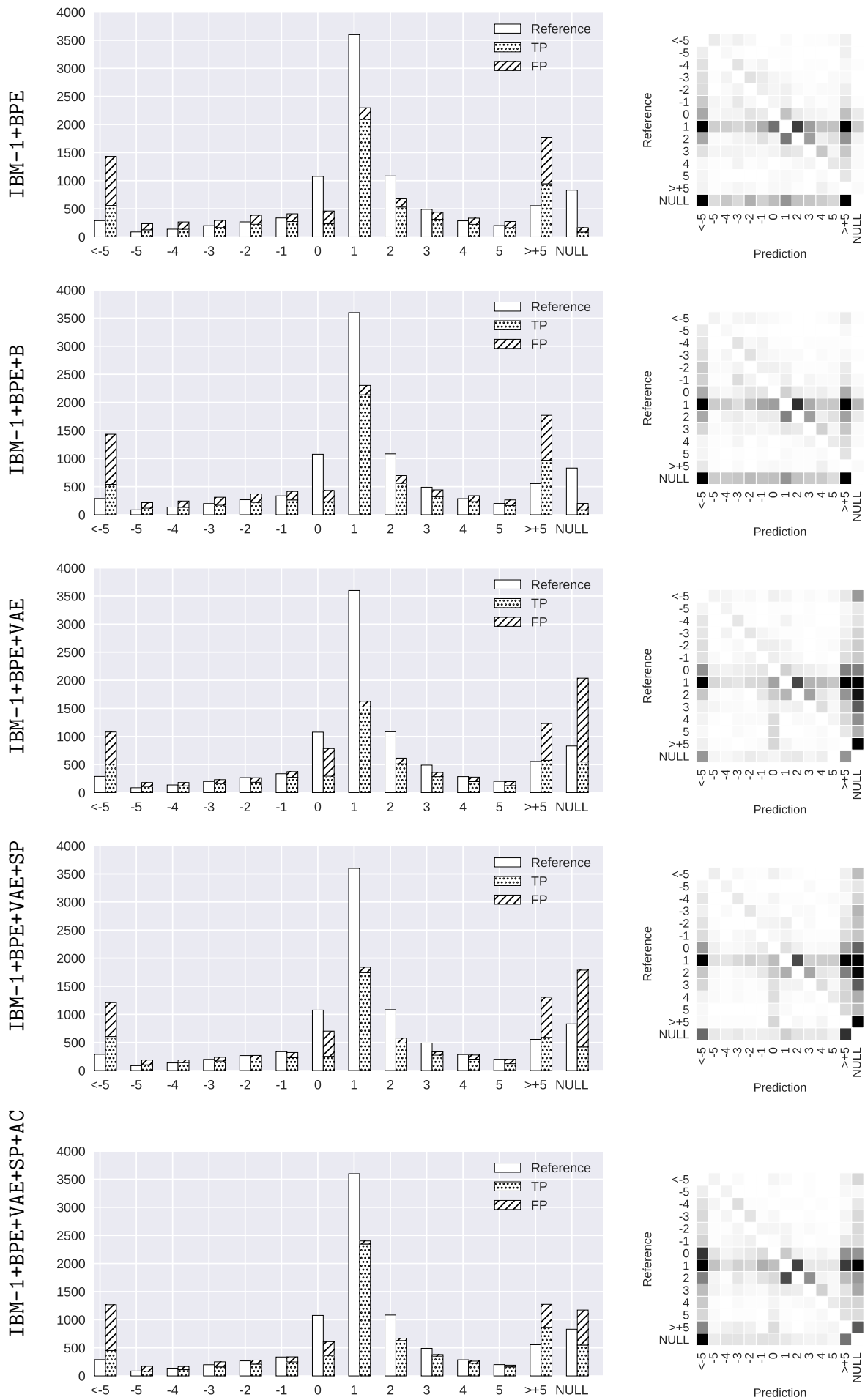


Figure D.33: Variational models for the direction English-German: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the right

graph: The darker the cell, the greater the number of confusions.



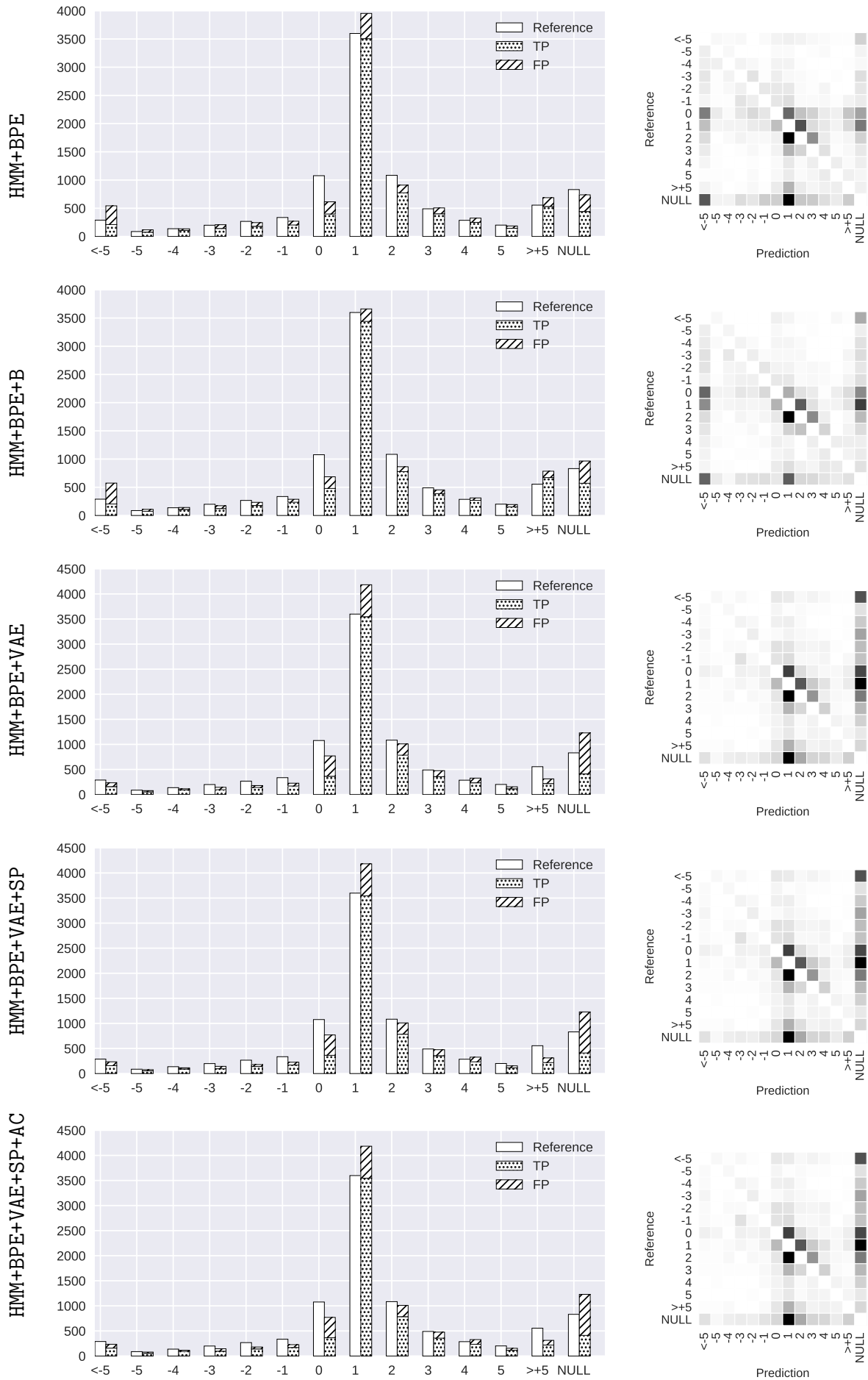
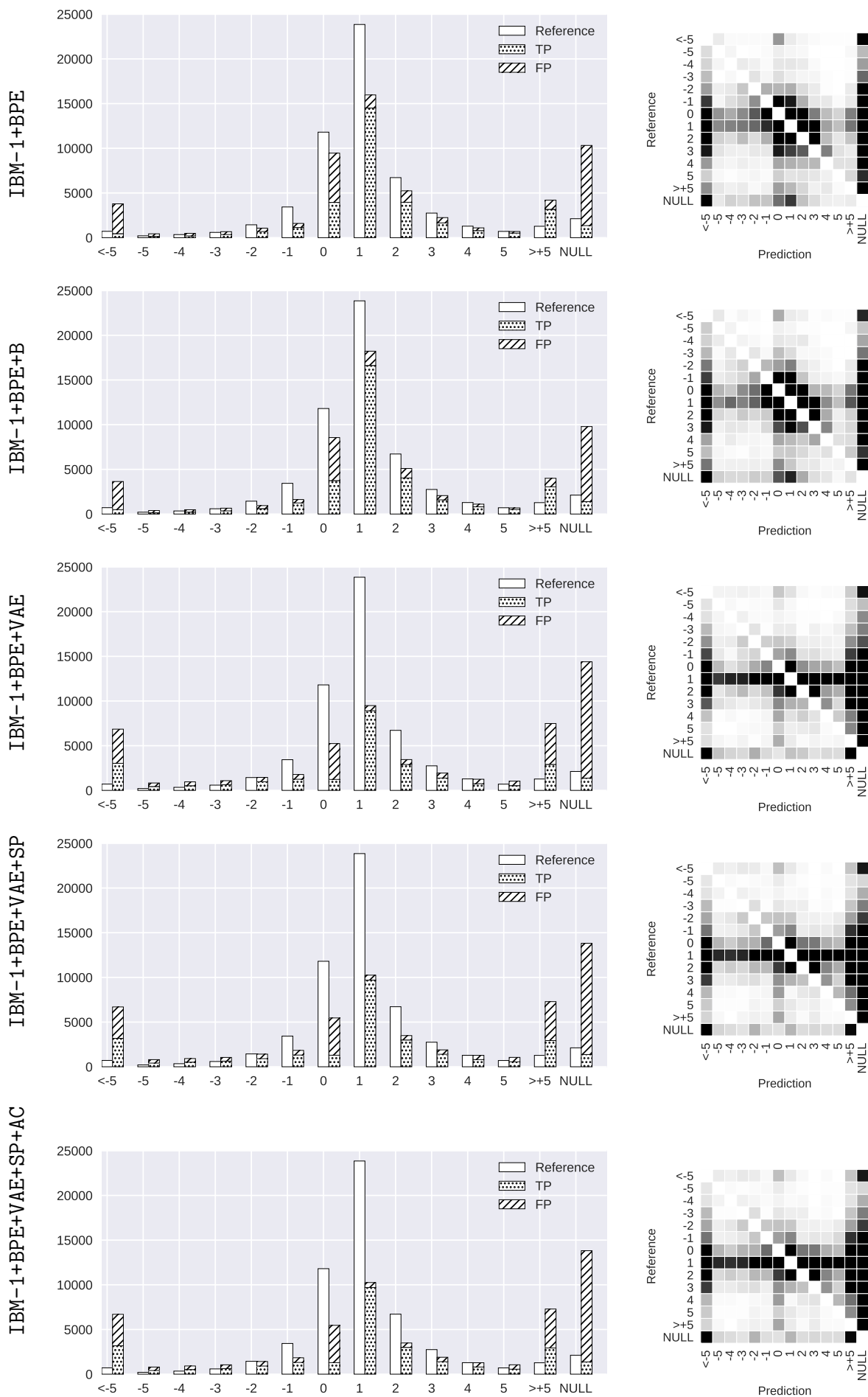


Figure D.34: Variational models for the direction German-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the right

graph: The darker the cell, the greater the number of confusions.



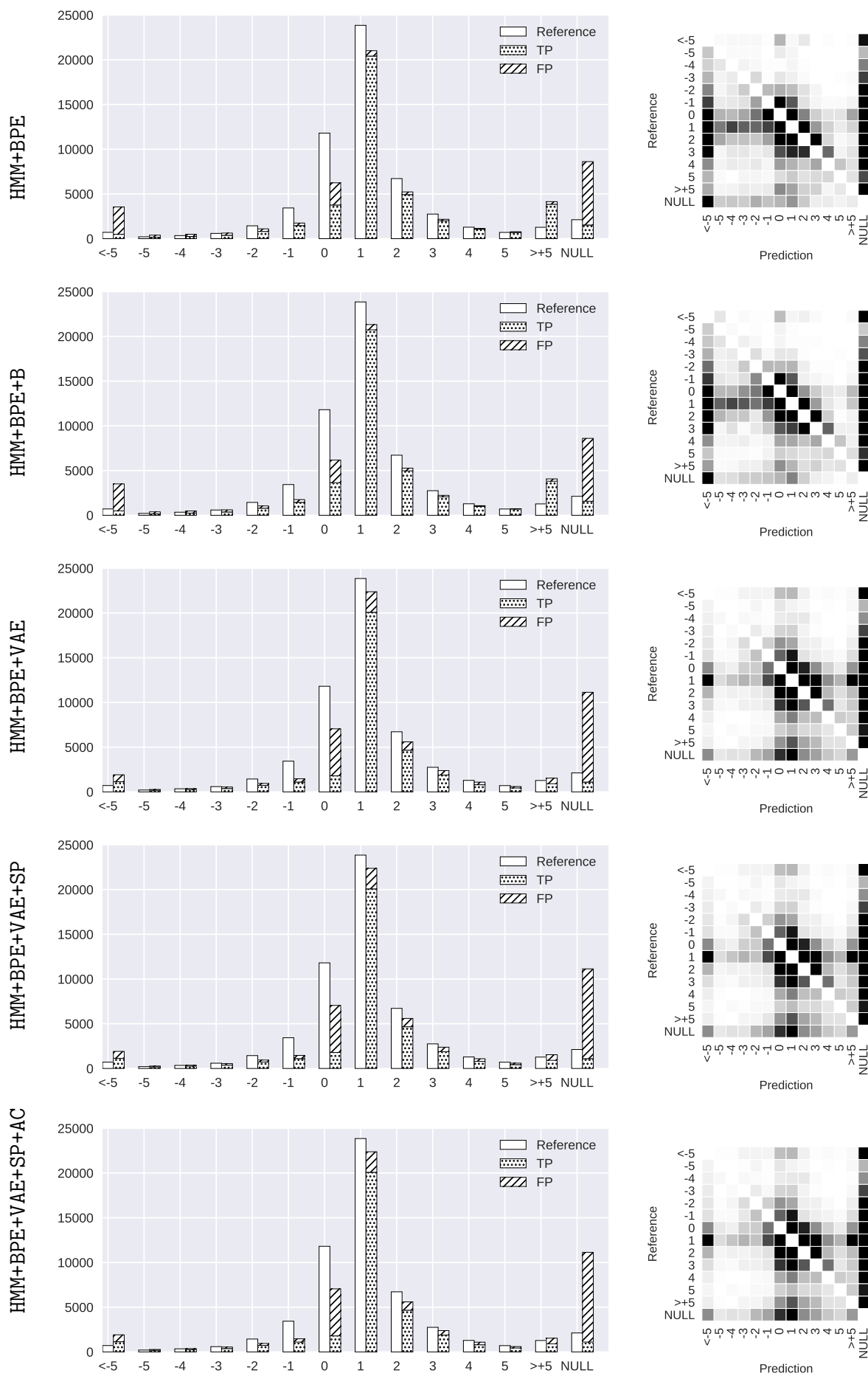
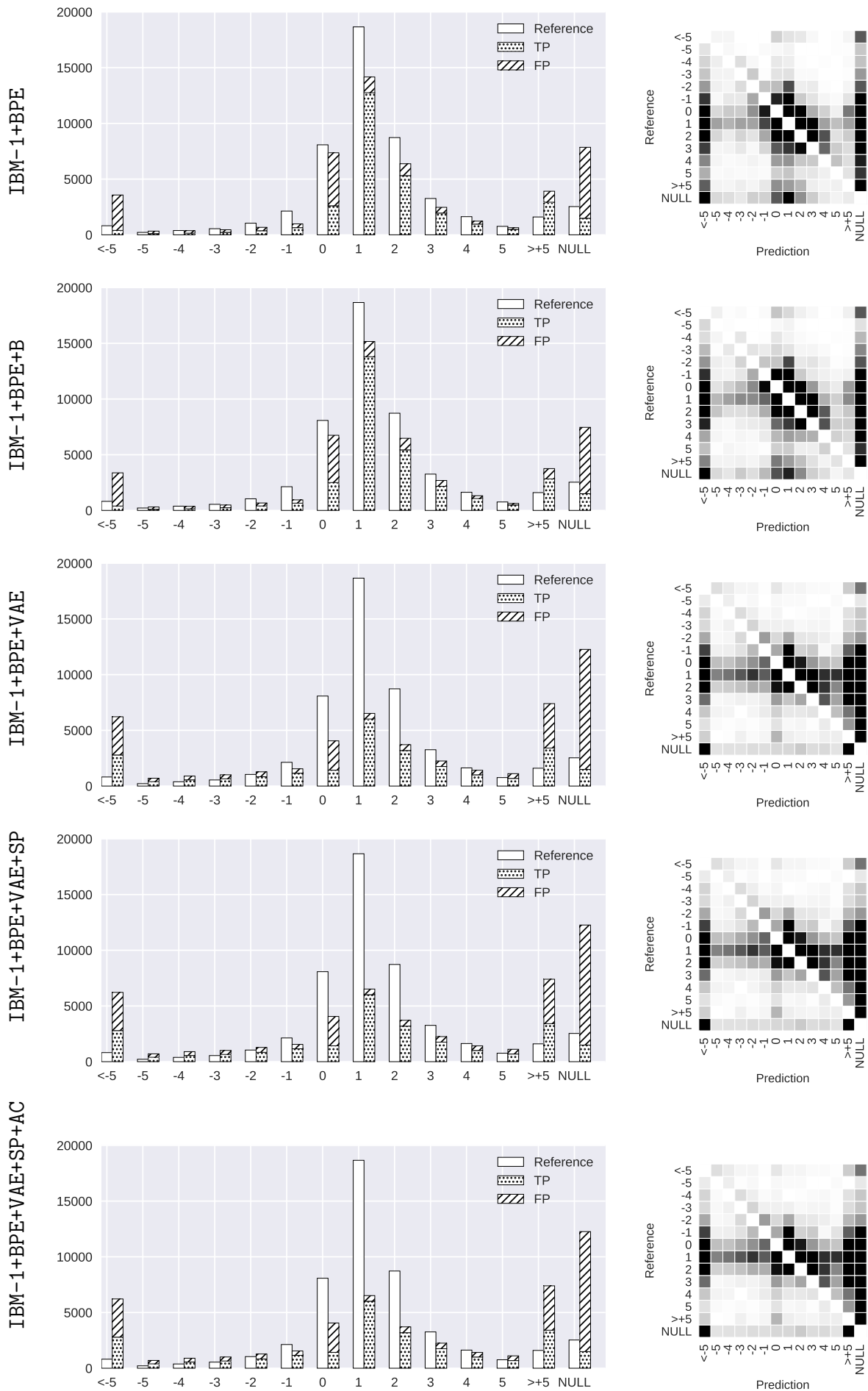


Figure D.35: Variational models for the direction English-Czech: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the right

graph: The darker the cell, the greater the number of confusions.



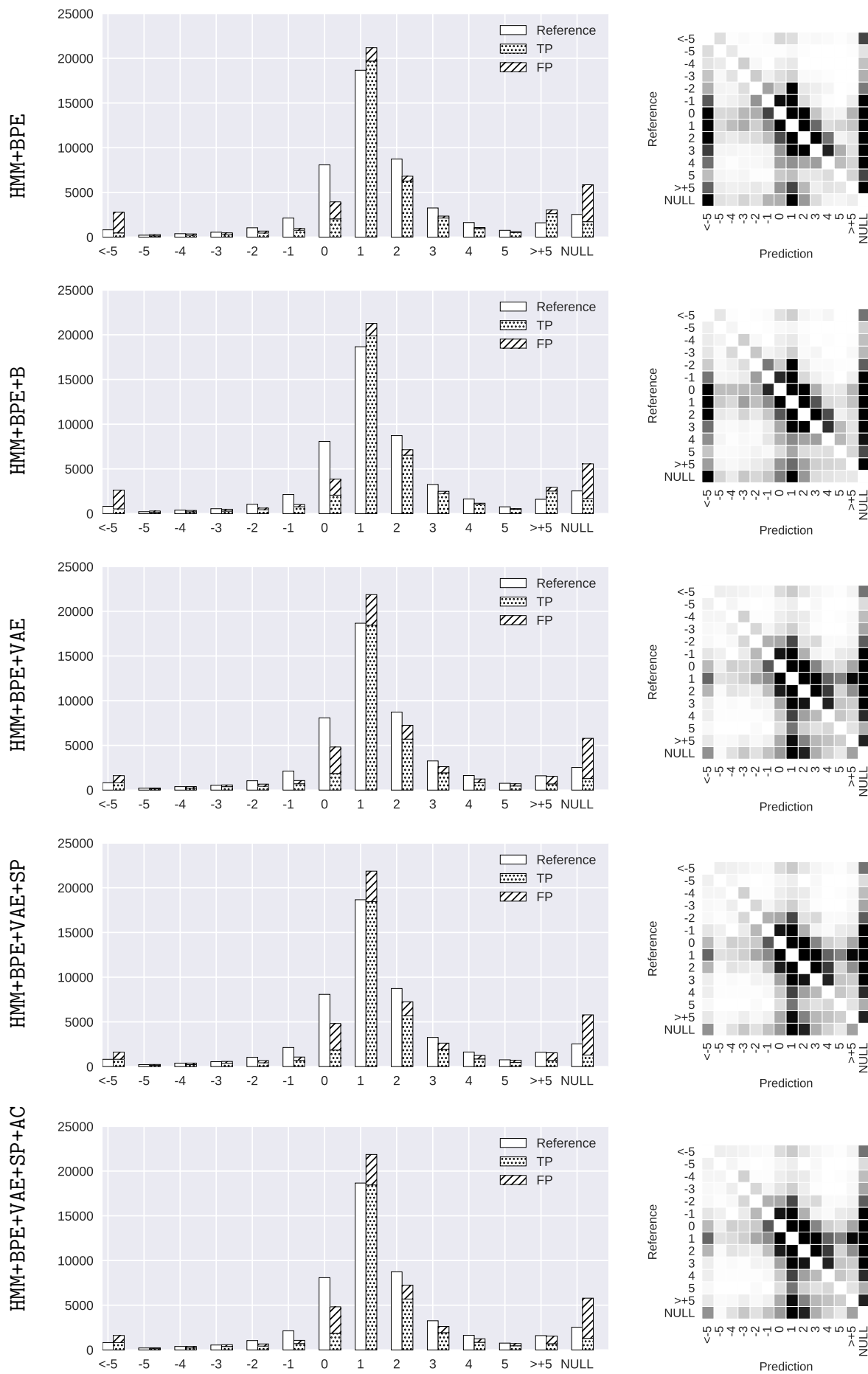
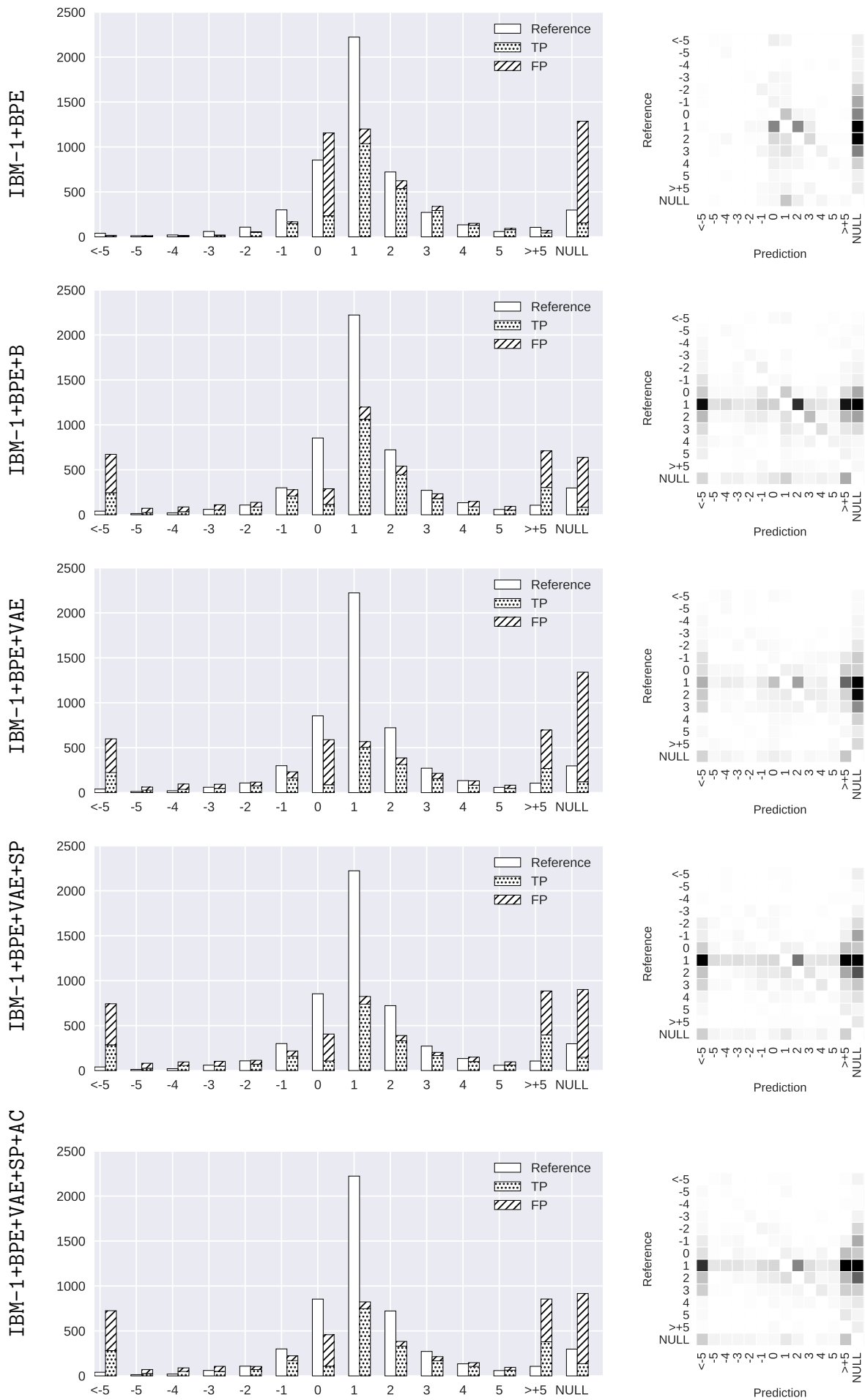


Figure D.36: Variational models for the direction Czech-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the right

graph: The darker the cell, the greater the number of confusions.



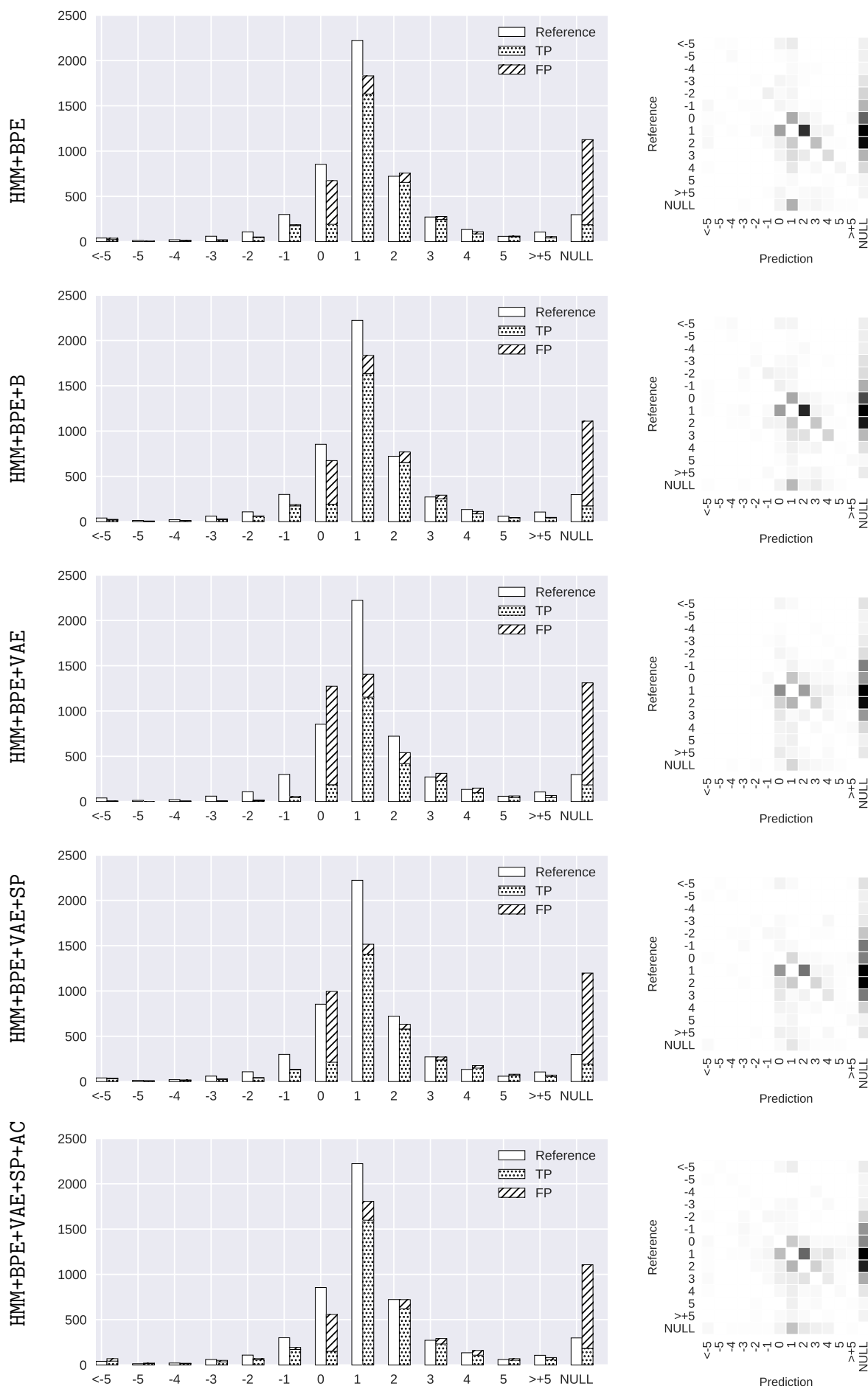
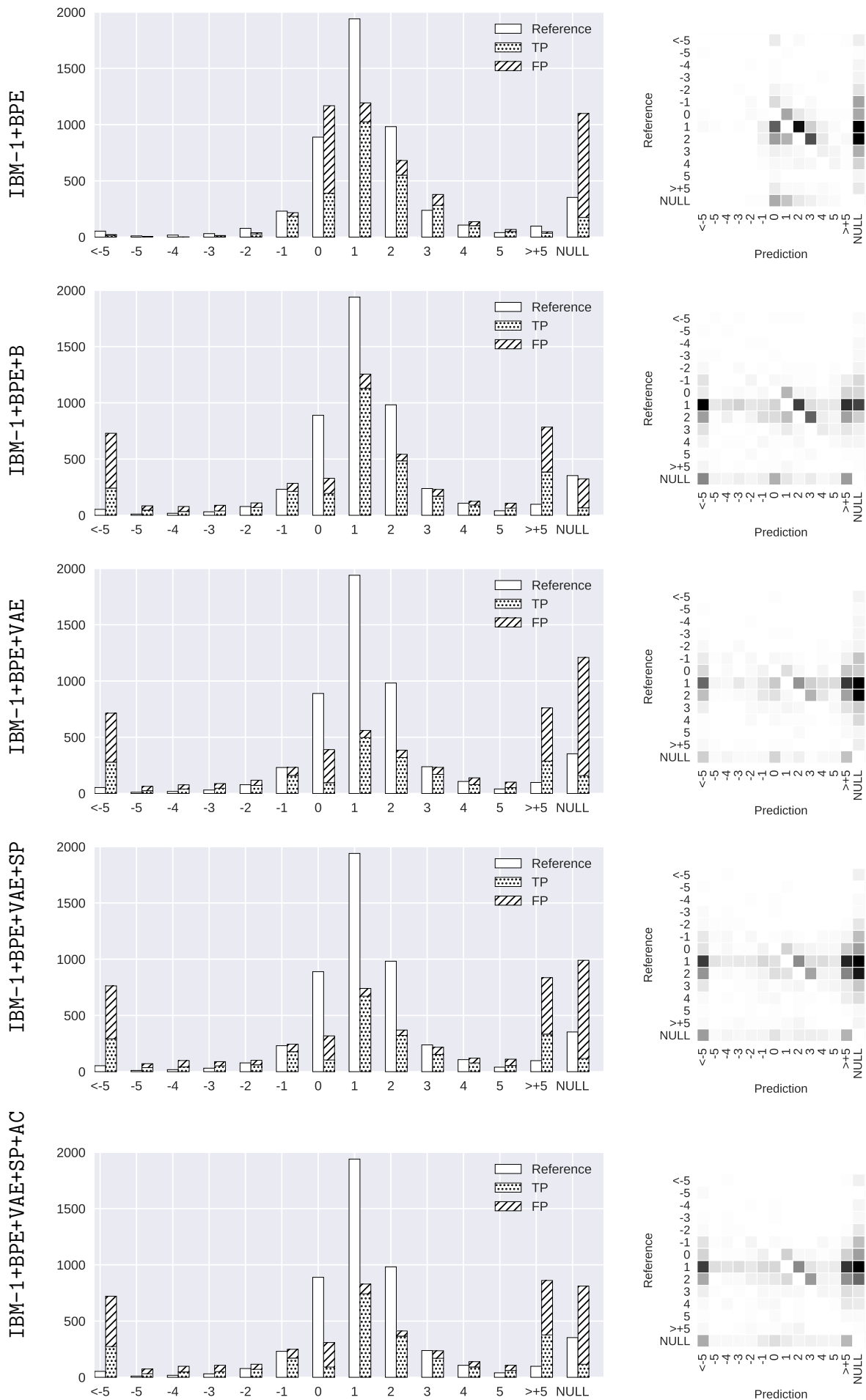


Figure D.37: Variational models for the direction English-Romanian: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the right

graph: The darker the cell, the greater the number of confusions.



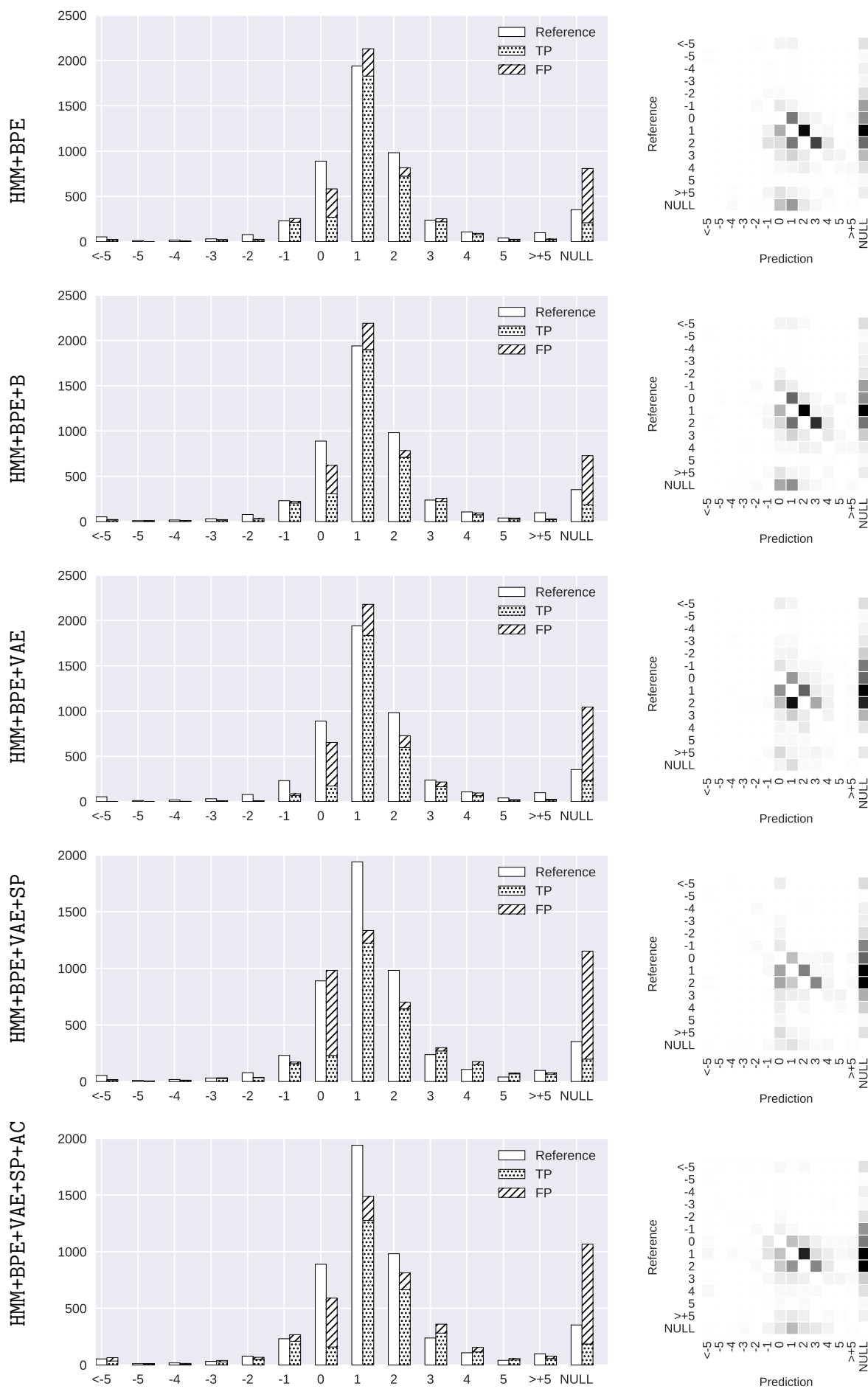
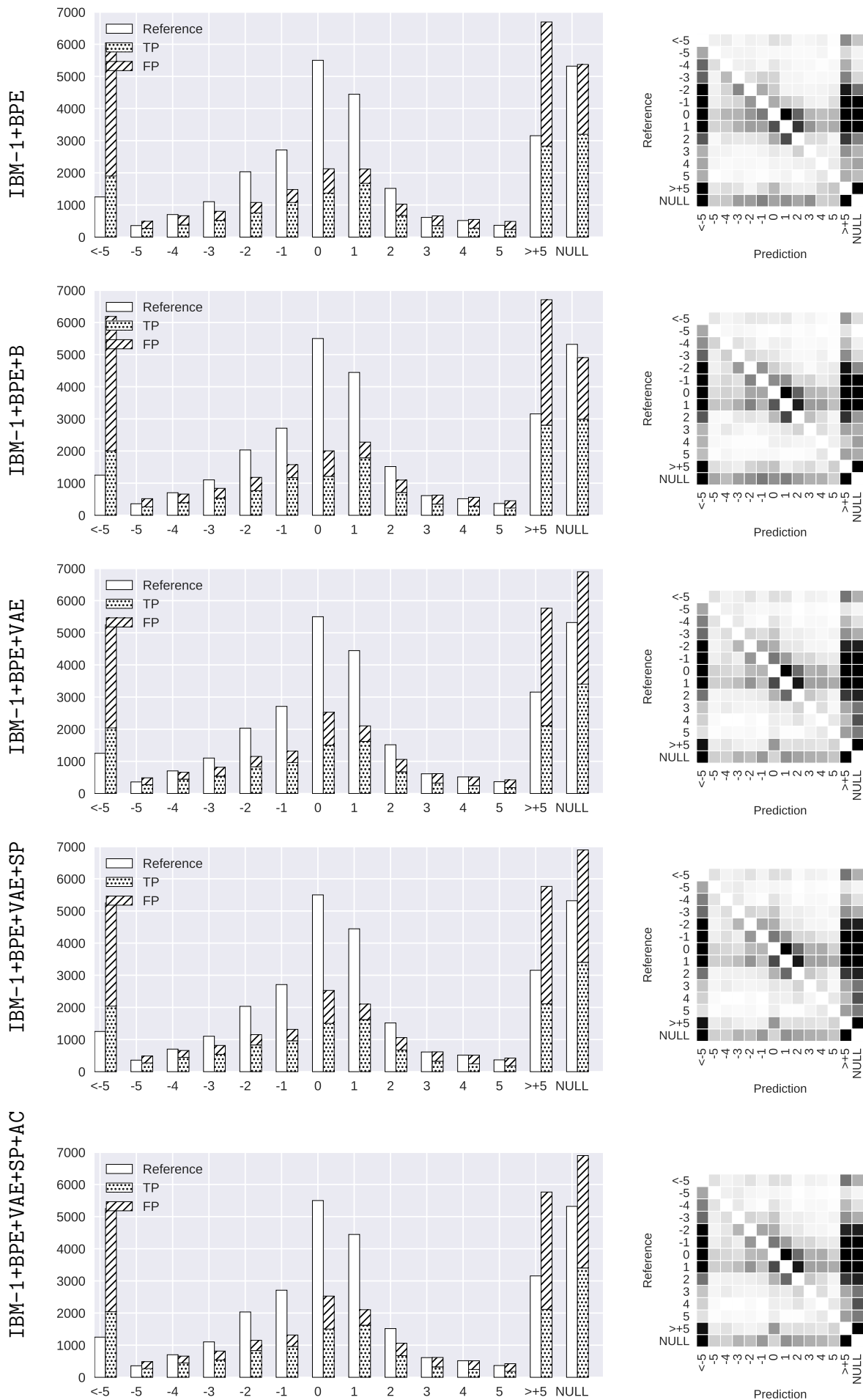


Figure D.38: Variational models for the direction Romanian-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the right

graph: The darker the cell, the greater the number of confusions.



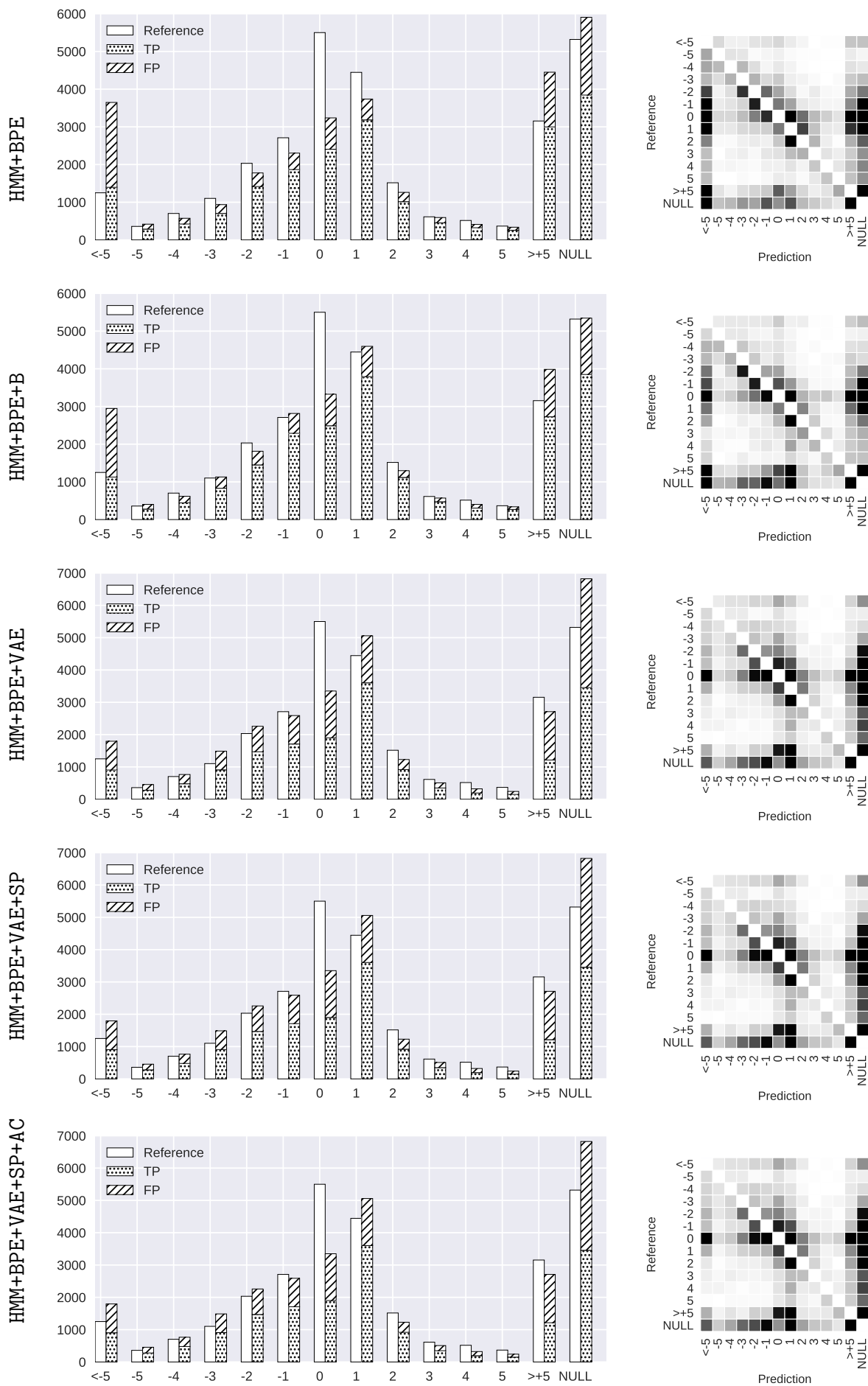
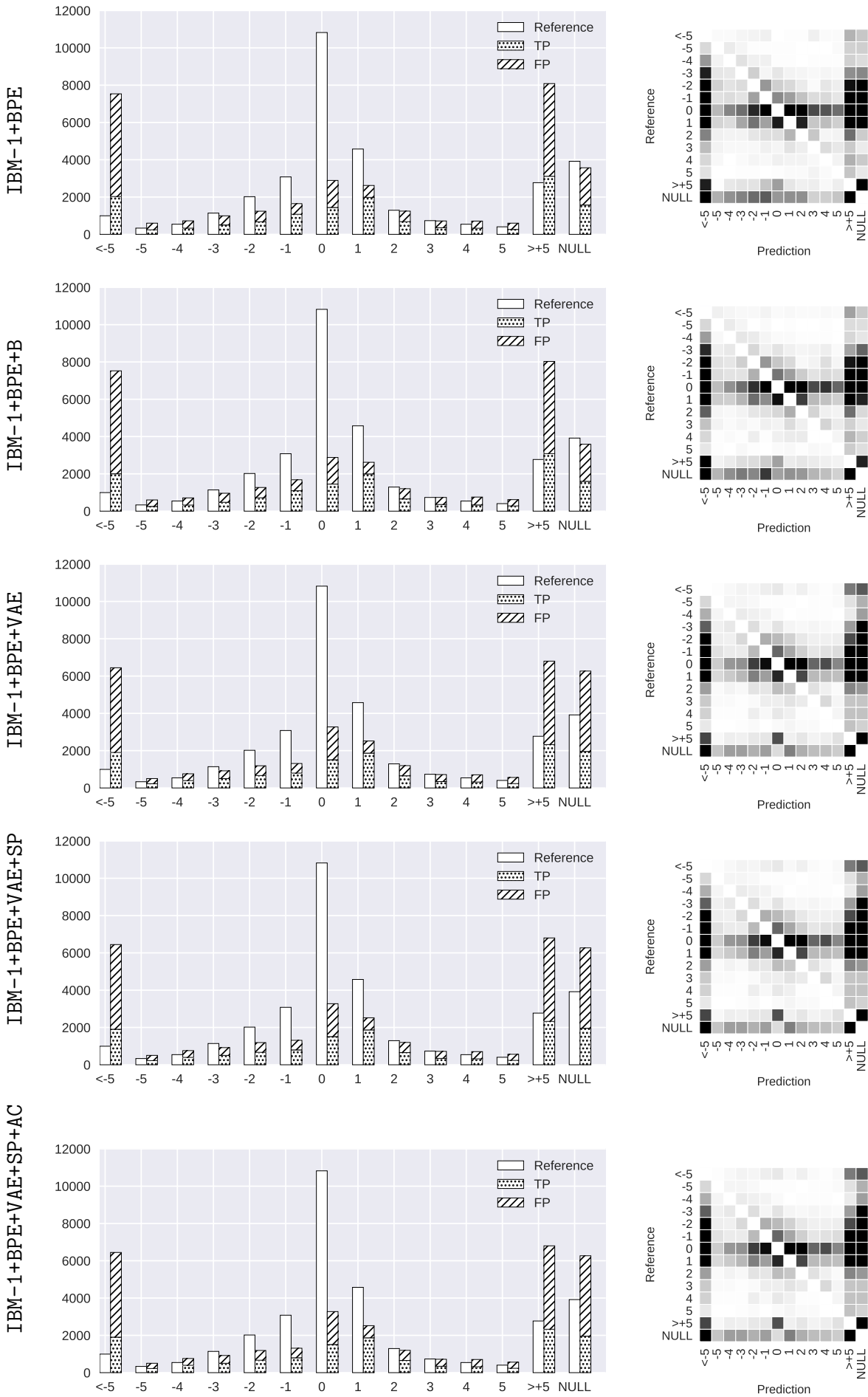


Figure D.39: Variational models for the direction English-Japanese: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the right

graph: The darker the cell, the greater the number of confusions.



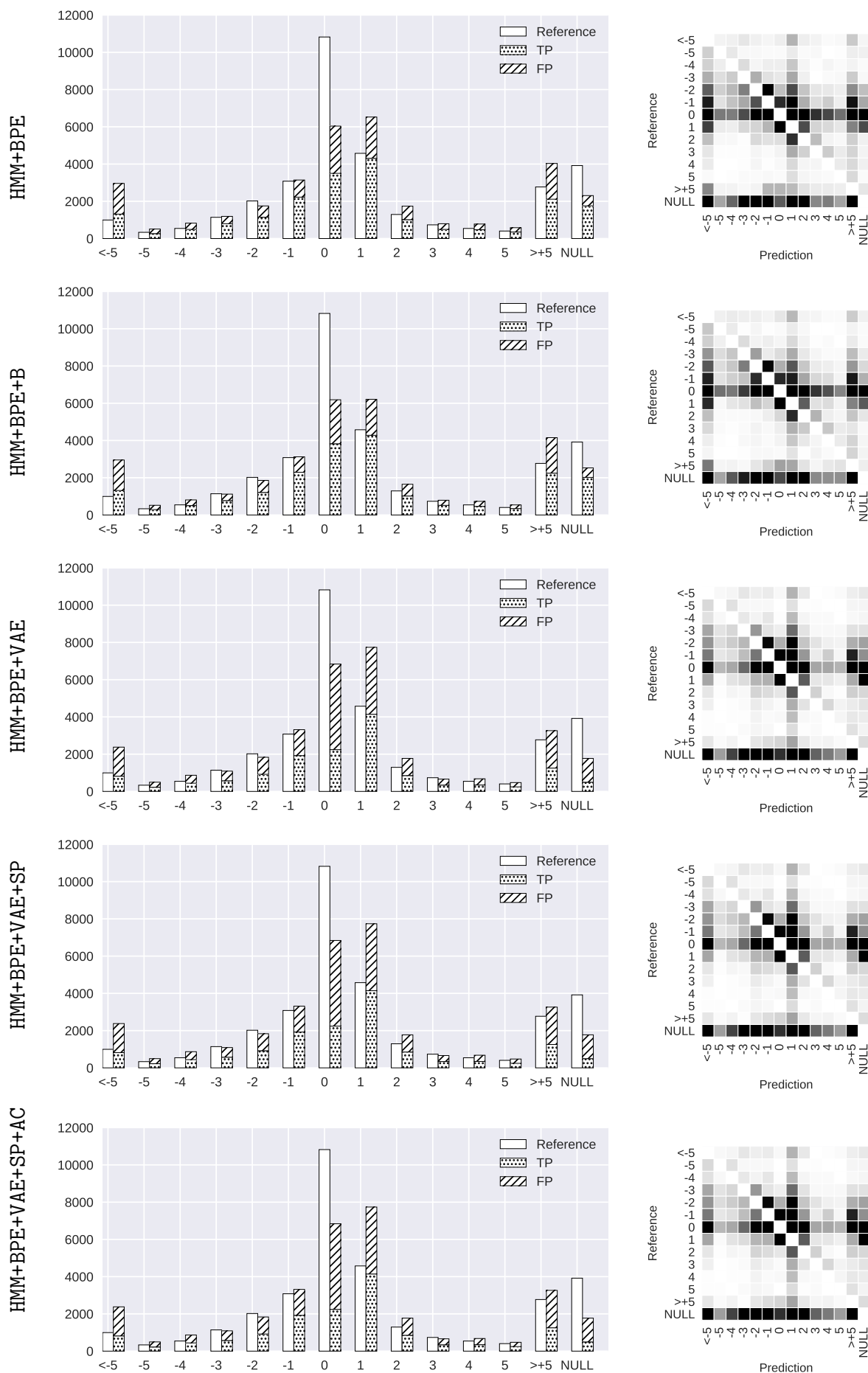
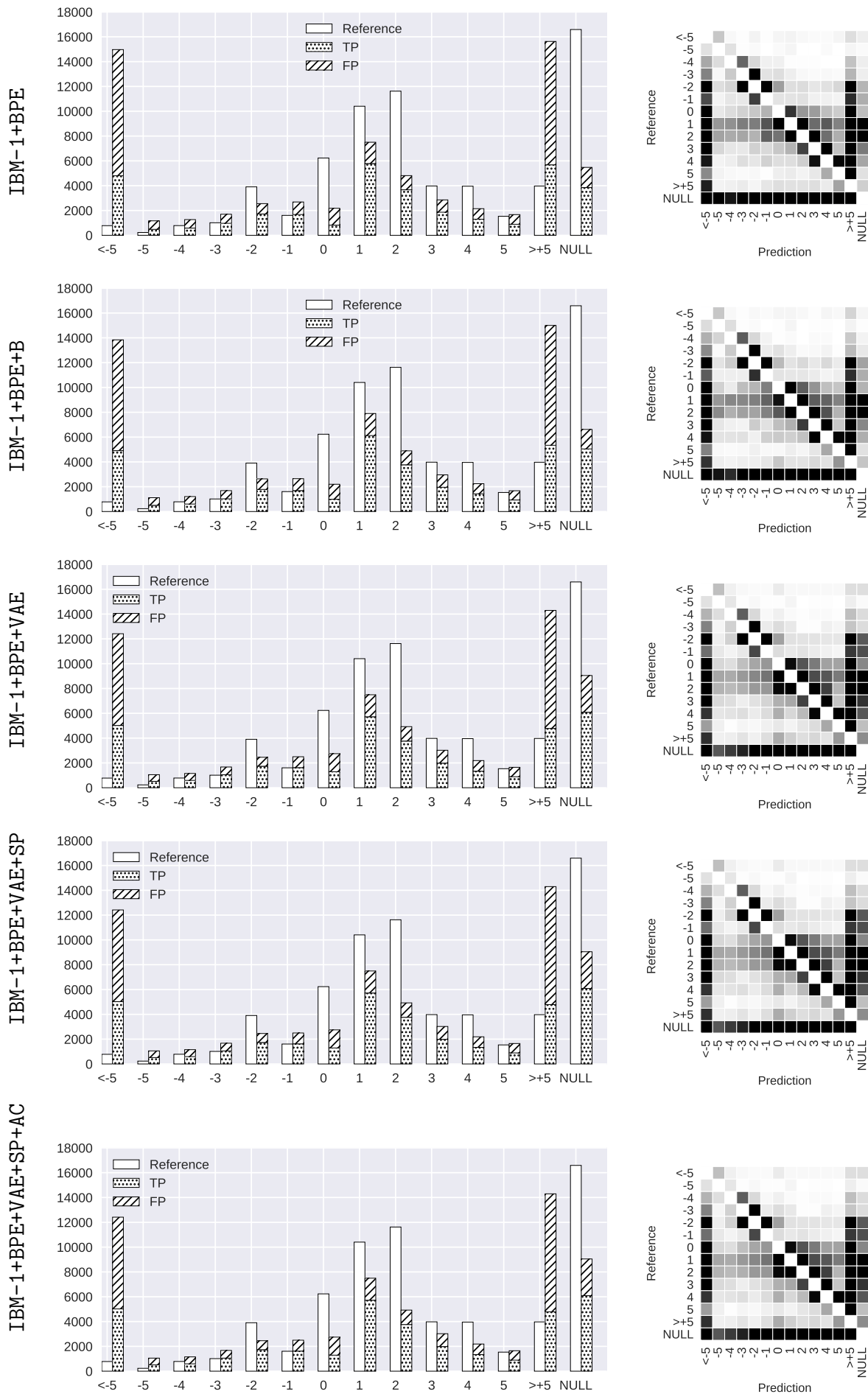


Figure D.40: Variational models for the direction Japanese-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the right

graph: The darker the cell, the greater the number of confusions.



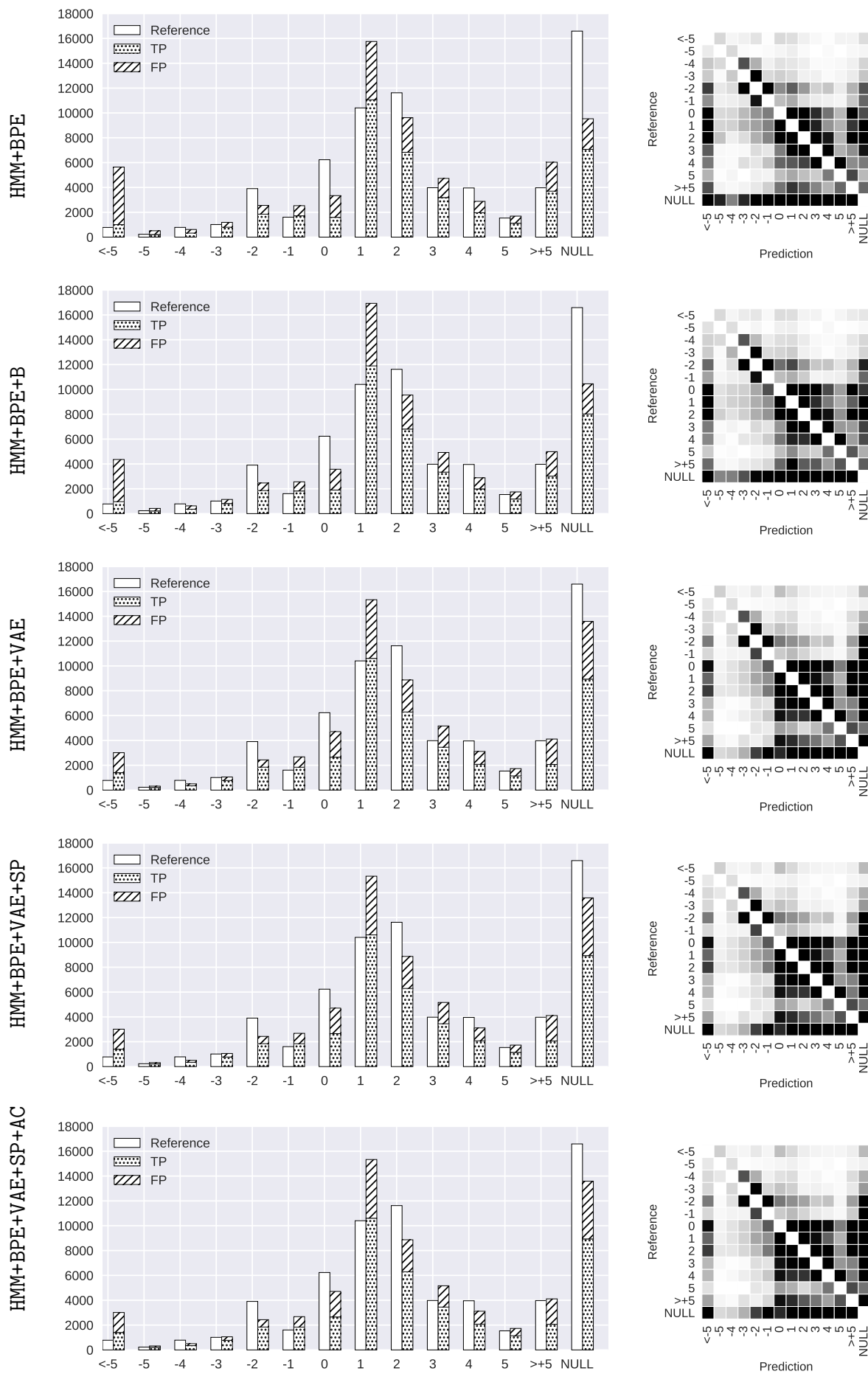
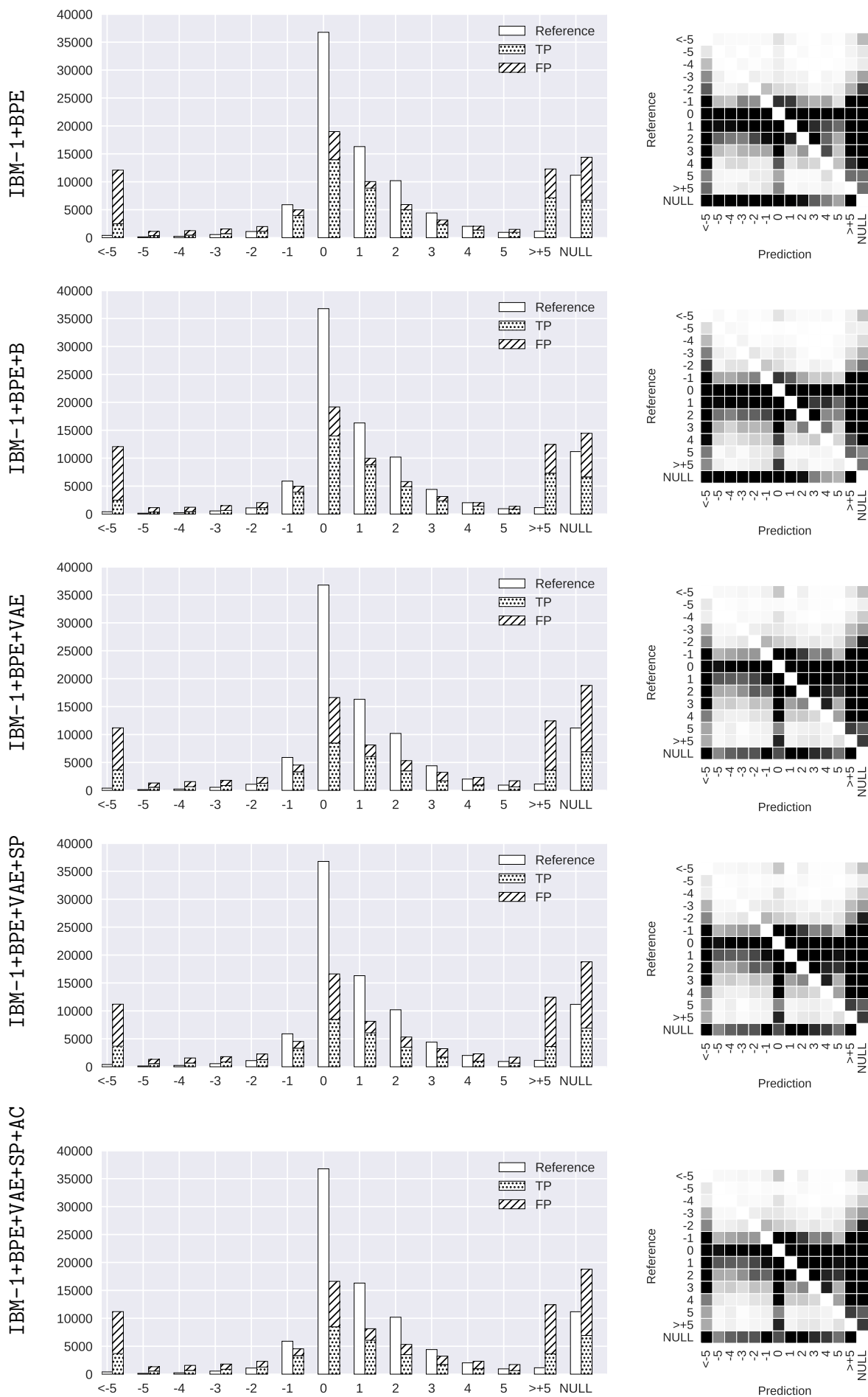


Figure D.41: Variational models for the direction English-Vietnamese: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the

right graph: The darker the cell, the greater the number of confusions.



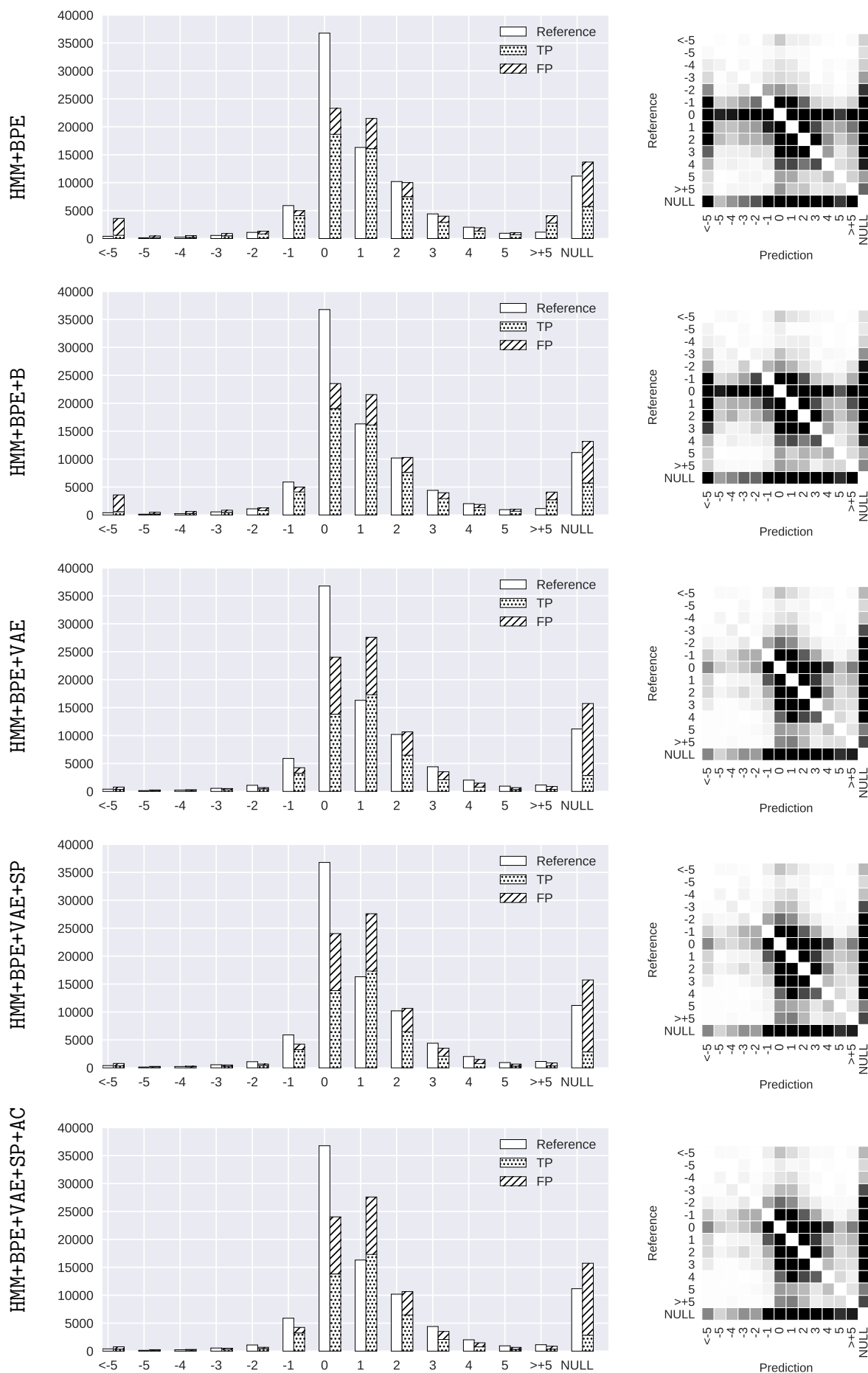


Figure D.42: Variational models for the direction Vietnamese-English: Correct (TP) and incorrect (FP) jump widths for source words on the left graph. The confusion matrix is on the

right graph: The darker the cell, the greater the number of confusions.

D.6 Rare words

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	45	4.09	86.61	31.11	93.33	46.67	64	2.78	86.21	23.44	40.54	29.7
IBM-1+NN	7	0.64	93.31	42.86	20	27.27	38	1.65	89.71	28.95	29.73	29.33
IBM-1+NNChar	7	0.64	94.98	71.43	33.33	45.45	45	1.96	87.57	20	24.32	21.95
IBM-1+BPE	11	1	95.82	72.73	53.33	61.54	17	0.74	95.34	88.24	40.54	55.56
IBM-1+BPE+B	20	1.82	91.21	35	46.67	40	36	1.57	92.04	44.44	43.24	43.84
IBM-1+BPE+VAE	13	1.18	94.98	61.54	53.33	57.14	27	1.17	92.62	48.15	35.14	40.62
IBM-1+BPE+VAE+SP	12	1.09	94.56	58.33	46.67	51.85	25	1.09	93.01	52	35.14	41.94
IBM-1+BPE+VAE+SP+AC	14	1.27	93.72	50	46.67	48.28	20	0.87	93.59	60	32.43	42.11
Fastalign	11	1	95.82	72.73	53.33	61.54	20	0.87	95.15	80	43.24	56.14
HMM Giza++	33	3	91.63	42.42	93.33	58.33	72	3.13	87.38	30.56	59.46	40.37
HMM+NN	11	1	96.65	81.82	60	69.23	34	1.48	92.82	50	45.95	47.89
HMM+NNCharTgt	9	0.82	96.65	88.89	53.33	66.67	26	1.13	95.53	76.92	54.05	63.49
HMM+NNCharJB	8	0.73	97.07	100	53.33	69.57	18	0.78	94.76	77.78	37.84	50.91
HMM+BPE	9	0.82	96.65	88.89	53.33	66.67	18	0.78	95.15	83.33	40.54	54.55
HMM+BPE+B	8	0.73	96.23	87.5	46.67	60.87	16	0.7	95.53	93.75	40.54	56.6
HMM+BPE+VAE	13	1.18	94.98	61.54	53.33	57.14	19	0.83	93.4	57.89	29.73	39.29
HMM+BPE+VAE+SP	6	0.55	96.23	100	40	57.14	18	0.78	95.15	83.33	40.54	54.55
HMM+BPE+VAE+SP+AC	7	0.64	95.82	85.71	40	54.55	18	0.78	95.53	88.89	43.24	58.18
IBM-4 Giza++	21	1.91	94.14	52.38	73.33	61.11	55	2.39	89.9	36.36	54.05	43.48

Table D.19: Models for English-French: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction French-English and in the direction English-French

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	26	4.33	78.02	23.08	100	37.5	88	4.63	85.37	35.23	77.5	48.44
IBM-1+NN	4	0.67	95.6	75	50	60	35	1.84	92.24	57.14	50	53.33
IBM-1+NNChar	9	1.5	92.31	44.44	66.67	53.33	38	2	93.79	65.79	62.5	64.1
IBM-1+BPE	4	0.67	95.6	75	50	60	31	1.63	96.23	87.1	67.5	76.06
IBM-1+BPE+B	7	1.17	96.7	71.43	83.33	76.92	31	1.63	96.67	90.32	70	78.87
IBM-1+BPE+VAE	7	1.17	96.7	71.43	83.33	76.92	45	2.37	94.9	68.89	77.5	72.94
IBM-1+BPE+VAE+SP	4	0.67	97.8	100	66.67	80	41	2.16	97.12	82.93	85	83.95
IBM-1+BPE+VAE+SP+AC	4	0.67	97.8	100	66.67	80	47	2.47	96.23	74.47	87.5	80.46
Fastalign	8	1.33	95.6	62.5	83.33	71.43	27	1.42	95.34	85.19	57.5	68.66
HMM Giza++	22	3.67	80.22	22.73	83.33	35.71	79	4.16	88.69	43.04	85	57.14
HMM+NN	4	0.67	97.8	100	66.67	80	36	1.89	94.68	72.22	65	68.42
HMM+NNCharTgt	5	0.83	96.7	80	66.67	72.73	31	1.63	96.67	90.32	70	78.87
HMM+NNCharJB	5	0.83	96.7	80	66.67	72.73	32	1.68	96.45	87.5	70	77.78
HMM+BPE	5	0.83	98.9	100	83.33	90.91	43	2.26	97.12	81.4	87.5	84.34
HMM+BPE+B	5	0.83	98.9	100	83.33	90.91	39	2.05	98.45	92.31	90	91.14
HMM+BPE+VAE	5	0.83	98.9	100	83.33	90.91	44	2.32	95.12	70.45	77.5	73.81
HMM+BPE+VAE+SP	4	0.67	97.8	100	66.67	80	28	1.47	96.9	96.43	67.5	79.41
HMM+BPE+VAE+SP+AC	4	0.67	97.8	100	66.67	80	34	1.79	97.78	94.12	80	86.49
IBM-4 Giza++	15	2.5	87.91	33.33	83.33	47.62	72	3.79	90.24	47.22	85	60.71

Table D.20: Models for English-German: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction German-English and in the direction English-German

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	1961	4.25	85.54	15.96	56.09	24.85	3365	2.86	90.68	23.6	46.06	31.2
IBM-1+NN	582	1.26	93.19	21.31	22.22	21.75	1131	0.96	93.55	19.01	12.47	15.06
IBM-1+NNChar	637	1.38	92.98	21.66	24.73	23.1	1817	1.55	93.54	30.6	32.25	31.4
IBM-1+BPE	450	0.98	96.43	60	48.39	53.57	1273	1.08	96.23	62.14	45.88	52.79
IBM-1+BPE+B	446	0.97	96.56	62.11	49.64	55.18	1290	1.1	96.45	65.12	48.72	55.74
IBM-1+BPE+VAE	716	1.55	94.32	37.01	47.49	41.6	1695	1.44	94.54	40.29	39.62	39.95
IBM-1+BPE+VAE+SP	845	1.83	93.5	32.66	49.46	39.34	1756	1.49	94.59	41.23	42	41.61
IBM-1+BPE+VAE+SP+AC	759	1.65	95.24	45.72	62.19	52.7	2101	1.79	94.78	44.36	54.06	48.73
Fastalign	700	1.52	95.94	51.86	65.05	57.71	1489	1.27	95.84	55.41	47.85	51.35
HMM Giza++	1623	3.52	89.42	24.52	71.33	36.5	2878	2.45	93.61	38.26	63.86	47.85
HMM+NN	521	1.13	96.63	61.23	57.17	59.13	1409	1.2	96.32	62.17	50.81	55.92
HMM+NNCharTgt	461	1	97.14	69.85	57.71	63.2	1205	1.02	97.23	78.42	54.81	64.53
HMM+NNCharJB	520	1.13	96.2	55.77	51.97	53.8	1272	1.08	97.17	75.94	56.03	64.49
HMM+BPE	485	1.05	97.56	74.64	64.87	69.42	1282	1.09	97.62	82.37	61.25	70.26
HMM+BPE+B	510	1.11	97.68	74.9	68.46	71.54	1288	1.1	97.61	82.07	61.31	70.19
HMM+BPE+VAE	837	1.82	94.81	42.77	64.16	51.33	1764	1.5	95.39	49.77	50.93	50.34
HMM+BPE+VAE+SP	482	1.05	97.78	77.8	67.2	72.12	1419	1.21	97.34	75.55	62.18	68.22
HMM+BPE+VAE+SP+AC	483	1.05	97.79	77.85	67.38	72.24	1473	1.25	97.7	79.16	67.63	72.94
IBM-4 Giza++	1468	3.18	90.83	28.13	74.01	40.77	2430	2.07	95	46.79	65.95	54.74

Table D.21: Models for English-Czech: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Czech-English and in the direction English-Czech

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	82	6.31	78.74	19.51	76.19	31.07	137	2.49	89.77	31.39	48.86	38.22
IBM-1+NN	17	1.31	92.81	41.18	33.33	36.84	42	0.76	93.38	47.62	22.73	30.77
IBM-1+NNChar	12	0.92	92.51	33.33	19.05	24.24	77	1.4	90.8	25.97	22.73	24.24
IBM-1+BPE	10	0.77	96.71	100	47.62	64.52	40	0.73	94.7	70	31.82	43.75
IBM-1+BPE+B	18	1.38	94.31	55.56	47.62	51.28	71	1.29	93.6	50.7	40.91	45.28
IBM-1+BPE+VAE	25	1.92	93.41	48	57.14	52.17	70	1.27	93.38	48.57	38.64	43.04
IBM-1+BPE+VAE+SP	22	1.69	91.92	36.36	38.1	37.21	77	1.4	92.27	38.96	34.09	36.36
IBM-1+BPE+VAE+SP+AC	26	2	91.32	34.62	42.86	38.3	71	1.29	92.42	39.44	31.82	35.22
Fastalign	17	1.31	97.01	82.35	66.67	73.68	57	1.04	95.36	71.93	46.59	56.55
HMM Giza++	53	4.08	87.43	30.19	76.19	43.24	182	3.31	88.96	32.97	68.18	44.44
HMM+NN	15	1.15	94.01	53.33	38.1	44.44	71	1.29	95.22	66.2	53.41	59.12
HMM+NNCharTgt	14	1.08	95.51	71.43	47.62	57.14	70	1.27	96.03	74.29	59.09	65.82
HMM+NNCharJB	12	0.92	96.11	83.33	47.62	60.61	71	1.29	95.51	69.01	55.68	61.64
HMM+BPE	16	1.23	95.51	68.75	52.38	59.46	49	0.89	94.78	67.35	37.5	48.18
HMM+BPE+B	15	1.15	96.41	80	57.14	66.67	47	0.85	95.36	76.6	40.91	53.33
HMM+BPE+VAE	16	1.23	94.31	56.25	42.86	48.65	44	0.8	93.97	56.82	28.41	37.88
HMM+BPE+VAE+SP	11	0.85	94.01	54.55	28.57	37.5	37	0.67	94.48	67.57	28.41	40
HMM+BPE+VAE+SP+AC	15	1.15	94.61	60	42.86	50	52	0.95	94.41	61.54	36.36	45.71
IBM-4 Giza++	50	3.85	89.52	36	85.71	50.7	167	3.04	89.92	35.33	67.05	46.27

Table D.22: Models for English-Romanian: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Romanian-English and in the direction English-Romanian

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	642	3.75	92.43	31.62	65.48	42.65	412	4.12	89.35	16.99	51.47	25.55
IBM-1+NN	653	3.82	91.77	28.33	59.68	38.42	581	5.81	84.1	9.29	39.71	15.06
IBM-1+NNChar	664	3.88	91.51	27.26	58.39	37.17	107	1.07	96.74	55.14	43.38	48.56
IBM-1+BPE	323	1.89	95.88	52.01	54.19	53.08	158	1.58	96.08	45.57	52.94	48.98
IBM-1+BPE+B	331	1.94	95.77	50.76	54.19	52.42	139	1.39	96.58	51.8	52.94	52.36
IBM-1+BPE+VAE	318	1.86	95.81	51.26	52.58	51.91	132	1.32	96.66	53.03	51.47	52.24
IBM-1+BPE+VAE+SP	292	1.71	95.98	53.42	50.32	51.83	138	1.38	96.92	56.52	57.35	56.93
IBM-1+BPE+VAE+SP+AC	327	1.91	95.94	52.6	55.48	54	139	1.39	96.79	54.68	55.88	55.27
Fastalign	231	1.35	96.8	67.1	50	57.3	100	1	96.45	50	36.76	42.37
HMM Giza++	639	3.74	92.58	32.39	66.77	43.62	442	4.42	89.03	17.87	58.09	27.34
HMM+NN	291	1.7	96.82	63.92	60	61.9	88	0.88	97.34	69.32	44.85	54.46
HMM+NNCharTgt	280	1.64	96.34	58.21	52.58	55.25	160	1.6	96.87	55	64.71	59.46
HMM+NNCharJB	553	3.23	94.88	44.67	79.68	57.24	154	1.54	96.14	46.1	52.21	48.97
HMM+BPE	293	1.71	97.6	73.38	69.35	71.31	132	1.32	98.07	73.48	71.32	72.39
HMM+BPE+B	295	1.73	97.54	72.54	69.03	70.74	135	1.35	97.78	68.89	68.38	68.63
HMM+BPE+VAE	292	1.71	96.75	63.01	59.35	61.13	132	1.32	96.92	56.82	55.15	55.97
HMM+BPE+VAE+SP	296	1.73	96.84	63.85	60.97	62.38	120	1.2	97.55	67.5	59.56	63.28
HMM+BPE+VAE+SP+AC	297	1.74	96.93	64.98	62.26	63.59	129	1.29	97.78	69.77	66.18	67.92
IBM-4 Giza++	667	3.9	92.74	34.03	73.23	46.47	425	4.25	89.48	18.59	58.09	28.16

Table D.23: Models for English-Japanese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Japanese-English and in the direction English-Japanese

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	3054	3.39	91.52	32.65	56.94	41.5	1147	2.76	90.04	23.63	64.68	34.61
IBM-1+NN	1512	1.68	96.23	66.53	57.45	61.66	390	0.94	96.97	63.85	59.43	61.56
IBM-1+NNChar	2474	2.74	93.47	41.63	58.82	48.76	415	1	97.2	65.78	65.16	65.47
IBM-1+BPE	1287	1.43	96.23	69.39	51	58.79	430	1.04	95.81	48.6	49.88	49.23
IBM-1+BPE+B	1310	1.45	96.16	68.24	51.06	58.41	422	1.02	95.83	48.82	49.16	48.99
IBM-1+BPE+VAE	1077	1.19	94.47	46.15	28.38	35.15	428	1.03	95.75	47.9	48.93	48.41
IBM-1+BPE+VAE+SP	1371	1.52	95.47	59.01	46.2	51.83	423	1.02	96.5	56.97	57.52	57.24
IBM-1+BPE+VAE+SP+AC	1383	1.53	95.63	60.88	48.09	53.73	407	0.98	96.54	57.74	56.09	56.9
Fastalign	1280	1.42	96.32	70.78	51.74	59.78	488	1.18	97.11	62.5	72.79	67.25
HMM Giza++	2853	3.16	93.37	42.13	68.65	52.22	1090	2.63	91.1	27.25	70.88	39.36
HMM+NN	2715	3.01	94.22	46.96	72.82	57.1	361	0.87	97.04	65.93	56.8	61.03
HMM+NNCharTgt	2007	2.23	95.95	60.14	68.93	64.24	427	1.03	97.57	69.79	71.12	70.45
HMM+NNCharJB	2438	2.7	95.2	53.28	74.19	62.02	324	0.78	98.26	87.04	67.3	75.91
HMM+BPE	1307	1.45	96.68	74.9	55.91	64.03	424	1.02	96.7	59.43	60.14	59.79
HMM+BPE+B	1339	1.48	96.7	74.53	57	64.6	426	1.03	96.74	59.86	60.86	60.36
HMM+BPE+VAE	1149	1.27	95.33	58.83	38.61	46.62	396	0.95	96.74	60.61	57.28	58.9
HMM+BPE+VAE+SP	1179	1.31	96.32	72.52	48.83	58.36	402	0.97	97.4	68.91	66.11	67.48
HMM+BPE+VAE+SP+AC	1192	1.32	96.59	76.01	51.74	61.57	394	0.95	97.6	71.83	67.54	69.62
IBM-4 Giza++	2586	2.87	94.71	49.92	73.73	59.53	956	2.3	92.75	32.95	75.18	45.82

Table D.24: Models for English-Vietnamese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for the rare target words in the direction Vietnamese-English and in the direction English-Vietnamese

D.7 Unknown words

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	680	4.33	79.06	22.65	52.38	31.62	298	4.66	78.04	21.14	62.38	31.58
IBM-1+NN	128	0.82	89.37	32.81	14.29	19.91	93	1.45	90.02	37.63	34.65	36.08
IBM-1+NNChar	188	1.2	89.31	37.77	24.15	29.46	109	1.7	88.42	30.28	32.67	31.43
IBM-1+BPE	166	1.06	91.95	61.45	34.69	44.35	69	1.08	94.53	73.91	50.5	60
IBM-1+BPE+B	189	1.2	91.48	56.08	36.05	43.89	76	1.19	93.32	61.84	46.53	53.11
IBM-1+BPE+VAE	184	1.17	91.64	57.61	36.05	44.35	100	1.56	91.87	50	49.5	49.75
IBM-1+BPE+VAE+SP	183	1.17	92.49	65.03	40.48	49.9	88	1.38	93	57.95	50.5	53.97
IBM-1+BPE+VAE+SP+AC	190	1.21	92.58	65.26	42.18	51.24	91	1.42	92.76	56.04	50.5	53.12
Fastalign	269	1.71	91.86	56.51	51.7	54	82	1.28	93.81	64.63	52.48	57.92
HMM Giza++	432	2.75	88.05	40.05	58.84	47.66	226	3.53	83.67	27.43	61.39	37.92
HMM+NN	176	1.12	92.46	65.34	39.12	48.94	94	1.47	93	57.45	53.47	55.38
HMM+NNCharTgt	153	0.97	93.43	77.78	40.48	53.24	70	1.09	94.45	72.86	50.5	59.65
HMM+NNCharJB	128	0.82	93.59	85.16	37.07	51.66	71	1.11	93.56	64.79	45.54	53.49
HMM+BPE	152	0.97	92.64	69.74	36.05	47.53	66	1.03	94.77	77.27	50.5	61.08
HMM+BPE+B	144	0.92	93.27	77.78	38.1	51.14	62	0.97	94.93	80.65	49.5	61.35
HMM+BPE+VAE	188	1.2	91.83	59.04	37.76	46.06	64	1	93.81	68.75	43.56	53.33
HMM+BPE+VAE+SP	138	0.88	93.65	83.33	39.12	53.24	63	0.98	95.33	84.13	52.48	64.63
HMM+BPE+VAE+SP+AC	178	1.13	93.02	70.22	42.52	52.97	70	1.09	96.06	87.14	60.4	71.35
IBM-4 Giza++	388	2.47	88.81	42.01	55.44	47.8	194	3.03	86.4	32.47	62.38	42.71

Table D.25: Models for English-French: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction French-English and in the direction English-French.

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	112	7.47	73.74	13.39	68.18	22.39	393	6.78	79.13	25.45	77.52	38.31
IBM-1+NN	15	1	92.17	20	13.64	16.22	94	1.62	91.51	48.94	35.66	41.26
IBM-1+NNChar	20	1.33	93.94	45	40.91	42.86	120	2.07	93.84	64.17	59.69	61.85
IBM-1+BPE	27	1.8	93.18	40.74	50	44.9	105	1.81	97.41	92.38	75.19	82.91
IBM-1+BPE+B	33	2.2	93.18	42.42	63.64	50.91	104	1.79	97.73	95.19	76.74	84.98
IBM-1+BPE+VAE	24	1.6	95.45	58.33	63.64	60.87	111	1.91	94.56	70.27	60.47	65
IBM-1+BPE+VAE+SP	23	1.53	93.18	39.13	40.91	40	128	2.21	97.21	83.59	82.95	83.27
IBM-1+BPE+VAE+SP+AC	22	1.47	96.46	68.18	68.18	68.18	122	2.1	97.6	87.7	82.95	85.26
Fastalign	15	1	94.19	46.67	31.82	37.84	48	0.83	92.81	68.75	25.58	37.29
HMM Giza++	88	5.87	81.82	21.59	86.36	34.55	323	5.57	84.83	33.75	84.5	48.23
HMM+NN	10	0.67	95.45	70	31.82	43.75	102	1.76	96.18	84.31	66.67	74.46
HMM+NNCharTgt	14	0.93	96.97	85.71	54.55	66.67	89	1.53	96.63	93.26	64.34	76.15
HMM+NNCharJB	16	1.07	96.46	75	54.55	63.16	98	1.69	97.21	93.88	71.32	81.06
HMM+BPE	28	1.87	96.46	64.29	81.82	72	114	1.97	97.21	87.72	77.52	82.3
HMM+BPE+B	28	1.87	96.46	64.29	81.82	72	107	1.84	97.67	93.46	77.52	84.75
HMM+BPE+VAE	29	1.93	94.19	48.28	63.64	54.9	110	1.9	95.79	79.09	67.44	72.8
HMM+BPE+VAE+SP	15	1	97.22	86.67	59.09	70.27	100	1.72	97.34	94	72.87	82.1
HMM+BPE+VAE+SP+AC	19	1.27	98.23	89.47	77.27	82.93	98	1.69	97.47	95.92	72.87	82.82
IBM-4 Giza++	76	5.07	84.85	25	86.36	38.78	270	4.66	88.66	41.48	86.82	56.14

Table D.26: Models for English-German: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction German-English and in the direction English-German.

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	6931	4.33	85.4	16.87	55.53	25.87	8487	3.33	89.86	20.9	48.91	29.29
IBM-1+NN	2073	1.3	92.8	21.13	20.81	20.97	2779	1.09	93.7	19.5	14.94	16.92
IBM-1+NNChar	2210	1.38	92.9	23.94	25.13	24.52	4393	1.73	92.9	23.06	27.93	25.26
IBM-1+BPE	1707	1.07	95.63	52.9	42.9	47.38	2766	1.09	96.27	58.6	44.69	50.71
IBM-1+BPE+B	1695	1.06	95.9	56.64	45.61	50.53	2814	1.11	96.23	57.92	44.94	50.61
IBM-1+BPE+VAE	2470	1.54	93.64	33.56	39.38	36.24	3964	1.56	94.12	33.12	36.2	34.59
IBM-1+BPE+VAE+SP	2930	1.83	92.83	29.76	41.43	34.64	3997	1.57	94.33	35.45	39.07	37.17
IBM-1+BPE+VAE+SP+AC	2682	1.68	94.68	43.77	55.77	49.05	4935	1.94	94.34	38.28	52.08	44.13
Fastalign	2118	1.32	96.29	59.54	59.9	59.72	3056	1.2	96.22	57.04	48.06	52.16
HMM Giza++	5702	3.57	89.75	27.24	73.78	39.78	7488	2.94	92.59	32.41	66.91	43.67
HMM+NN	1925	1.2	96.24	59.84	54.73	57.17	3115	1.22	96.34	58.62	50.34	54.17
HMM+NNCharTgt	1665	1.04	96.49	64.86	51.31	57.29	2489	0.98	97.18	74.97	51.45	61.02
HMM+NNCharJB	1931	1.21	95.91	55.88	51.26	53.47	3016	1.18	96.82	65.58	54.54	59.55
HMM+BPE	1711	1.07	97.05	72	58.53	64.57	2775	1.09	97.59	78.7	60.22	68.23
HMM+BPE+B	1739	1.09	97.1	72.23	59.67	65.35	2672	1.05	97.69	81.32	59.91	69
HMM+BPE+VAE	3101	1.94	93.93	39.08	57.58	46.56	4005	1.57	95.37	46.42	51.25	48.72
HMM+BPE+VAE+SP	1720	1.08	97.27	74.77	61.09	67.24	3235	1.27	97.23	69.86	62.31	65.87
HMM+BPE+VAE+SP+AC	1743	1.09	97.27	74.41	61.62	67.41	3311	1.3	97.6	74.15	67.69	70.77
IBM-4 Giza++	5132	3.21	91.11	30.79	75.06	43.66	6058	2.38	94.39	40.82	68.18	51.07

Table D.27: Models for English-Czech: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Czech-English and in the direction English-Czech.

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	238	6.61	78.68	13.45	65.31	22.3	263	4.24	86.77	21.29	58.33	31.2
IBM-1+NN	32	0.89	95.12	46.88	30.61	37.04	68	1.1	93.79	35.29	25	29.27
IBM-1+NNChar	37	1.03	93.12	18.92	14.29	16.28	76	1.23	94.11	40.79	32.29	36.05
IBM-1+BPE	30	0.83	96.46	70	42.86	53.16	44	0.71	96.36	81.82	37.5	51.43
IBM-1+BPE+B	49	1.36	95.98	57.14	57.14	57.14	79	1.27	95.13	53.16	43.75	48
IBM-1+BPE+VAE	62	1.72	93.98	38.71	48.98	43.24	74	1.19	94.43	44.59	34.38	38.82
IBM-1+BPE+VAE+SP	55	1.53	94.84	45.45	51.02	48.08	97	1.56	93.09	32.99	33.33	33.16
IBM-1+BPE+VAE+SP+AC	50	1.39	95.51	52	53.06	52.53	82	1.32	93.89	39.02	33.33	35.96
Fastalign	36	1	96.46	66.67	48.98	56.47	61	0.98	95.45	59.02	37.5	45.86
HMM Giza++	189	5.25	83.94	18.52	71.43	29.41	242	3.9	89.93	30.99	78.12	44.38
HMM+NN	42	1.17	96.65	66.67	57.14	61.54	70	1.13	96.57	72.86	53.12	61.45
HMM+NNCharTgt	35	0.97	97.32	80	57.14	66.67	81	1.31	97.27	77.78	65.62	71.19
HMM+NNCharJB	43	1.19	97.32	74.42	65.31	69.57	81	1.31	96.73	71.6	60.42	65.54
HMM+BPE	34	0.94	97.42	82.35	57.14	67.47	56	0.9	95.93	67.86	39.58	50
HMM+BPE+B	31	0.86	97.32	83.87	53.06	65	58	0.94	96.46	75.86	45.83	57.14
HMM+BPE+VAE	38	1.06	96.46	65.79	51.02	57.47	60	0.97	95.07	53.33	33.33	41.03
HMM+BPE+VAE+SP	24	0.67	97.04	87.5	42.86	57.53	40	0.65	96.25	82.5	34.38	48.53
HMM+BPE+VAE+SP+AC	34	0.94	97.04	76.47	53.06	62.65	48	0.77	96.36	79.17	39.58	52.78
IBM-4 Giza++	159	4.42	87.76	25.16	81.63	38.46	214	3.45	91.22	34.11	76.04	47.1

Table D.28: Models for English-Romanian: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Romanian-English and in the direction English-Romanian.

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	2649	4.73	89.29	25.78	57.44	35.59	1192	4.97	88.47	15.44	58.04	24.39
IBM-1+NN	2326	4.15	90.46	28.2	55.17	37.33	1377	5.74	85.33	8.79	38.17	14.29
IBM-1+NNChar	2439	4.36	89.91	26.61	54.58	35.78	219	0.91	97.45	64.84	44.79	52.99
IBM-1+BPE	1250	2.23	94.81	49.68	52.23	50.92	376	1.57	96.13	41.22	48.9	44.73
IBM-1+BPE+B	1247	2.23	94.9	50.44	52.9	51.64	329	1.37	96.48	45.29	47	46.13
IBM-1+BPE+VAE	1184	2.11	94.83	49.83	49.62	49.73	295	1.23	96.87	51.19	47.63	49.35
IBM-1+BPE+VAE+SP	1291	2.31	95.11	52.36	56.85	54.52	349	1.45	96.73	49	53.94	51.35
IBM-1+BPE+VAE+SP+AC	1338	2.39	95	51.27	57.7	54.29	355	1.48	96.64	47.89	53.63	50.6
Fastalign	684	1.22	95.6	62.72	36.08	45.81	167	0.7	96.97	55.09	29.02	38.02
HMM Giza++	2610	4.66	89.95	28.35	62.24	38.96	1121	4.67	89.49	17.75	62.78	27.68
HMM+NN	1063	1.9	95.54	57.48	51.39	54.26	174	0.72	97.99	83.91	46.06	59.47
HMM+NNCharTgt	789	1.41	96.78	78.2	51.89	62.39	286	1.19	97.42	60.84	54.89	57.71
HMM+NNCharJB	1912	3.41	93.72	43.2	69.47	53.27	336	1.4	95.99	38.1	40.38	39.2
HMM+BPE	1117	1.99	97.25	74.84	70.31	72.51	303	1.26	97.76	65.68	62.78	64.19
HMM+BPE+B	1142	2.04	97.14	73.12	70.23	71.64	287	1.2	98.02	71.08	64.35	67.55
HMM+BPE+VAE	1151	2.06	96.06	62.12	60.13	61.11	278	1.16	97.24	57.91	50.79	54.12
HMM+BPE+VAE+SP	1142	2.04	96.25	64.19	61.65	62.89	273	1.14	97.61	64.84	55.84	60
HMM+BPE+VAE+SP+AC	1108	1.98	96.63	68.59	63.92	66.17	272	1.13	97.79	68.01	58.36	62.82
IBM-4 Giza++	2502	4.47	90.46	29.74	62.57	40.31	1081	4.5	90.06	19.15	65.3	29.61

Table D.29: Models for English-Japanese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Japanese-English and in the direction English-Japanese.

Models	English						Foreign					
	#	FE	ACC	PRE	REC	F1	#	FE	ACC	PRE	REC	F1
IBM-1 Giza++	14589	3	92.22	23.11	56.59	32.82	7181	2.55	90.93	17.1	64.56	27.04
IBM-1+NN	6023	1.24	97.15	57.5	58.11	57.8	2444	0.87	97.1	45.58	58.57	51.27
IBM-1+NNChar	8627	1.78	95.55	38.79	56.15	45.88	2054	0.73	97.89	58.76	63.46	61.02
IBM-1+BPE	6958	1.43	96.04	42.34	49.44	45.61	3927	1.39	94.04	18.74	38.7	25.25
IBM-1+BPE+B	6970	1.44	96.08	42.88	50.16	46.24	3617	1.28	94.46	20.29	38.59	26.6
IBM-1+BPE+VAE	6918	1.42	94.37	20.95	24.32	22.51	3394	1.2	94.64	20.33	36.28	26.06
IBM-1+BPE+VAE+SP	8116	1.67	94.94	31.39	42.76	36.21	3336	1.18	95.24	26.44	46.37	33.68
IBM-1+BPE+VAE+SP+AC	7996	1.65	95.24	34.43	46.2	39.46	3239	1.15	95.4	27.45	46.74	34.58
Fastalign	5808	1.2	96.62	49.62	48.36	48.98	3141	1.11	96.89	44.09	72.82	54.93
HMM Giza++	13469	2.77	93.79	31.23	70.58	43.3	6888	2.44	91.85	20.6	74.61	32.29
HMM+NN	13112	2.7	94.1	32.84	72.26	45.16	2141	0.76	97.87	58.15	65.46	61.59
HMM+NNCharTgt	7931	1.63	96.58	49.3	65.62	56.3	2278	0.81	97.92	58.34	69.87	63.59
HMM+NNCharJB	9402	1.94	96.32	46.99	74.14	57.52	1720	0.61	98.77	79.19	71.61	75.21
HMM+BPE	6003	1.24	97.18	58	58.43	58.22	3951	1.4	95.51	32.52	67.56	43.91
HMM+BPE+B	5966	1.23	97.21	58.53	58.6	58.57	3804	1.35	95.75	34.2	68.4	45.6
HMM+BPE+VAE	5637	1.16	96.19	42.91	40.59	41.72	3188	1.13	96.3	37.45	62.78	46.92
HMM+BPE+VAE+SP	5916	1.22	96.77	51.93	51.55	51.74	3475	1.23	96.26	38.1	69.61	49.25
HMM+BPE+VAE+SP+AC	5723	1.18	97.01	55.71	53.5	54.58	3188	1.13	96.71	42.16	70.66	52.81
IBM-4 Giza++	11283	2.32	95.07	37.69	71.35	49.32	5977	2.12	93.41	25.63	80.55	38.89

Table D.30: Models for English-Vietnamese: # links, fertility (FE), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for unknown target words in the direction Vietnamese-English and in the direction English-Vietnamese.

D.8 Part-of-Speech

D.8.1 Alignment links for content and function words

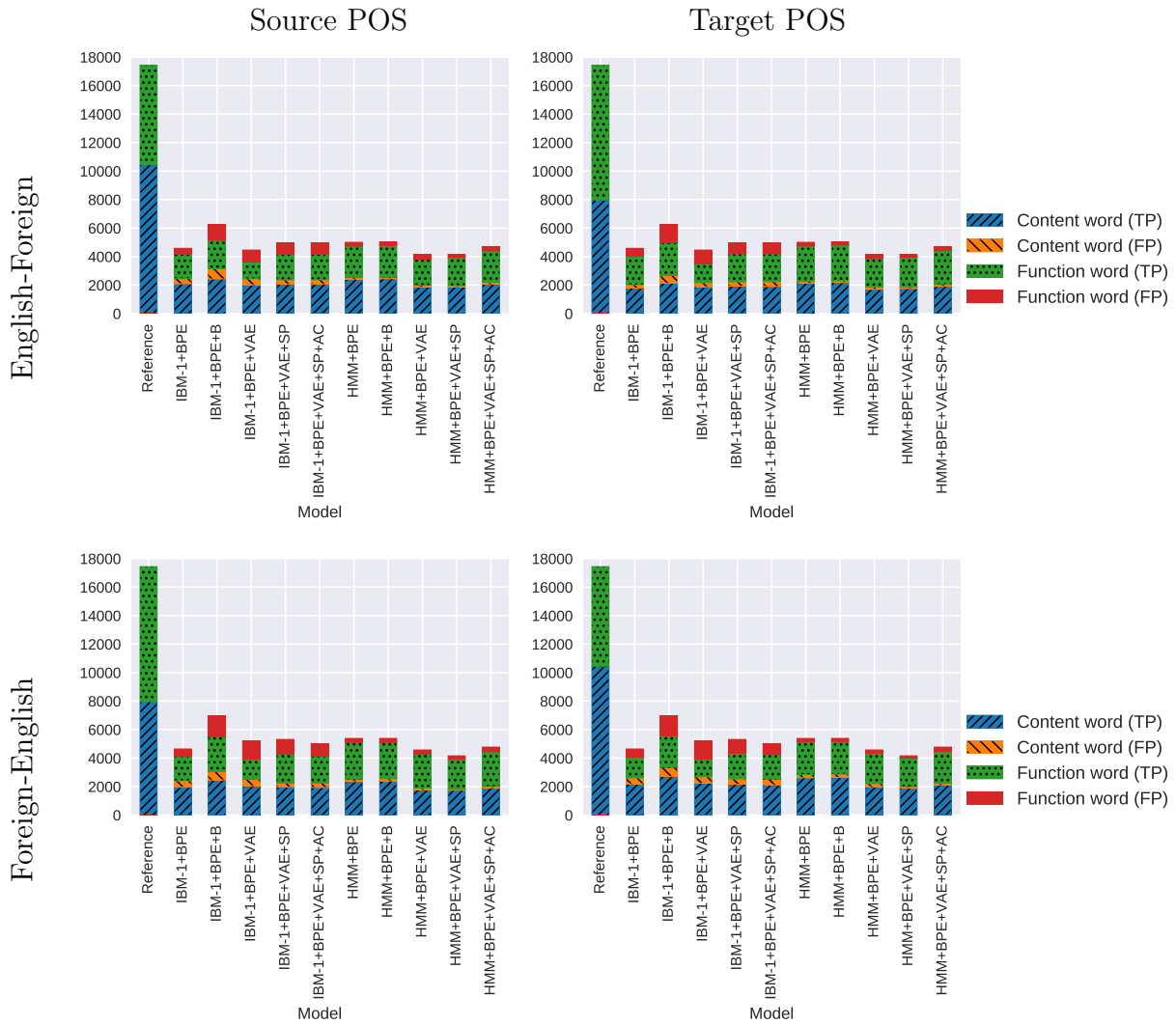


Figure D.43: Models for English-French: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).

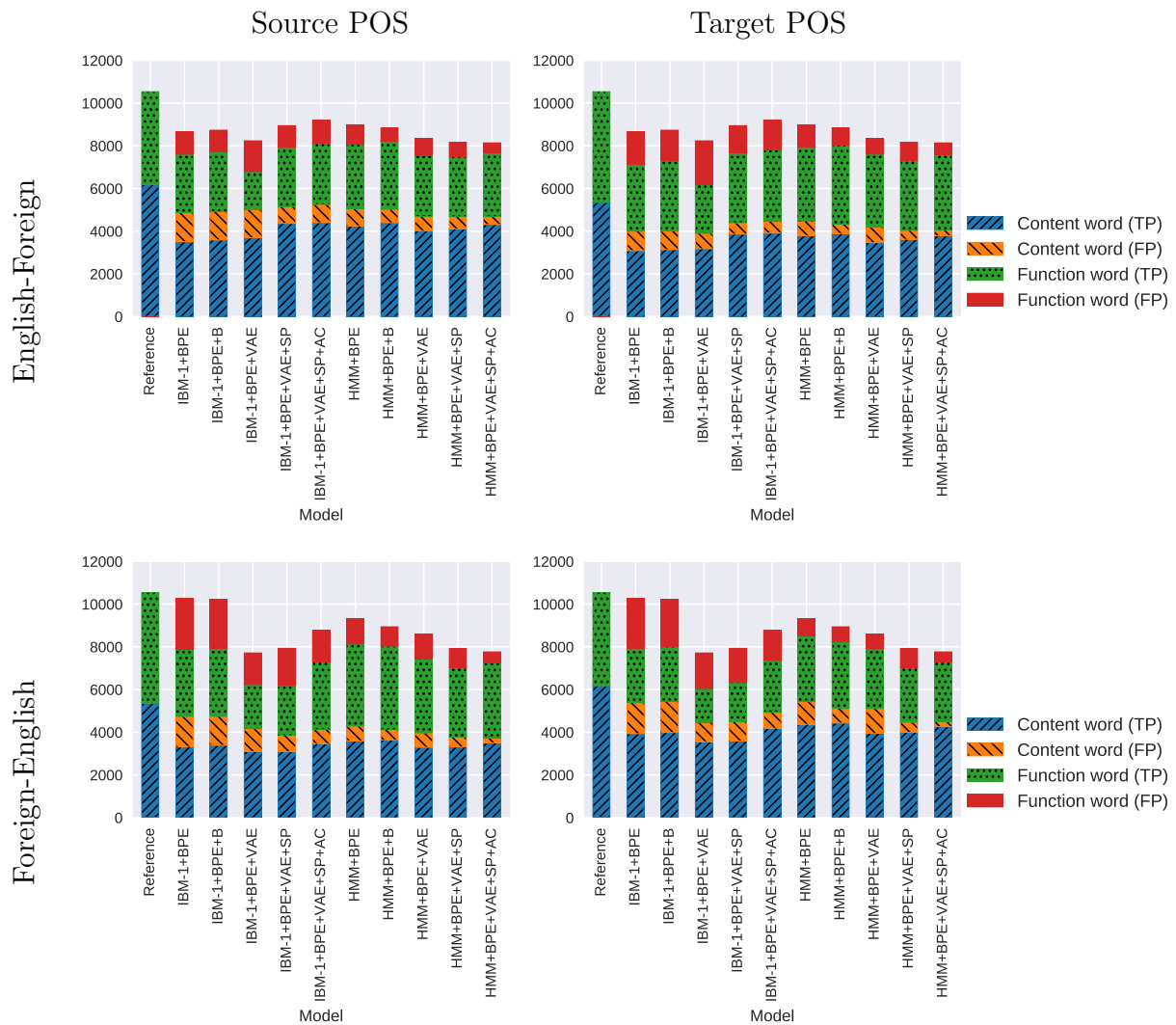


Figure D.44: Models for English-German: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).

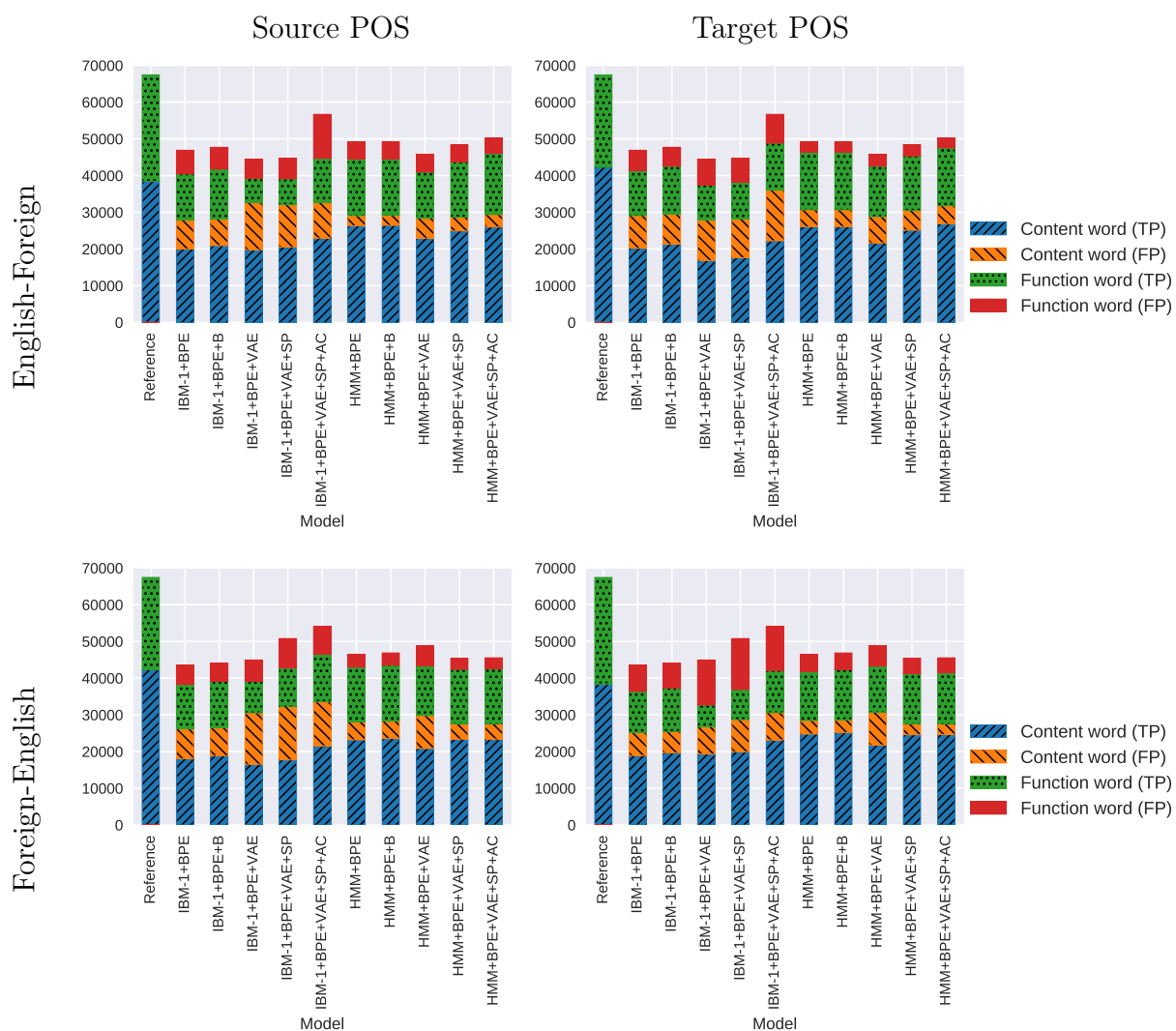


Figure D.45: Models for English-Czech: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).

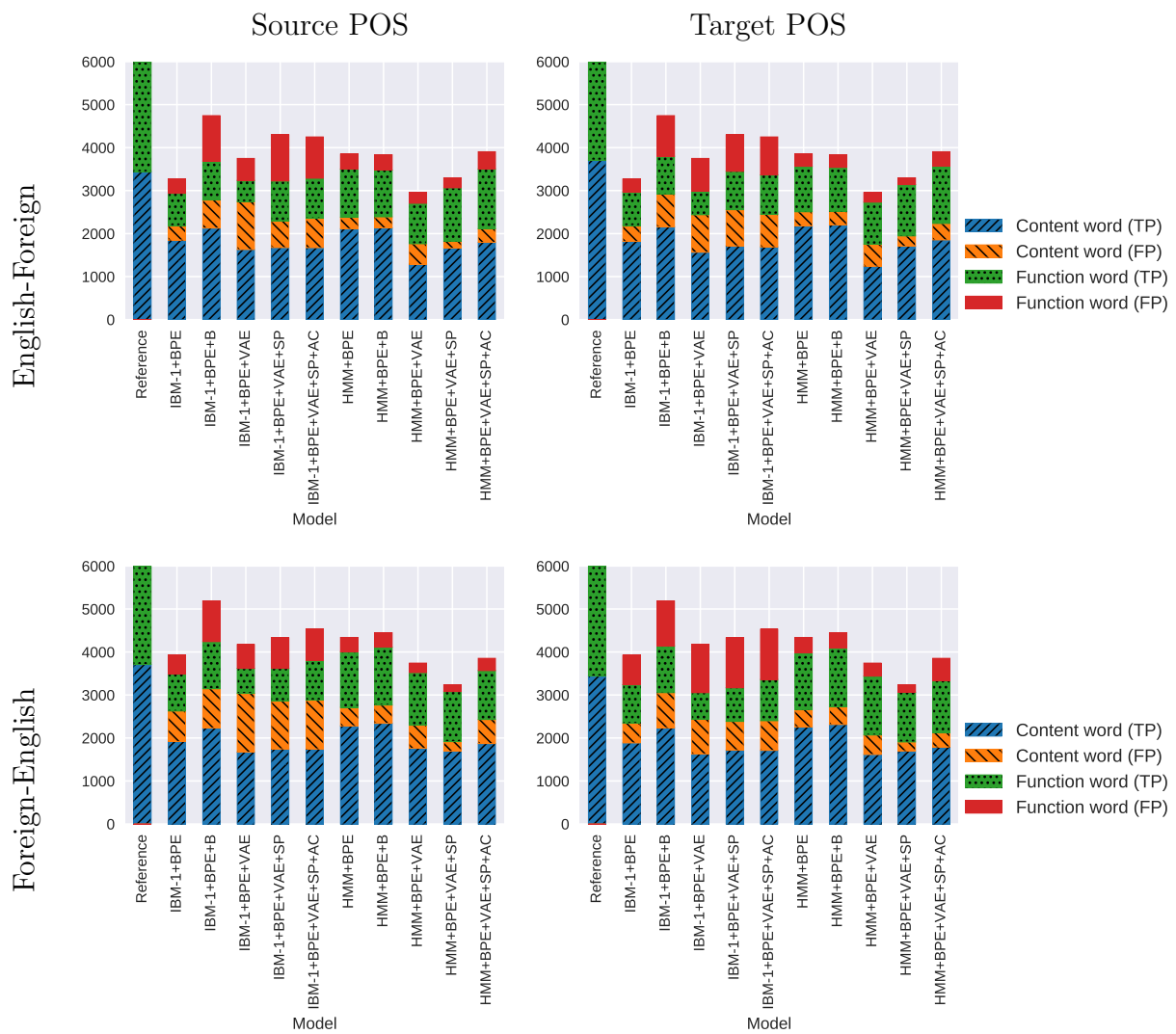


Figure D.46: Models for English-Romanian: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).

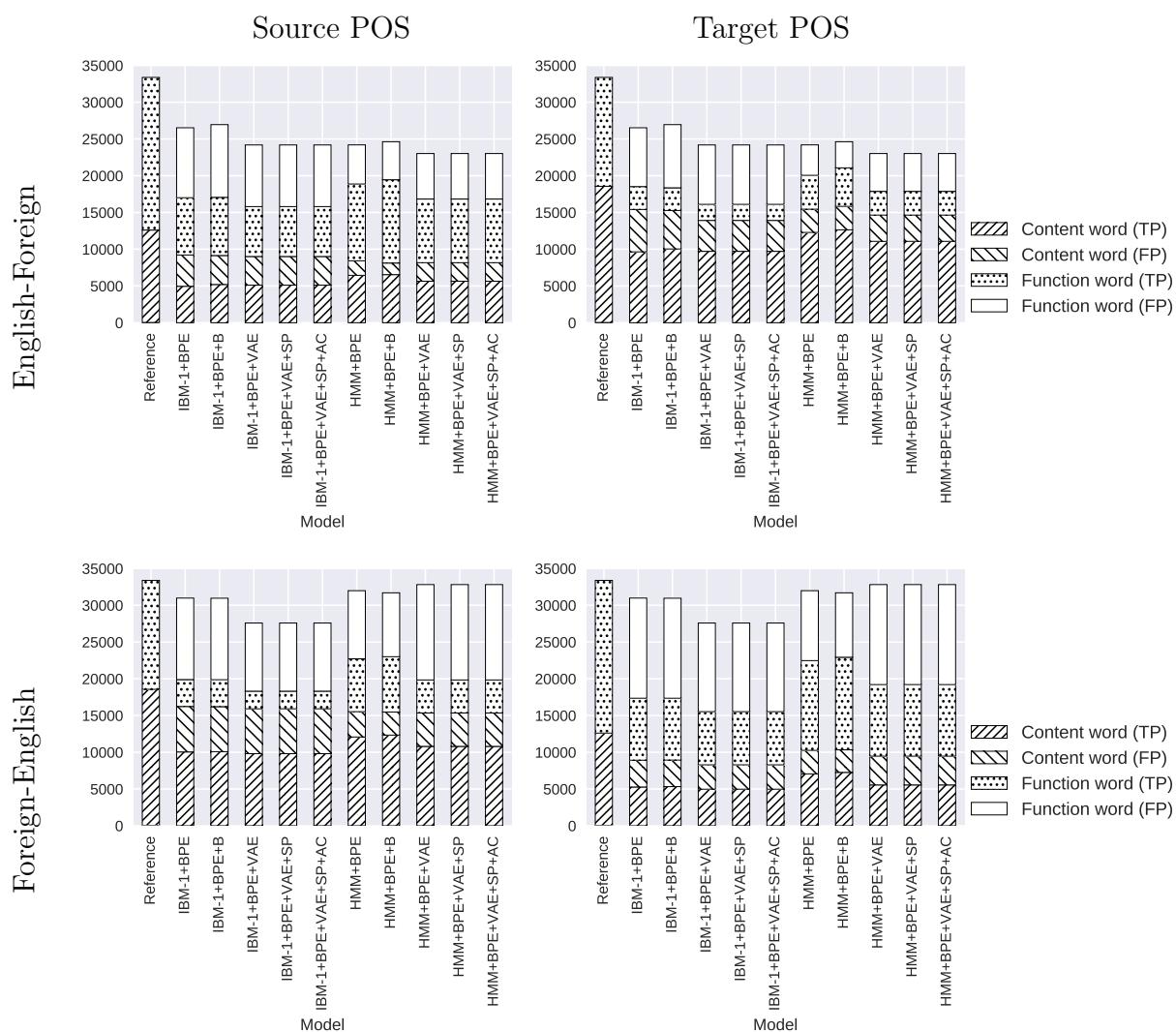


Figure D.47: Models for English-Japanese: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).

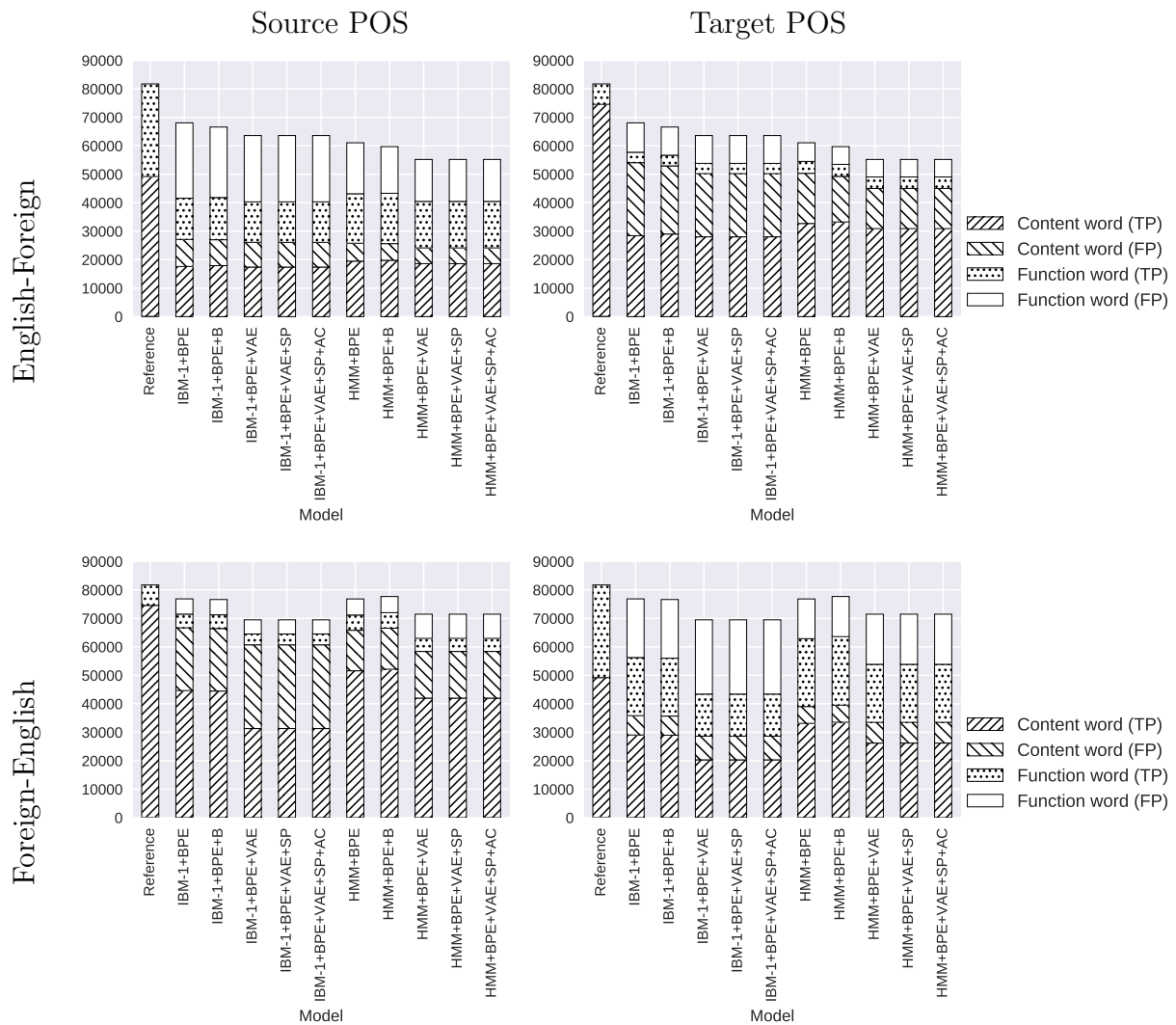


Figure D.48: Models for English-Vietnamese: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).

D.8.2 Unaligned words for content and function words

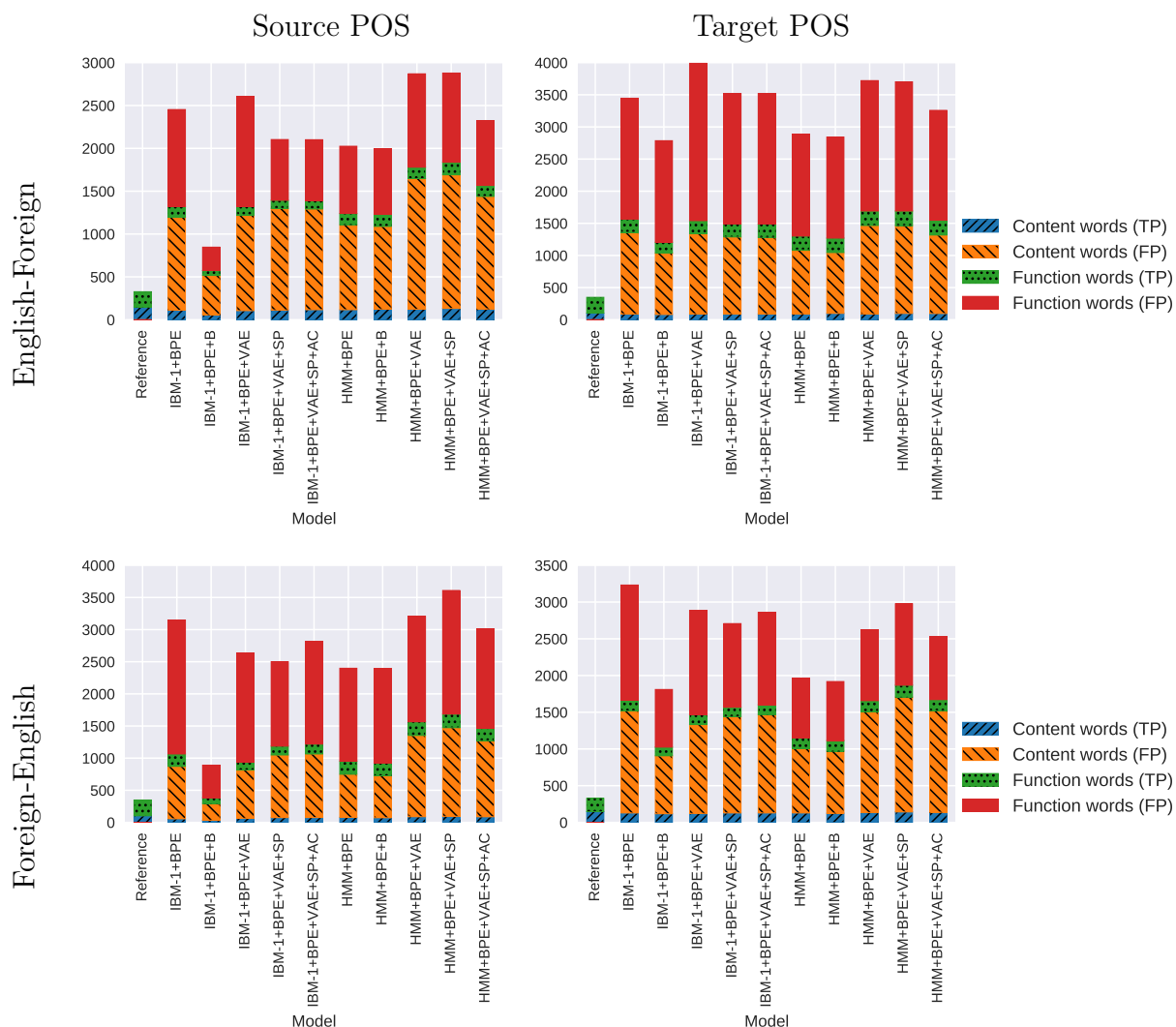


Figure D.49: Models for English-French: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).

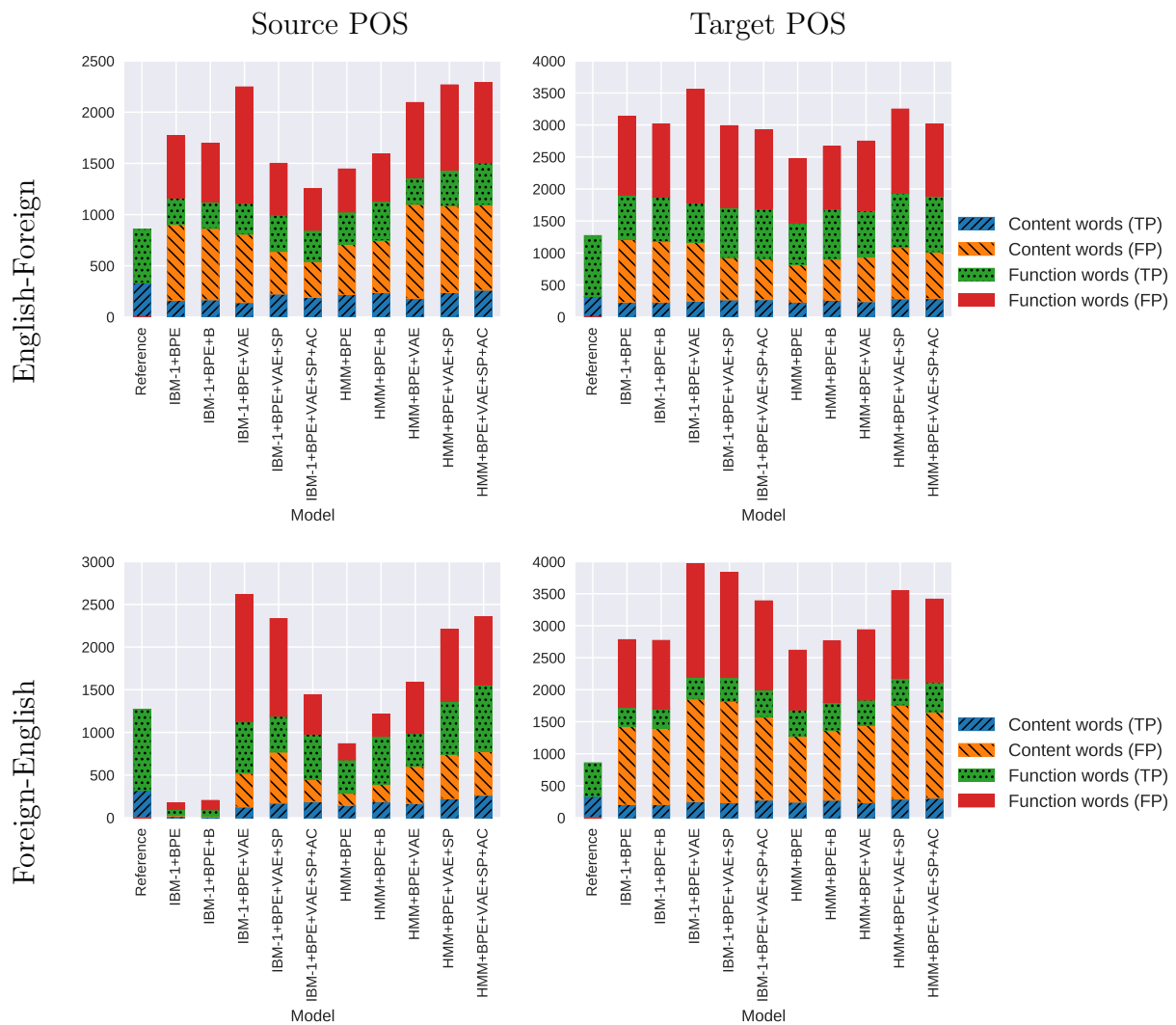


Figure D.50: Models for English-German: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).

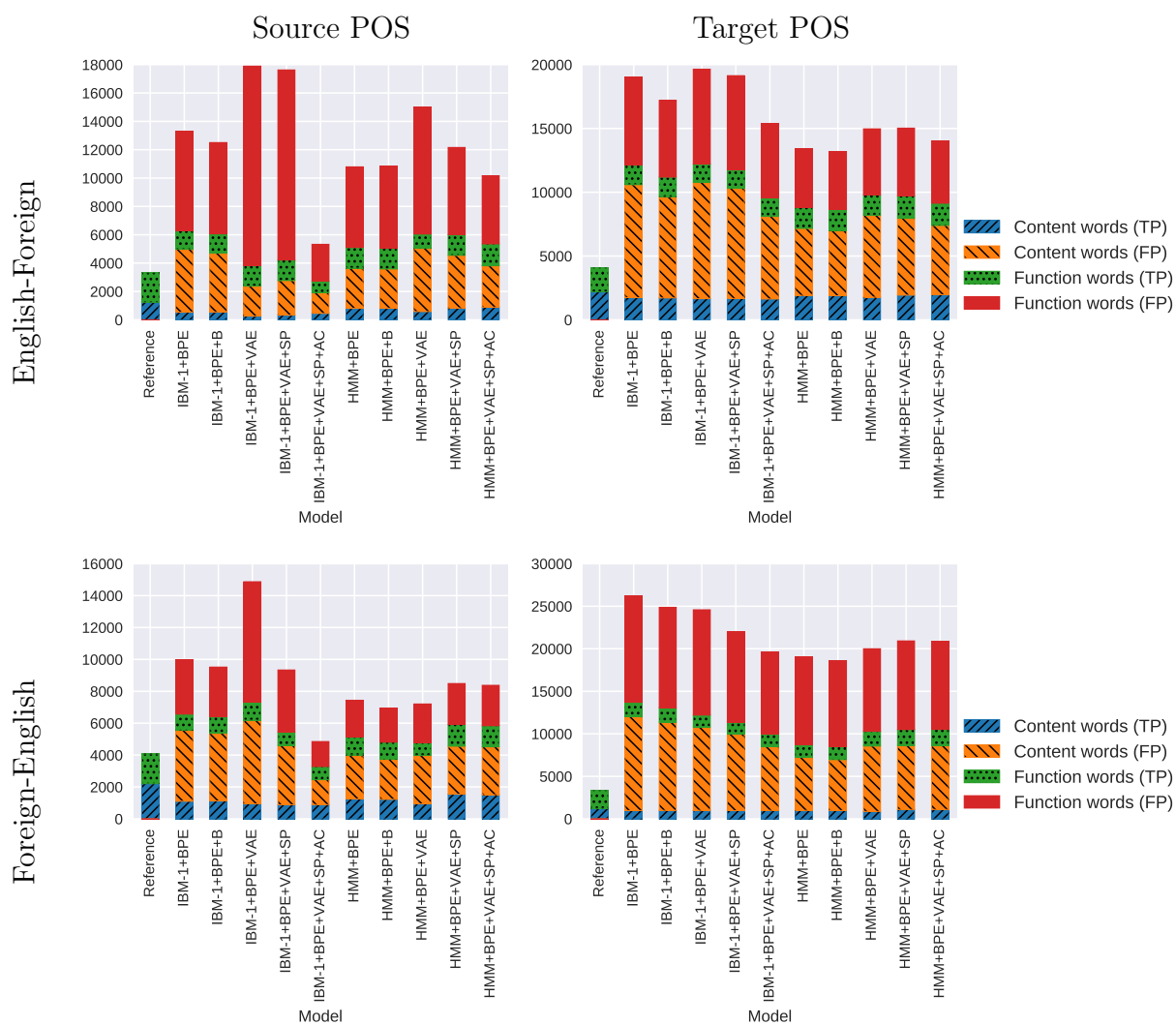


Figure D.51: Models for English-Czech: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).

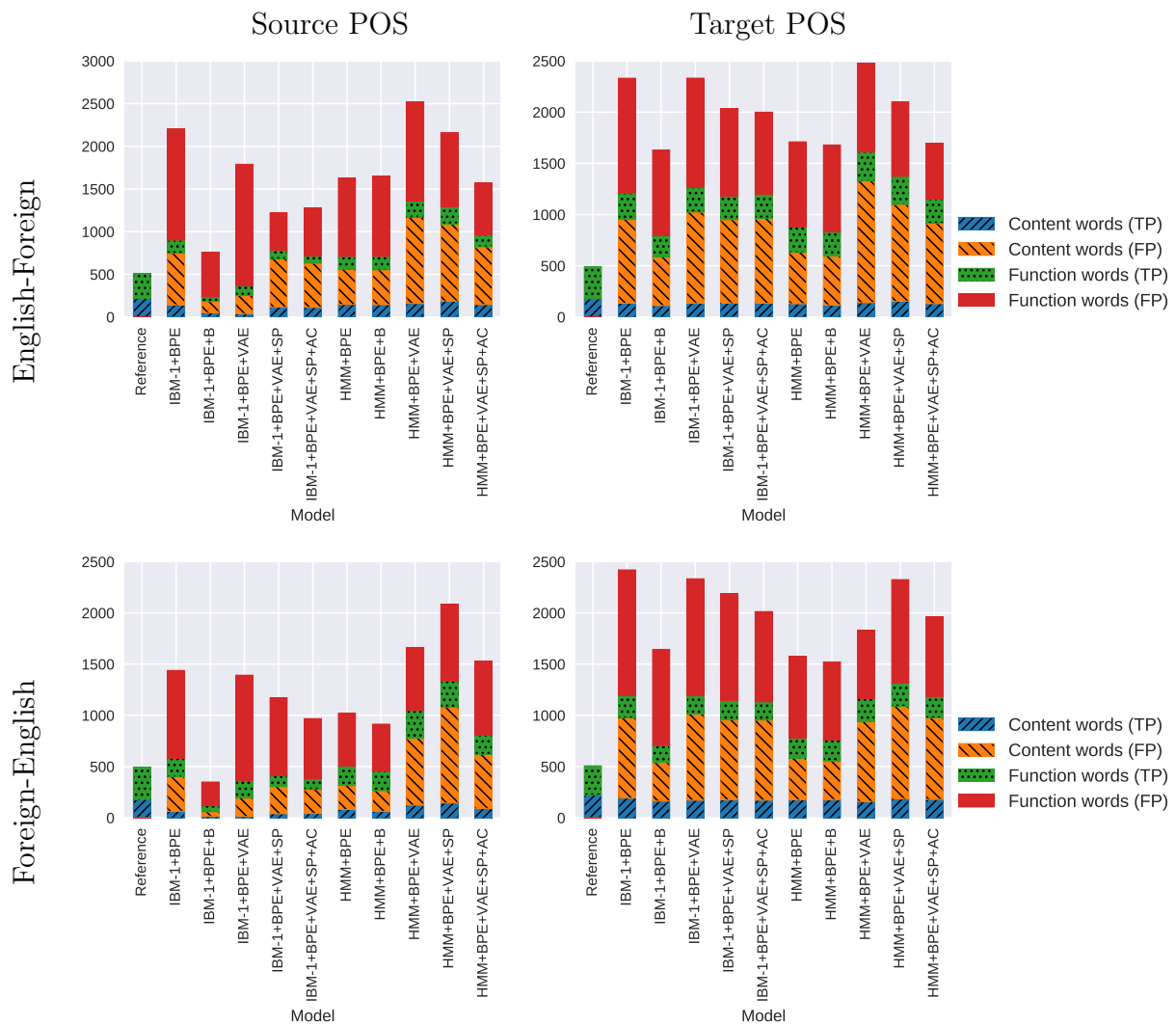


Figure D.52: Models for English-Romanian: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).

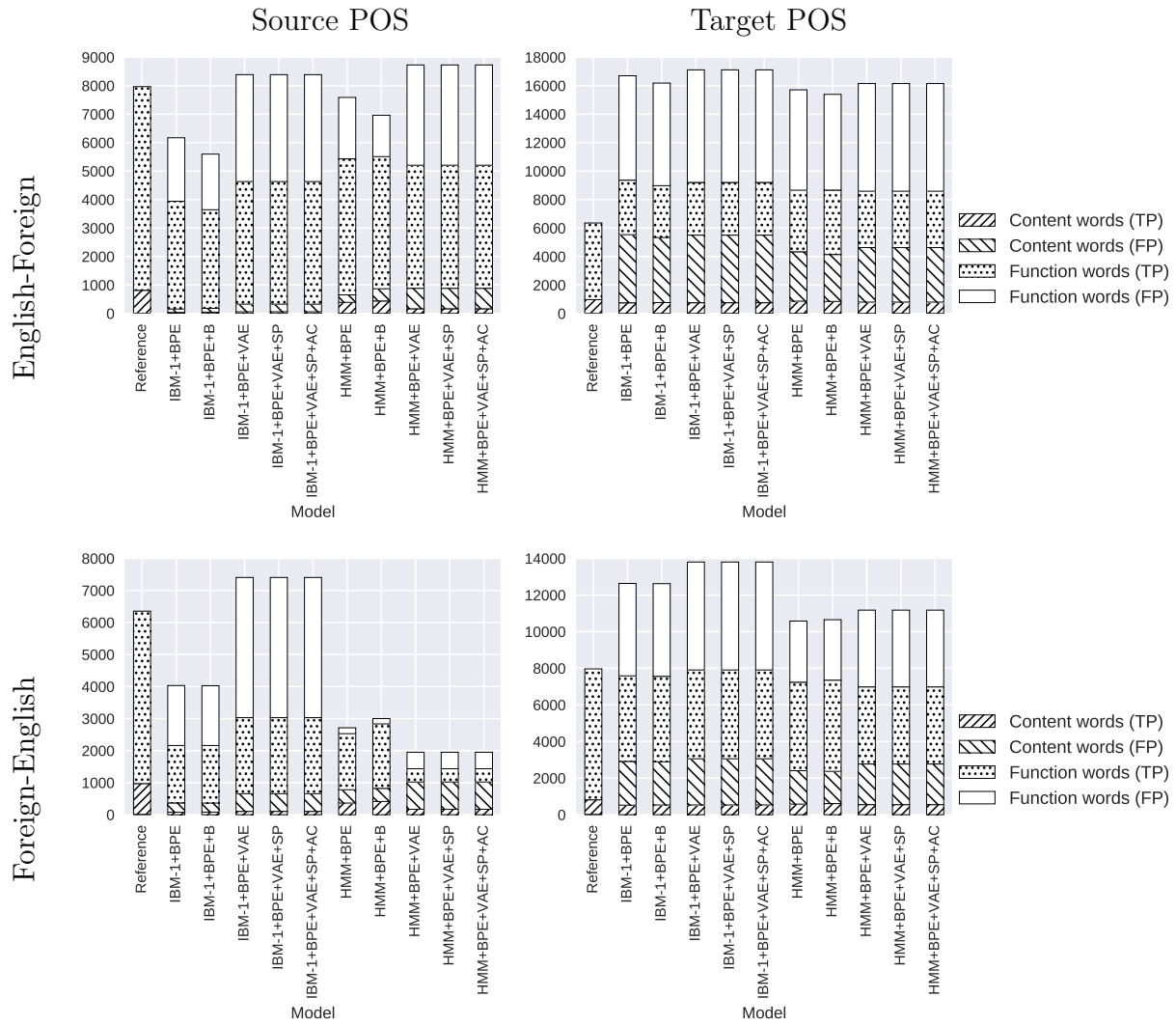


Figure D.53: Models for English-Japanese: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).

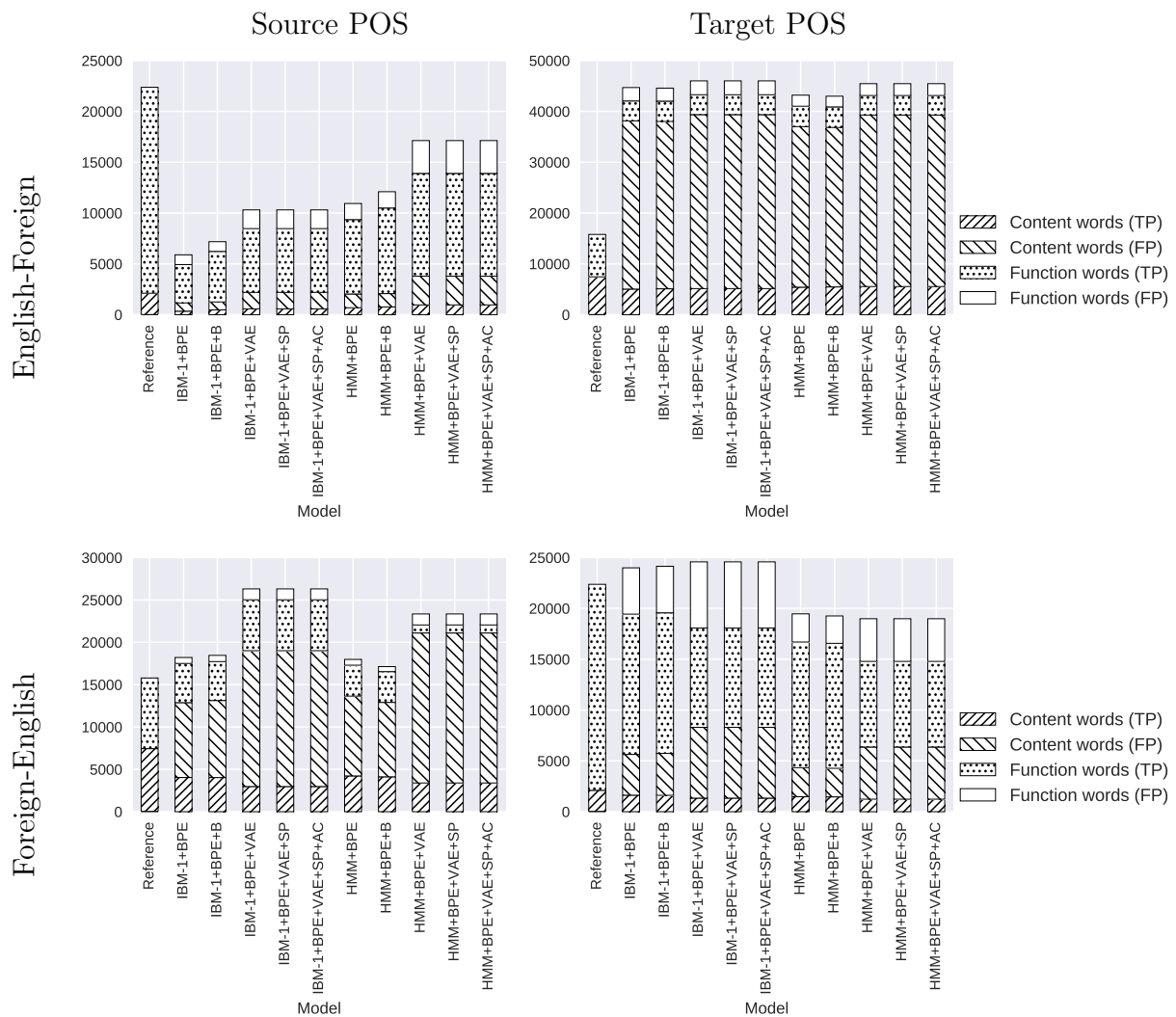


Figure D.54: Models for English-Vietnamese: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).

D.9 Symmetrization

D.9.1 Intersection

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
IBM-1 Giza++	2441	0.31	0.61	36.9	23.04	93.81	13.13	89.3	0.43	0.55
IBM-1+NN	4023	0.59	0.65	21.46	32.78	87.42	20.17	89.91	0.52	0.36
IBM-1+NNChar	4106	0.56	0.61	22.86	32.3	84.73	19.95	89.8	0.26	0.09
IBM-1+BPE	2404	0.52	0.53	32.3	23.54	97.13	13.39	89.39	0.65	0.64
IBM-1+BPE+B	4046	0.58	0.65	18.58	33.6	89.2	20.7	90.02	0.74	0.91
IBM-1+BPE+VAE	2750	0.53	0.62	32.81	24.66	90.51	14.27	89.36	0.65	0.64
IBM-1+BPE+VAE+SP	3949	0.74	0.8	20.26	32.52	88.05	19.94	89.91	0.74	0.64
IBM-1+BPE+VAE+SP+AC	3886	0.78	0.78	21.11	32.02	87.85	19.58	89.86	0.7	0.64
Fastalign	4879	0.69	0.75	11.09	40.48	92.58	25.9	90.71	0.7	0.73
HMM Giza++	4683	0.73	0.76	7.59	41.16	97.2	26.1	90.9	0.65	0.82
HMM+NN	4771	0.73	0.77	7.42	41.53	96.67	26.45	90.92	0.57	0.64
HMM+NNCharTgt	5049	0.78	0.81	6	43.54	96.95	28.07	91.12	0.43	0.64
HMM+NNCharJB	4698	0.8	0.85	6.27	41.62	98.06	26.42	90.96	0.39	0.64
HMM+BPE	3898	0.72	0.78	11.54	36.09	98.77	22.08	90.46	0.65	0.64
HMM+BPE+B	4040	0.75	0.8	10.5	37.12	98.66	22.86	90.56	0.65	0.64
HMM+BPE+VAE	3160	0.69	0.76	18.73	30.16	98.29	17.81	89.94	0.48	0.64
HMM+BPE+VAE+SP	3586	0.86	0.87	13.09	33.5	98.22	20.2	90.22	0.61	0.55
HMM+BPE+VAE+SP+AC	3989	0.84	0.85	10.17	36.35	97.62	22.33	90.46	0.65	0.55
IBM-4 Giza++	4588	0.77	0.81	7.76	40.88	98.13	25.82	90.89	0.65	0.64

Table D.31: Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-French

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
IBM-1 Giza++	5531	0.56	0.53	37.11	59.5	86.4	45.37	97.29	0.68	1
IBM-1+NN	5908	0.63	0.62	36.51	60.22	83.78	47	97.28	0.47	0.5
IBM-1+NNChar	5851	0.61	0.58	36.16	60.55	84.77	47.09	97.31	0.21	0.5
IBM-1+BPE	4926	0.48	0.57	35.52	60.9	95.55	44.69	97.48	1.11	0.5
IBM-1+BPE+B	5085	0.5	0.58	34	62.36	95.77	46.24	97.55	1.26	0.67
IBM-1+BPE+VAE	5101	0.66	0.62	39.05	57.54	88.18	42.7	97.24	1.42	0.83
IBM-1+BPE+VAE+SP	5442	0.69	0.61	33.04	63.41	93.07	48.09	97.57	1.32	0.67
IBM-1+BPE+VAE+SP+AC	5815	0.66	0.63	27.78	68.78	96.68	53.38	97.88	1.32	0.67
Fastalign	6746	0.72	0.68	28.41	68.31	87.49	56.03	97.72	0.84	1
HMM Giza++	6265	0.74	0.69	25.46	70.91	95.07	56.55	97.97	0.84	0.83
HMM+NN	6696	0.76	0.69	27.39	69.43	89.32	56.78	97.81	0.84	0.67
HMM+NNCharTgt	6809	0.76	0.73	26.2	70.55	89.84	58.07	97.87	0.89	0.67
HMM+NNCharJB	6441	0.8	0.76	25.58	70.84	93.34	57.08	97.94	0.79	0.67
HMM+BPE	6147	0.66	0.68	25.33	71.15	96.53	56.34	98	1.58	0.67
HMM+BPE+B	6296	0.7	0.71	22.49	73.98	98.87	59.1	98.18	1.32	0.67
HMM+BPE+VAE	6416	0.75	0.77	24.81	71.49	94.42	57.51	97.99	1.37	0.83
HMM+BPE+VAE+SP	5751	0.73	0.7	27.09	69.17	97.93	53.47	97.91	1.32	0.67
HMM+BPE+VAE+SP+AC	6001	0.77	0.74	24.44	71.92	99.08	56.45	98.07	1.37	0.67
IBM-4 Giza++	6413	0.82	0.78	23.82	72.55	95.85	58.36	98.06	0.89	0.83

Table D.32: Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-German

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
IBM-1 Giza++	23298	0.45	0.4	40.22	45.12	87.85	30.36	97	0.39	0.55
IBM-1+NN	28673	0.56	0.52	34.94	50.34	84.36	35.88	97.13	0.18	0.27
IBM-1+NNChar	27439	0.54	0.47	35.66	49.55	85.65	34.86	97.12	0.12	0.15
IBM-1+BPE	21564	0.5	0.46	38.45	46.61	96.17	30.76	97.14	0.37	0.43
IBM-1+BPE+B	23921	0.54	0.5	33.83	50.55	96.52	34.24	97.28	0.4	0.43
IBM-1+BPE+VAE	22750	0.51	0.51	40.14	45.12	89.41	30.17	97.02	0.55	0.64
IBM-1+BPE+VAE+SP	26170	0.52	0.59	35.68	49.13	87.85	34.1	97.13	0.62	0.7
IBM-1+BPE+VAE+SP+AC	30012	0.56	0.53	29.59	55.26	89.7	39.93	97.37	0.74	0.75
Fastalign	36091	0.71	0.65	20.68	63.06	90.43	48.41	97.7	0.63	0.77
HMM Giza++	28415	0.6	0.55	25.65	57.2	96.46	40.65	97.53	0.51	0.68
HMM+NN	36067	0.73	0.68	20.64	63.4	90.95	48.65	97.72	0.56	0.61
HMM+NNCharTgt	35571	0.68	0.75	18.44	64.83	93.85	49.52	97.82	0.5	0.53
HMM+NNCharJB	33178	0.74	0.7	19.61	63.23	95.86	47.17	97.77	0.34	0.48
HMM+BPE	32573	0.7	0.66	18.06	64.34	98.77	47.72	97.85	0.64	0.67
HMM+BPE+B	33496	0.72	0.68	16.78	65.5	98.67	49.02	97.9	0.65	0.7
HMM+BPE+VAE	30436	0.62	0.66	23.69	59.32	95.36	43.05	97.6	0.63	0.65
HMM+BPE+VAE+SP	31942	0.7	0.66	20.22	63.17	98.25	46.55	97.8	0.67	0.69
HMM+BPE+VAE+SP+AC	32532	0.71	0.65	19.24	64.22	98.67	47.61	97.85	0.68	0.7
IBM-4 Giza++	30648	0.69	0.65	21.43	60.84	97.33	44.24	97.69	0.58	0.73

Table D.33: Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Czech

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
IBM-1 Giza++	1811	0.35	0.51	59.32	40.7	87.63	26.5	97.12	0.49	0.69
IBM-1+NN	3074	0.61	0.59	42.95	57.07	84.12	43.19	97.58	0.22	0.54
IBM-1+NNChar	2680	0.52	0.53	48.03	51.98	84.07	37.63	97.41	0.18	0.08
IBM-1+BPE	2135	0.54	0.65	50.48	49.54	94.24	33.6	97.45	0.29	0.38
IBM-1+BPE+B	2929	0.56	0.62	40.94	59.08	89.93	43.99	97.73	0.55	0.77
IBM-1+BPE+VAE	1878	0.45	0.5	57.86	42.16	88.29	27.69	97.17	0.49	0.62
IBM-1+BPE+VAE+SP	2706	0.63	0.63	49.32	50.7	81.45	36.81	97.33	0.69	0.62
IBM-1+BPE+VAE+SP+AC	2890	0.64	0.68	47.1	52.92	81.28	39.23	97.4	0.67	0.77
Fastalign	3659	0.73	0.72	32.85	67.17	88.55	54.11	98.03	0.67	0.92
HMM Giza++	3119	0.67	0.66	35.7	64.32	93.91	48.91	97.98	0.6	0.92
HMM+NN	3247	0.61	0.66	36.11	63.91	90.88	49.28	97.92	0.29	0.23
HMM+NNCharTgt	3751	0.81	0.72	28.52	71.51	92.83	58.15	98.27	0.58	0.62
HMM+NNCharJB	3587	0.78	0.77	29.53	70.5	94.09	56.36	98.24	0.29	0.46
HMM+BPE	3087	0.71	0.8	36.22	63.8	93.78	48.35	97.95	0.47	0.69
HMM+BPE+B	3153	0.71	0.82	35.76	64.26	93.15	49.05	97.96	0.55	0.54
HMM+BPE+VAE	1963	0.52	0.66	53.78	46.23	93.63	30.69	97.34	0.36	0.54
HMM+BPE+VAE+SP	2816	0.87	0.85	39.59	60.43	94.46	44.42	97.83	0.49	0.46
HMM+BPE+VAE+SP+AC	2970	0.77	0.76	38.44	61.58	92.86	46.06	97.86	0.56	0.54
IBM-4 Giza++	3315	0.77	0.72	32.19	67.83	95.17	52.69	98.14	0.69	0.85

Table D.34: Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Romanian

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
IBM-1 Giza++	9540	0.31	0.34	64.7	35.3	79.4	22.7	98.01	0.58	0.57
IBM-1+NN	3921	0.15	0.19	79.82	20.18	96	11.28	97.87	0.18	0.11
IBM-1+NNChar	4789	0.16	0.25	75.63	24.37	97.1	13.93	97.94	0.28	0.33
IBM-1+BPE	10703	0.35	0.4	59.25	40.75	83.91	26.91	98.13	0.78	0.79
IBM-1+BPE+B	10960	0.35	0.41	58.22	41.78	84.52	27.75	98.15	0.71	0.74
IBM-1+BPE+VAE	10548	0.38	0.44	59.7	40.3	83.9	26.52	98.13	0.69	0.78
IBM-1+BPE+VAE+SP	12515	0.42	0.43	56.79	43.21	79.22	29.71	98.14	0.84	0.78
IBM-1+BPE+VAE+SP+AC	12736	0.4	0.43	56.15	43.85	79.38	30.29	98.15	0.85	0.88
Fastalign	15622	0.54	0.59	51.23	48.77	76.49	35.8	98.21	0.53	0.67
HMM Giza++	9372	0.37	0.4	61.12	38.88	88.67	24.9	98.13	0.53	0.57
HMM+NN	6579	0.23	0.35	67.51	32.49	98.65	19.44	98.07	0.23	0.33
HMM+NNCharTgt	8057	0.34	0.32	61.45	38.55	99.13	23.93	98.18	0.36	0.4
HMM+NNCharJB	8652	0.3	0.48	58.96	41.04	99.69	25.84	98.23	0.31	0.18
HMM+BPE	13048	0.41	0.54	46.09	53.91	95.91	37.49	98.47	0.78	0.92
HMM+BPE+B	14168	0.45	0.58	43.03	56.97	95.6	40.58	98.54	0.8	0.9
HMM+BPE+VAE	12458	0.38	0.54	52.41	47.59	87.55	32.68	98.28	0.75	0.8
HMM+BPE+VAE+SP	14597	0.46	0.58	47	53	87.09	38.09	98.39	0.65	0.81
HMM+BPE+VAE+SP+AC	14600	0.47	0.6	45.32	54.68	89.84	39.3	98.45	0.69	0.85
IBM-4 Giza++	10137	0.43	0.48	56.29	43.71	93.82	28.5	98.25	0.56	0.63

Table D.35: Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Japanese

Models	# links	Ratio		AER	F1	PRE	REC	ACC	FE	
		En-XX	XX-En						En	XX
IBM-1 Giza++	20603	0.23	0.46	67.18	32.82	81.52	20.55	97.26	0.62	0.39
IBM-1+NN	15580	0.19	0.34	68.68	31.33	97.83	18.65	97.33	0.34	0.32
IBM-1+NNChar	17882	0.26	0.39	65.23	34.77	96.85	21.19	97.41	0.5	0.4
IBM-1+BPE	29083	0.38	0.43	53.2	46.8	89.17	31.73	97.65	0.41	0.38
IBM-1+BPE+B	29282	0.38	0.44	52.88	47.13	89.34	32.01	97.66	0.4	0.38
IBM-1+BPE+VAE	28837	0.42	0.45	58.72	41.28	79.15	27.92	97.41	0.41	0.35
IBM-1+BPE+VAE+SP	36607	0.44	0.56	50.5	49.51	80.03	35.84	97.62	0.46	0.46
IBM-1+BPE+VAE+SP+AC	36883	0.45	0.57	50.03	49.97	80.36	36.26	97.63	0.42	0.46
Fastalign	48975	0.55	0.73	46.21	53.79	71.78	43.01	97.59	0.78	0.71
HMM Giza++	35598	0.41	0.58	56.09	43.92	72.38	31.52	97.38	0.67	0.47
HMM+NN	24457	0.29	0.45	54.43	45.57	98.94	29.6	97.7	0.42	0.52
HMM+NNCharTgt	24144	0.35	0.47	55.2	44.81	98.24	29.02	97.67	0.55	0.52
HMM+NNCharJB	25998	0.32	0.63	52.04	47.96	99.38	31.61	97.76	0.56	0.58
HMM+BPE	36674	0.48	0.6	45.74	54.26	87.6	39.3	97.84	0.51	0.51
HMM+BPE+B	37129	0.48	0.62	45.24	54.77	87.66	39.82	97.86	0.55	0.51
HMM+BPE+VAE	36378	0.51	0.66	50.63	49.37	80.15	35.67	97.62	0.46	0.43
HMM+BPE+VAE+SP	42058	0.56	0.72	46.07	53.94	79.38	40.84	97.73	0.61	0.59
HMM+BPE+VAE+SP+AC	39910	0.55	0.7	44.31	55.7	84.88	41.45	97.85	0.62	0.59
IBM-4 Giza++	38925	0.48	0.68	51.18	48.82	75.67	36.04	97.54	0.72	0.55

Table D.36: Intersection alignment: The number of alignment links, their ratio to the total number of alignment links predicted by the model, alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE), recall (REC) and average fertility (FE) for English-Vietnamese

D.9.2 Grow-diag-final (GDF)

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
BM-1 Giza++	40.1	26.7	33.9	36.49	25.19	33.83	82.75	21.26
IBM-1+NN	27.96	36.42	27.21	38.08	17.86	39.48	82.89	25.91
IBM-1+NNChar	28.76	37.5	31.4	37.21	19.62	40.07	79.39	26.8
IBM-1+BPE	22.98	34.56	28.69	33.12	23.24	33.46	84.41	20.87
IBM-1+BPE+B	25.71	37.27	24.05	40.42	14.61	41.68	84.45	27.66
IBM-1+BPE+VAE	33.42	29.21	34.36	30.96	24.91	30.76	86.68	18.7
IBM-1+BPE+VAE+SP	22.12	34.73	23.89	34.97	16.89	35.98	86.2	22.74
IBM-1+BPE+VAE+SP+AC	22.87	34.44	23.61	34.35	17.85	35.27	86.52	22.15
Fastalign	15.19	44.98	16.23	46.32	14.25	49.05	81.84	35.02
HMM Giza++	11.99	45.18	11.97	45.98	8.54	47.5	90.78	32.17
HMM+NN	11.84	45.57	11.15	46.86	9.8	48.07	88.85	32.95
HMM+NNCharTgt	9.17	47.22	9.56	47.87	8.41	48.99	90.12	33.64
HMM+NNCharJB	8.47	44.38	7.74	46.26	6.81	45.83	93.32	30.38
HMM+BPE	10.91	41.1	10.97	42.98	10.22	42.48	92.76	27.55
HMM+BPE+B	9.84	41.74	10.48	43.2	9.15	43.01	93.76	27.9
HMM+BPE+VAE	18.92	34.11	12.94	37.53	13.92	35.47	93.94	21.86
HMM+BPE+VAE+SP	12.94	35.86	12.27	36.09	11.78	35.16	96.87	21.48
HMM+BPE+VAE+SP+AC	11.47	38.73	10.84	39.2	9.64	38.59	94.9	24.22
IBM-4 Giza++	10	44.43	9.64	45.43	7.03	46.32	93.55	30.78

Table D.37: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-French

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
BM-1 Giza++	39.03	58.76	42.66	55.39	32.31	65.06	72.19	59.2
IBM-1+NN	37.64	60.07	39.22	58.53	30.82	66.36	76.51	58.59
IBM-1+NNChar	36.22	61.55	40.88	56.99	30.29	66.89	76.44	59.47
IBM-1+BPE	32.91	64.95	35.54	62.35	26.71	70.57	82.89	61.44
IBM-1+BPE+B	31.36	66.52	34.46	63.34	25.95	71.29	82.9	62.54
IBM-1+BPE+VAE	38.92	58.4	40.87	56.51	33.82	62.94	82.07	51.05
IBM-1+BPE+VAE+SP	24.41	73.38	38.72	58.67	23.51	73.5	89.32	62.43
IBM-1+BPE+VAE+SP+AC	24.3	73.6	29.37	68.64	18.27	79.3	92.4	69.46
Fastalign	28.98	68.75	31.28	66.47	28.21	69.84	70.12	69.57
HMM Giza++	23.92	73.3	26.33	71.04	22.87	74.48	81.74	68.39
HMM+NN	26.78	70.95	29.44	68.21	26.92	71	73.93	68.29
HMM+NNCharTgt	26.04	71.57	28.11	69.48	26.26	71.54	75.18	68.23
HMM+NNCharJB	23.69	73.38	24.9	72.16	22.42	74.71	84.83	66.74
HMM+BPE	23.66	74.28	22.75	74.94	20.73	77.15	83.65	71.59
HMM+BPE+B	19.61	78.25	20.38	77.38	16.66	81.13	90.07	73.8
HMM+BPE+VAE	23.96	72.98	26.5	70.63	23.54	73.47	83.17	65.8
HMM+BPE+VAE+SP	23.73	73.73	26.4	71.13	20.98	76.11	93.68	64.08
HMM+BPE+VAE+SP+AC	19.13	78.38	20.58	76.98	17	80.36	96.15	69.03
IBM-4 Giza++	21.46	75.48	23.31	73.63	19.59	77.29	89.26	68.16

Table D.38: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-German

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
BM-1 Giza++	45.09	46.75	48.47	42.88	35.47	52.67	71.16	41.81
IBM-1+NN	42.29	48.64	40.97	49.08	30.89	56.25	75.65	44.77
IBM-1+NNChar	40.85	50.35	42.35	47.99	29.85	57.16	77.02	45.44
IBM-1+BPE	34.3	56.91	35.54	54.2	29	59.77	81.16	47.31
IBM-1+BPE+B	30.76	59.86	32.71	56.45	26.12	62.39	82.18	50.28
IBM-1+BPE+VAE	41.92	47.3	45.3	44.9	34.46	51.55	80.98	37.81
IBM-1+BPE+VAE+SP	39.61	49.2	42.63	47.6	31.13	54.67	80.42	41.41
IBM-1+BPE+VAE+SP+AC	35.41	56.47	33.83	56.95	22.73	65.83	83.87	54.18
Fastalign	25.75	64.09	25.3	62.86	23.3	66.93	72.95	61.82
HMM Giza++	27.86	61.22	30.38	57.28	25.25	62.96	75.67	53.91
HMM+NN	23.5	65.45	24.06	64.03	22.18	67.27	75.3	60.78
HMM+NNCharTgt	16.74	69.36	24.61	63.94	18.87	69.16	79.35	61.28
HMM+NNCharJB	16.38	69.35	23.88	62.72	16.18	69.5	86.22	58.21
HMM+BPE	16.68	71.66	19.87	67.02	13.66	73.62	90.83	61.89
HMM+BPE+B	16.24	71.87	18.74	67.93	12.93	74.2	90.82	62.72
HMM+BPE+VAE	23.71	62.57	29.01	59.05	22.35	63.92	81.03	52.78
HMM+BPE+VAE+SP	20.31	69.01	20.12	67.94	14.47	73.66	93.17	60.9
HMM+BPE+VAE+SP+AC	17.56	72.53	20.1	68.03	12.86	76.14	93.8	64.07
IBM-4 Giza++	20.92	65.7	26.5	59.81	19.13	66.67	84.22	55.17

Table D.39: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Czech

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
BM-1 Giza++	56.02	43.99	53.52	46.49	51.13	48.88	73.82	36.54
IBM-1+NN	46.4	53.62	44.9	55.11	38.38	61.64	75.8	51.94
IBM-1+NNChar	50.16	49.85	48.28	51.73	41.42	58.6	75.25	47.98
IBM-1+BPE	43.77	56.25	43.96	56.05	41.96	58.06	78.17	46.18
IBM-1+BPE+B	43.47	56.54	40.42	59.59	34.46	65.56	80.47	55.31
IBM-1+BPE+VAE	56.39	43.63	55.7	44.32	51.34	48.67	79.6	35.05
IBM-1+BPE+VAE+SP	49.3	50.71	51.49	48.52	45.21	54.81	76.23	42.79
IBM-1+BPE+VAE+SP+AC	49.12	50.89	49.2	50.81	43.37	56.65	76.91	44.84
Fastalign	33.36	66.65	32.91	67.1	30.42	69.6	72.65	66.8
HMM Giza++	33.36	66.65	36.38	63.64	32.4	67.62	76.63	60.5
HMM+NN	30.69	69.33	40.12	59.89	34.36	65.66	65.89	65.43
HMM+NNCharTgt	27.54	72.47	28.01	72.01	25.97	74.05	79.04	69.66
HMM+NNCharJB	26.85	73.17	29.76	70.26	26.13	73.89	83.22	66.43
HMM+BPE	34.23	65.79	30.75	69.27	30.26	69.76	83.34	59.99
HMM+BPE+B	34.41	65.61	29.34	70.68	29.45	70.57	83.31	61.21
HMM+BPE+VAE	50.29	49.73	38.64	61.38	42.73	57.29	83.13	43.7
HMM+BPE+VAE+SP	37.52	62.5	38.04	61.98	37.03	62.99	91.62	48
HMM+BPE+VAE+SP+AC	35.55	64.47	38.87	61.15	35.19	64.83	84.51	52.59
IBM-4 Giza++	31.04	68.98	32.3	67.72	29.08	70.94	82.98	61.96

Table D.40: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Romanian

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
BM-1 Giza++	63.12	36.88	61.55	38.45	55.31	44.69	64	34.33
IBM-1+NN	62.64	37.36	56.9	43.1	48.3	51.7	47.99	56.02
IBM-1+NNChar	57.96	42.04	54.91	45.09	45.16	54.84	50.4	60.16
IBM-1+BPE	57.41	42.59	57.39	42.61	51.45	48.55	74.13	36.09
IBM-1+BPE+B	56.51	43.49	57.27	42.73	50.69	49.31	73.85	37.01
IBM-1+BPE+VAE	58.66	41.34	59.95	40.05	54.33	45.67	73.53	33.12
IBM-1+BPE+VAE+SP	55.54	44.43	55.1	44.9	48.38	51.62	71.96	40.24
IBM-1+BPE+VAE+SP+AC	54.81	45.18	54.23	45.76	47.19	52.81	72.21	41.63
Fastalign	50.67	49.33	49.37	50.63	44.79	55.21	64.99	47.99
HMM Giza++	57.01	42.99	54.41	45.59	52.18	47.82	62.37	38.78
HMM+NN	49.68	50.32	47.09	52.91	35.97	64.03	60.63	67.84
HMM+NNCharTgt	43.92	56.08	43.6	56.4	31.55	68.45	64.21	73.29
HMM+NNCharJB	40.92	59.08	37.71	62.29	25.52	74.48	72.96	76.07
HMM+BPE	41.22	58.69	41.01	58.97	37.05	62.95	77.13	53.18
HMM+BPE+B	38.33	61.65	38.93	61.04	34.4	65.6	78.22	56.49
HMM+BPE+VAE	49.27	50.73	53.78	46.22	48.06	51.94	63.92	43.75
HMM+BPE+VAE+SP	43.67	56.32	48.99	51.01	41.54	58.46	75.57	47.67
HMM+BPE+VAE+SP+AC	40.86	59.12	45.24	54.76	38.52	61.48	79.08	50.28
IBM-4 Giza++	52.52	47.48	49.23	50.77	46.68	53.32	74.58	41.49

Table D.41: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Japanese

Models	English-Foreign		Foreign-English		GDF			
	AER	F1	AER	F1	AER	F1	PRE	REC
BM-1 Giza++	69.43	30.58	46.45	53.56	47.91	52.09	70.78	41.21
IBM-1+NN	58.87	41.13	42.25	57.75	47.3	52.7	88.34	37.56
IBM-1+NNChar	55.06	44.94	41.15	58.86	45.52	54.49	86.59	39.75
IBM-1+BPE	57.17	42.83	37.59	62.38	36.02	63.99	80.4	53.14
IBM-1+BPE+B	55.85	44.15	37.72	62.26	35.52	64.48	80.85	53.62
IBM-1+BPE+VAE	56.47	43.53	53.56	46.44	47.42	52.59	71.67	41.53
IBM-1+BPE+VAE+SP	51.34	48.61	41.38	58.63	35.44	64.57	76.18	56.03
IBM-1+BPE+VAE+SP+AC	50.84	49.09	38.77	61.23	33.92	66.09	77.09	57.83
Fastalign	48.89	51.11	32.82	67.19	32.9	67.11	64.41	70.05
HMM Giza++	57.86	42.14	37.57	62.43	39.79	60.21	61.51	58.97
HMM+NN	49.27	50.73	31.45	68.56	28.74	71.27	90.09	58.95
HMM+NNCharTgt	47.52	52.49	30.94	69.06	31.98	68.03	91.09	54.29
HMM+NNCharJB	43.28	56.73	27.59	72.42	25.32	74.69	93.47	62.19
HMM+BPE	48.34	51.66	28.16	71.85	29.62	70.38	77.02	64.8
HMM+BPE+B	47.03	52.97	27.76	72.24	28.87	71.14	77.04	66.07
HMM+BPE+VAE	48.97	51.04	39.2	60.8	37.36	62.65	71.36	55.83
HMM+BPE+VAE+SP	45.87	54.13	33.78	66.22	32.07	67.94	76.65	61
HMM+BPE+VAE+SP+AC	43.94	56.05	32.59	67.41	30.66	69.35	79.62	61.43
IBM-4 Giza++	51.91	48.1	33.19	66.82	33.29	66.71	70.25	63.51

Table D.42: Grow-diag-final: Alignment error rate (AER), F-score (F1) for English-Vietnamese

D.10 Training with monolingual data

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	R-ACC	AER	F1	PRE	REC	R-ACC
	IBM-1+BPE+VAE									
+SP	39.61	49.2	61.7	40.91	61.90	42.63	47.6	55.41	41.72	76.28
+Mono+Noise	37.25	53.6	61.22	47.67	62.04	39.28	51.12	58.84	45.19	76.43
	HMM+BPE+VAE+SP									
+SP	20.31	69.01	82.62	59.25	97.05	20.12	67.94	84.46	56.83	97.22
+Mono+Noise	19.11	69.25	86.8	57.6	97.39	18.51	68.64	88.95	55.88	97.28

Table D.43: Training with a monolingual corpus and the noise model (+Noise) on English-Czech corpus. R-Acc is the accuracy of the reconstruction model.

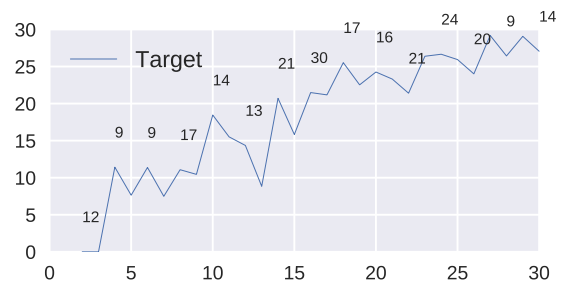
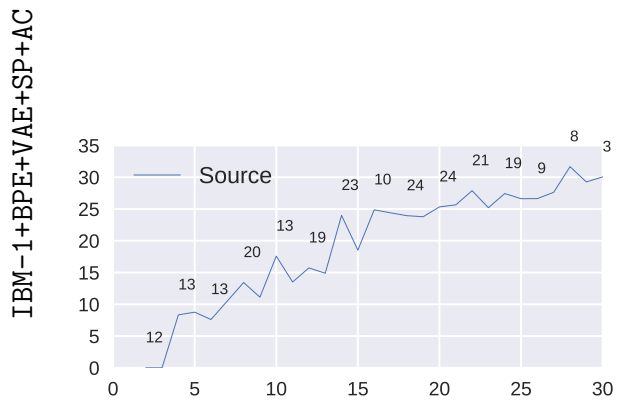
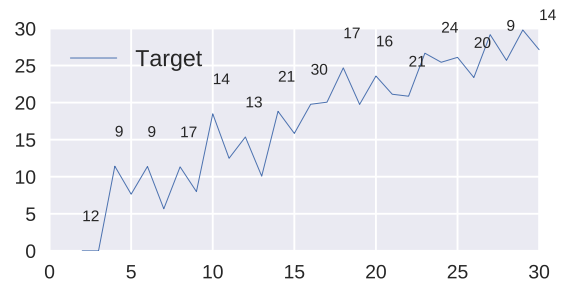
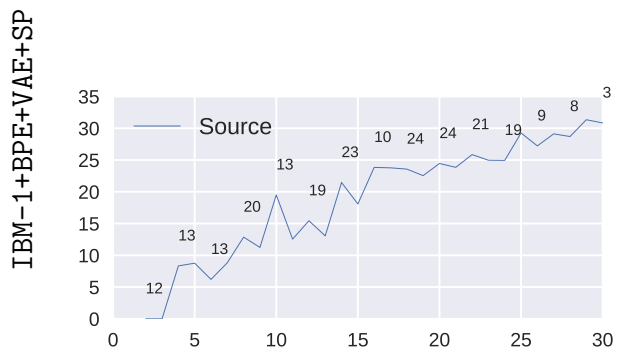
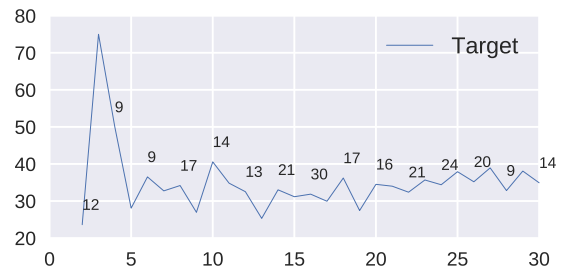
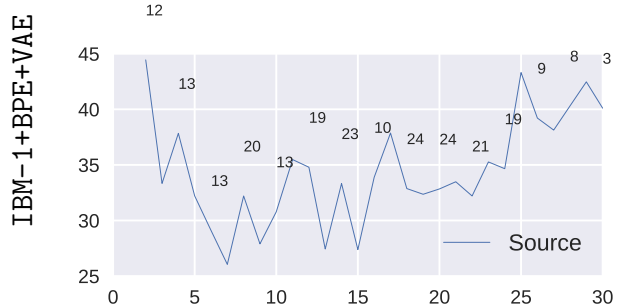
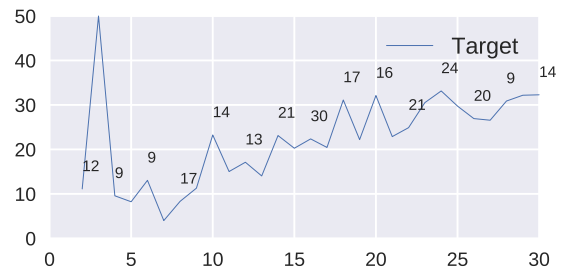
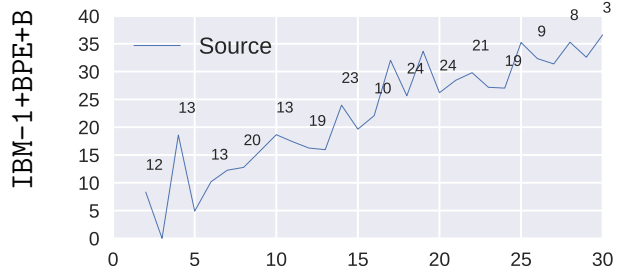
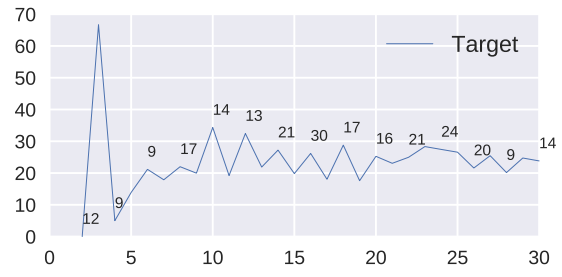
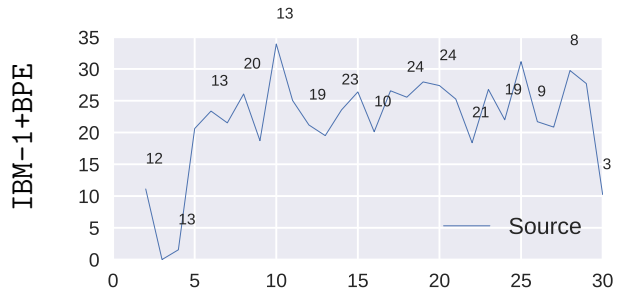
Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	R-ACC	AER	F1	PRE	REC	R-ACC
	IBM-1+BPE+VAE									
+SP	55.54	44.43	47.58	41.67	78.53	55.1	44.9	47.39	42.66	79.74
+Mono+Noise	54.74	45.26	48.47	42.44	77.68	54.06	45.92	47.99	44.02	80.34
	HMM+BPE+VAE+SP									
+SP	43.67	56.32	65.78	49.24	95.59	48.99	51.01	52.15	49.91	95.09
+Mono+Noise	42.87	57.13	71.18	47.71	95.77	46.94	53.06	55.52	50.81	95.19

Table D.44: Training with a monolingual corpus and the noise model (+Noise) on English-Japanese corpus. R-Acc is the accuracy of the reconstruction model.

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	R-ACC	AER	F1	PRE	REC	R-ACC
	IBM-1+BPE+VAE									
+SP	51.34	48.61	54.75	43.71	82.63	41.38	58.63	58.37	58.88	64.35
+Mono+Noise	53.93	46.06	51.83	41.44	83.44	40.2	59.8	63.27	56.69	64.28
	HMM+BPE+VAE+SP									
+SP	45.87	54.13	64.85	46.45	95.09	33.78	66.22	69.36	63.35	88.13
+Mono+Noise	45.23	54.77	67.73	45.98	94.89	33.26	66.74	75.25	59.96	88.08

Table D.45: Training with a monolingual corpus and the noise model (+Noise) on English-Vietnamese corpus. R-Acc is the accuracy of the reconstruction model.

D.11 AER as a function of sentence length



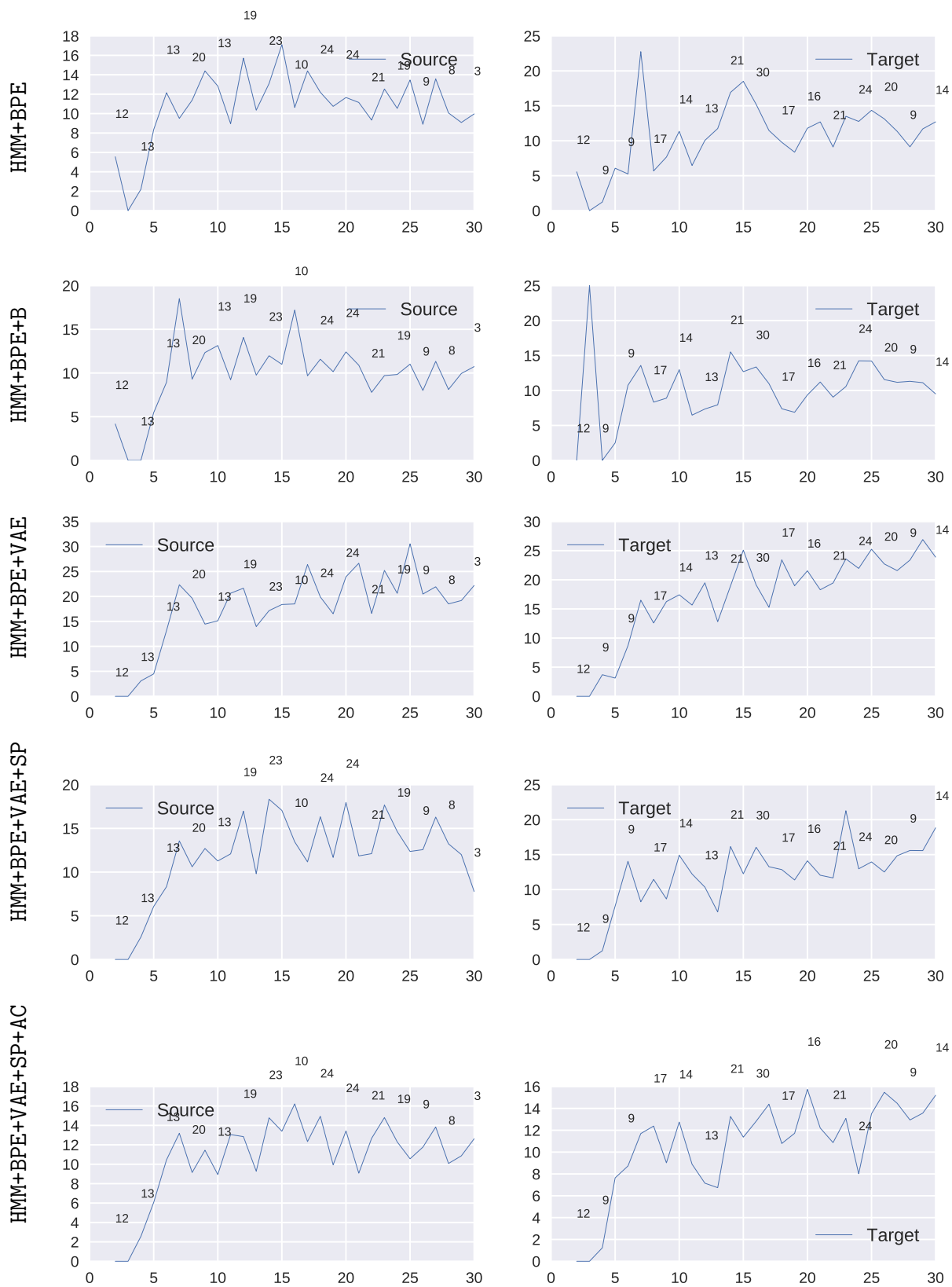
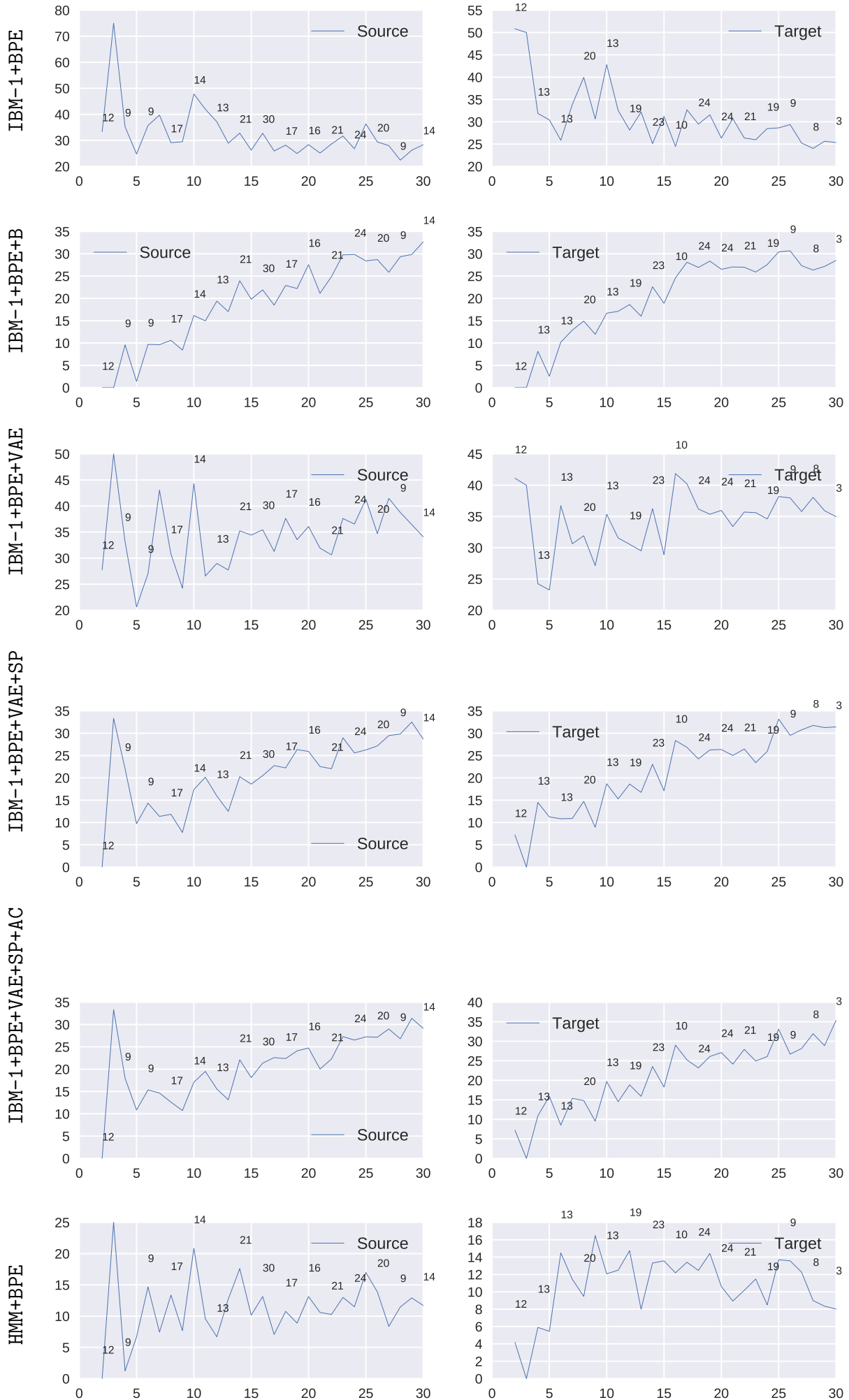


Figure D.55: Variational models for the direction English-French: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



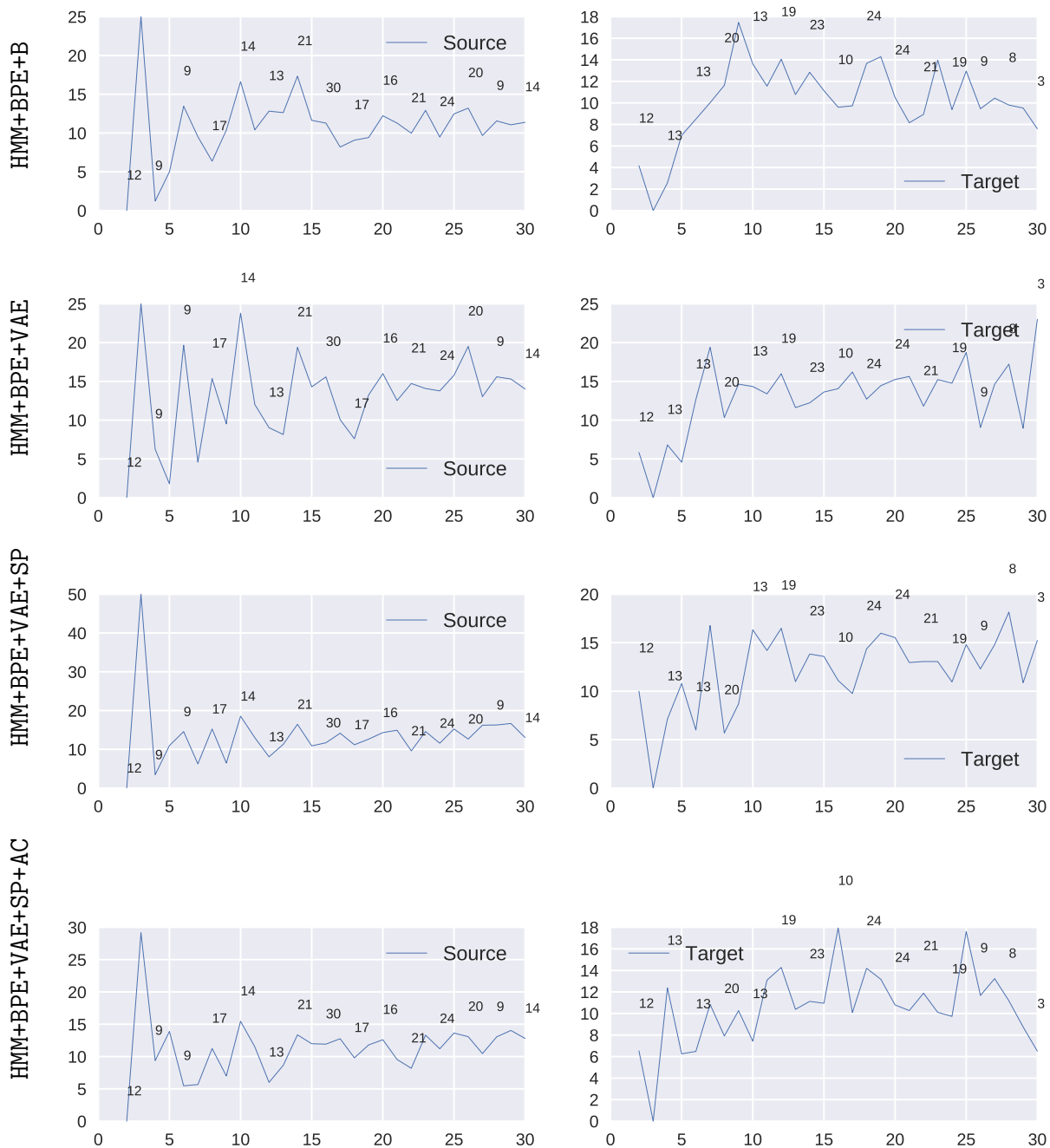
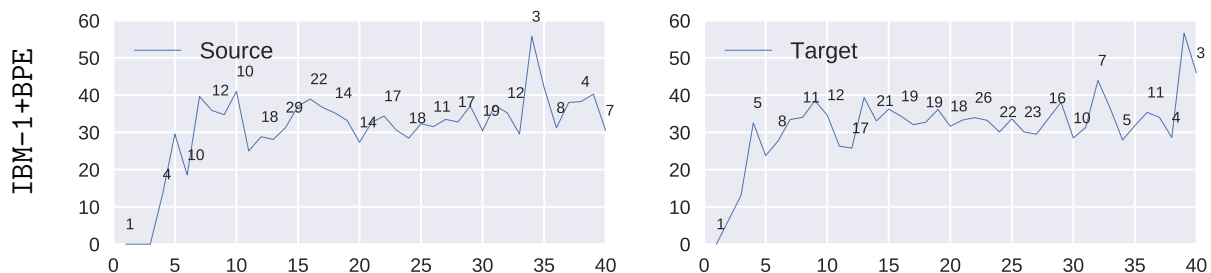
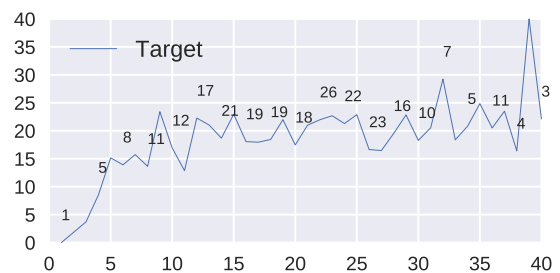
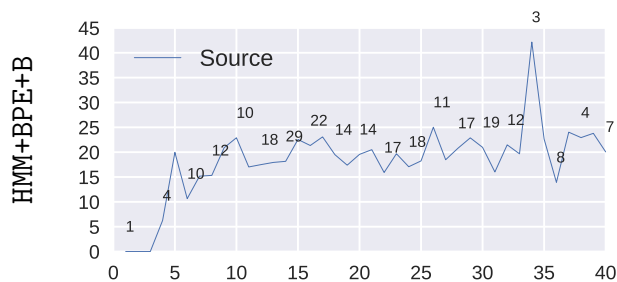
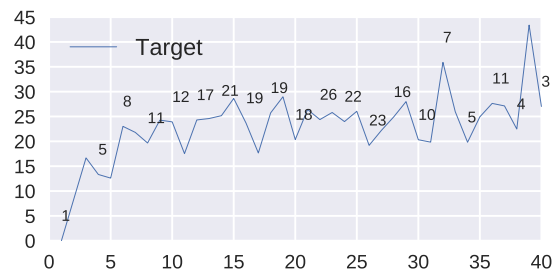
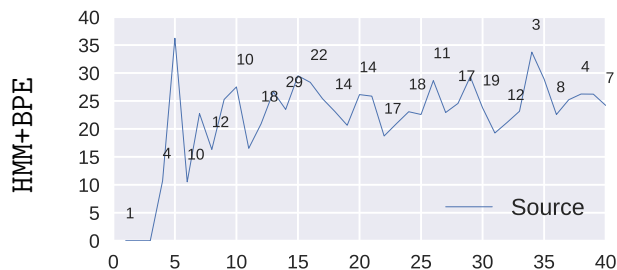
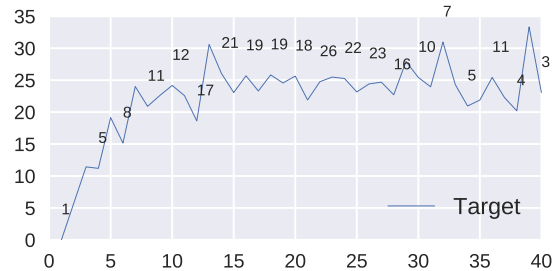
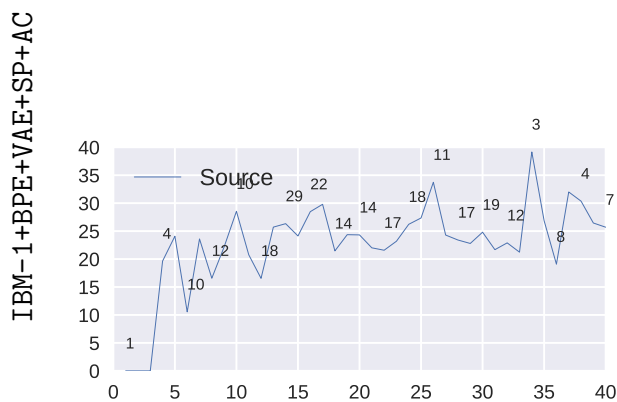
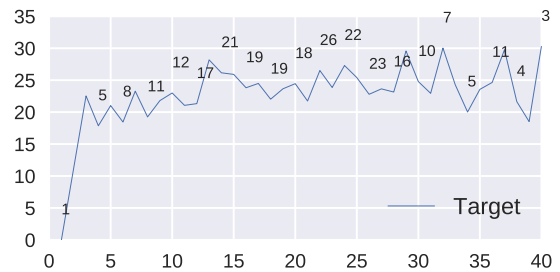
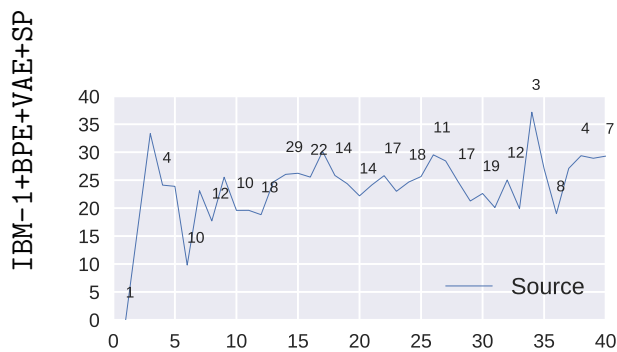
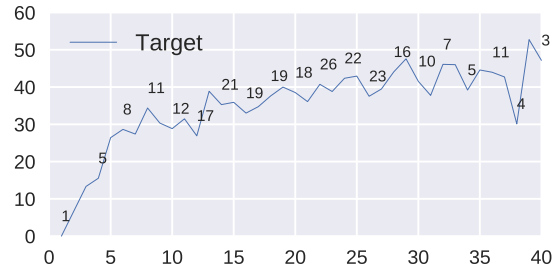
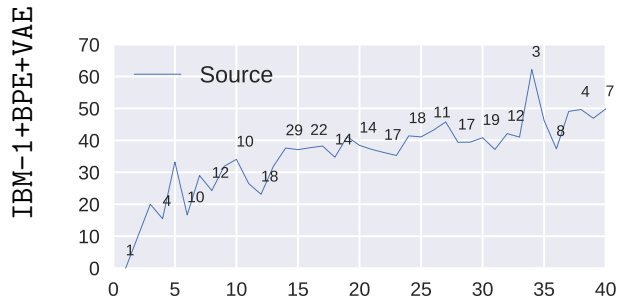
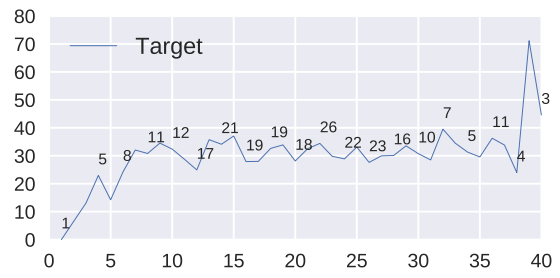
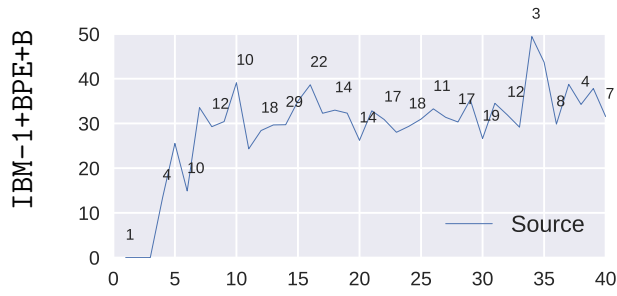


Figure D.56: Variational models for the direction French-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.





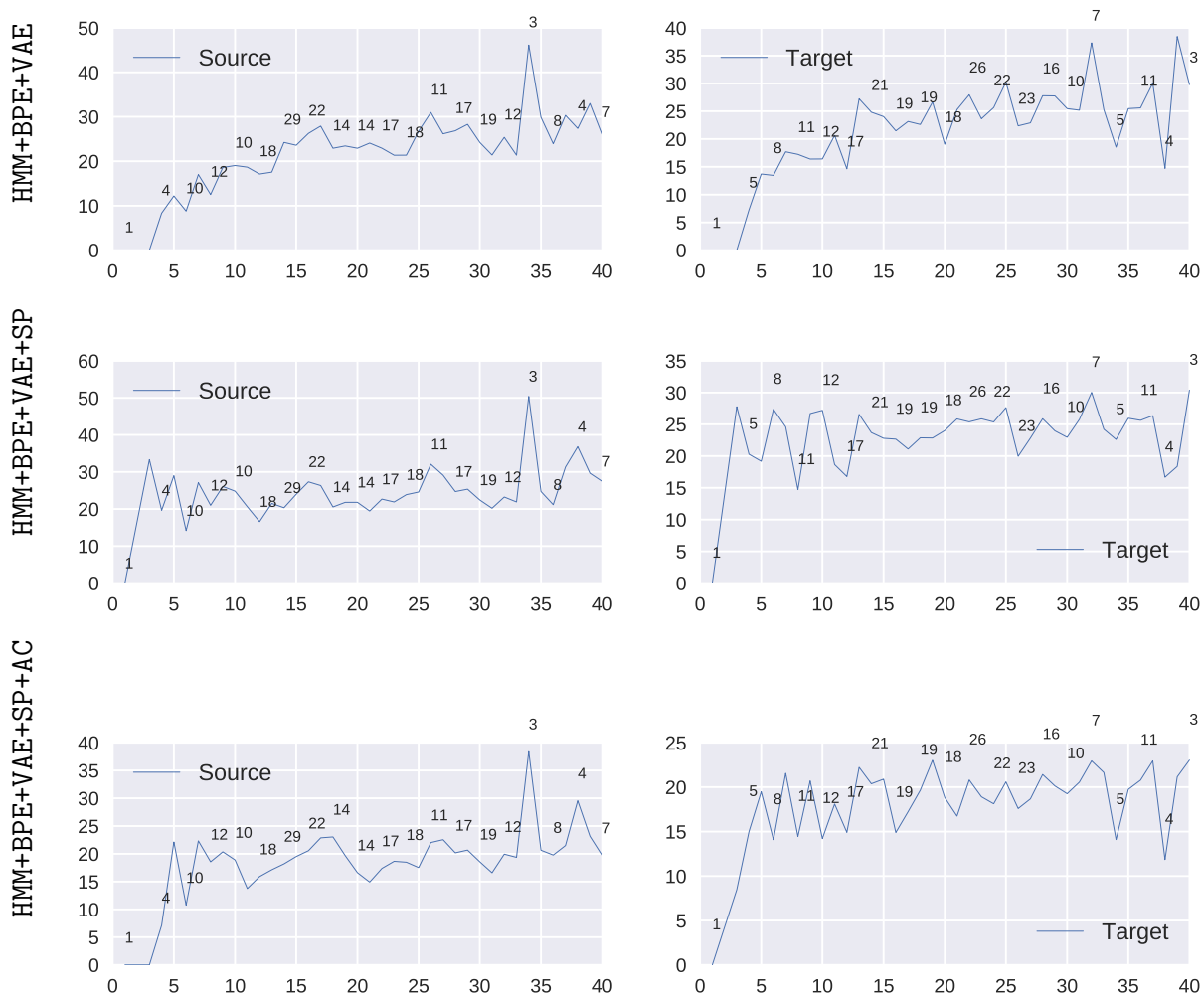
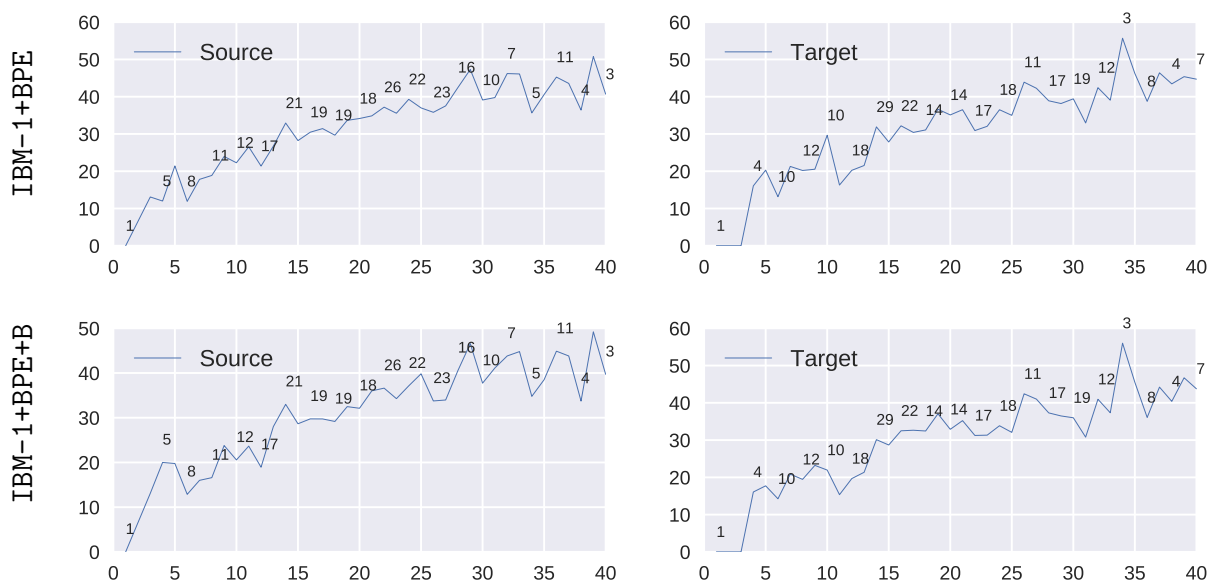
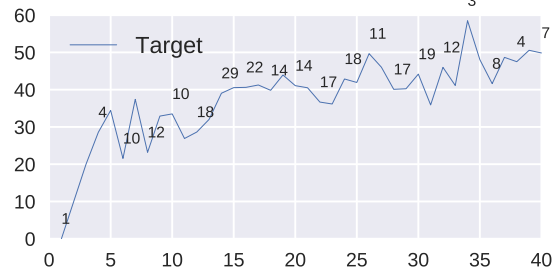
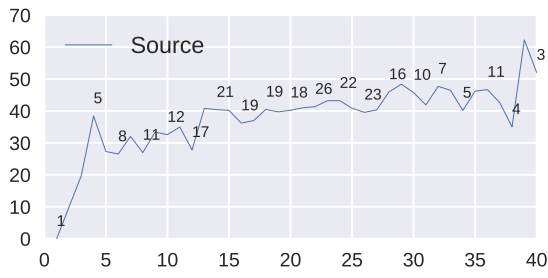


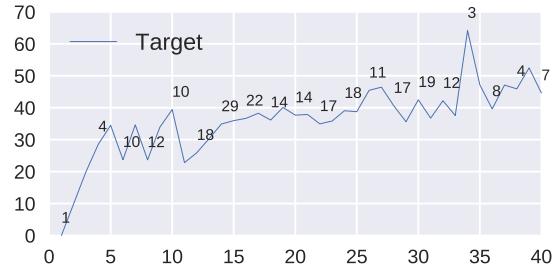
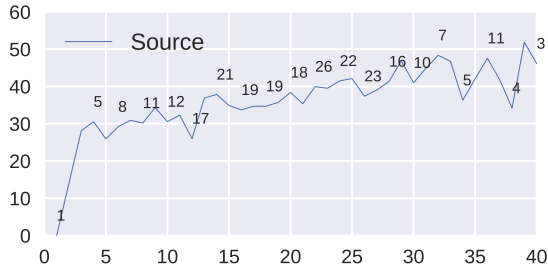
Figure D.57: Variational models for the direction English-German: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



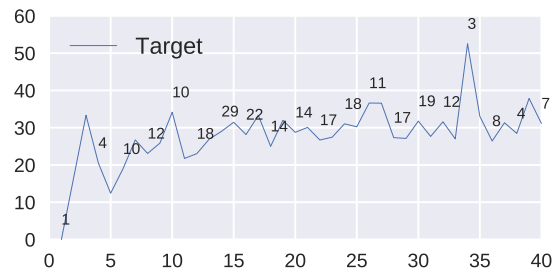
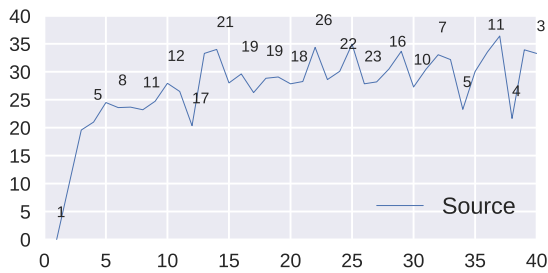
IBM-1+BPE+VAE



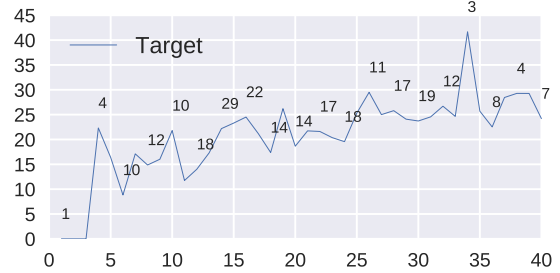
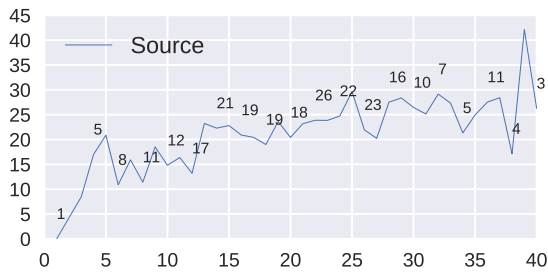
IBM-1+BPE+VAE+SP



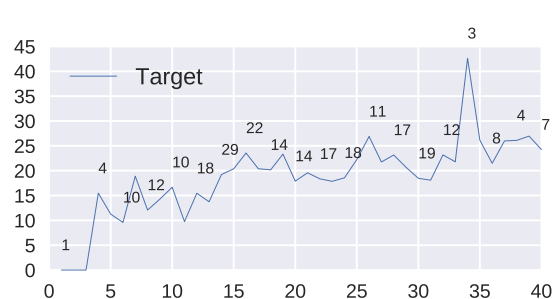
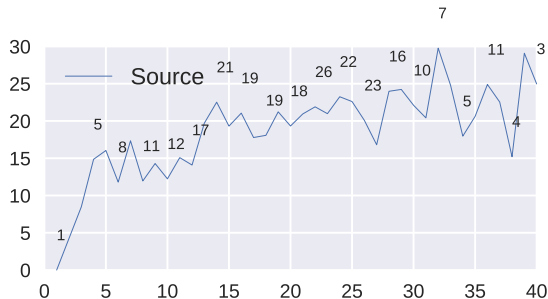
IBM-1+BPE+VAE+SP+AC



HMM+BPE



HMM+BPE+B



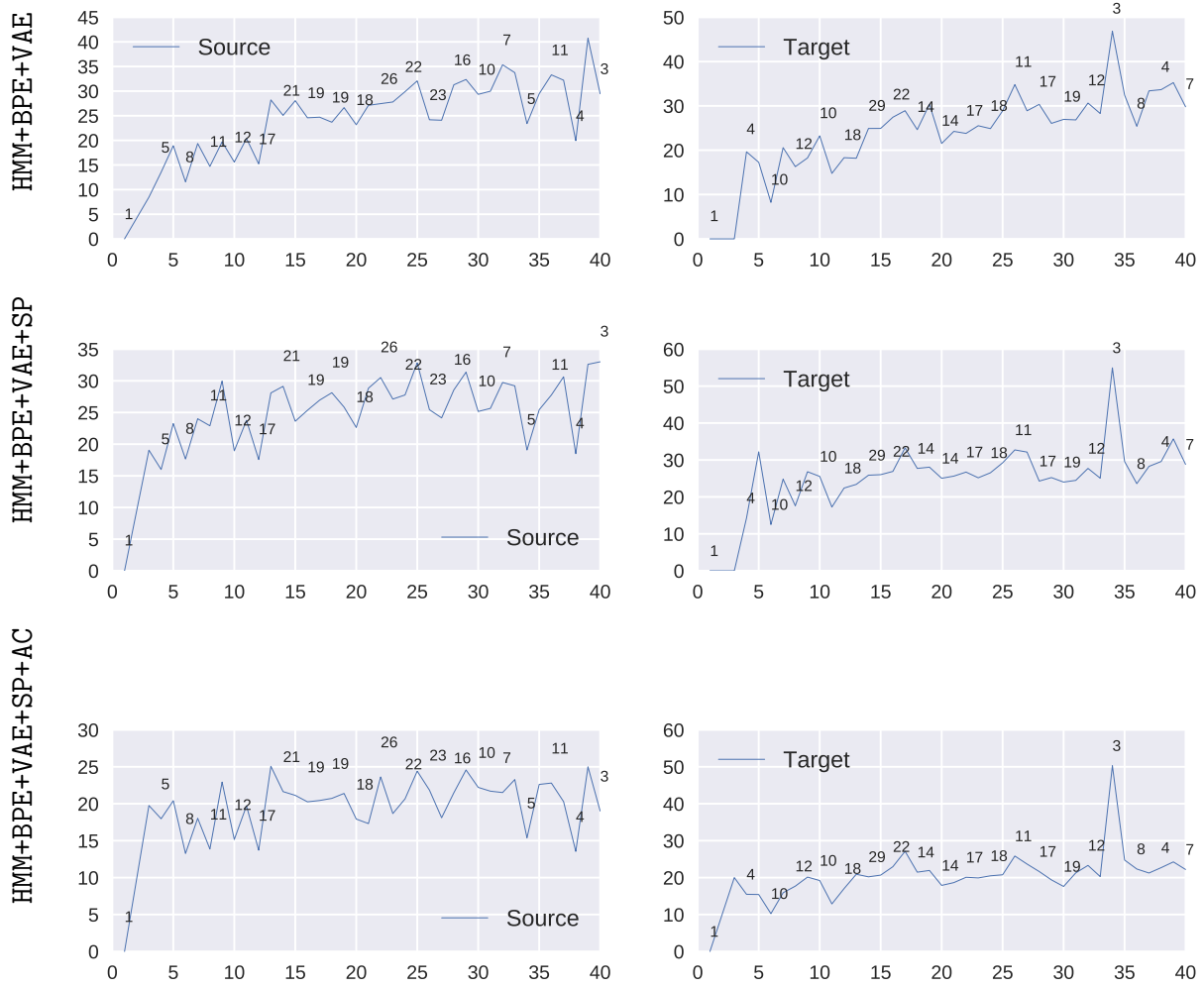
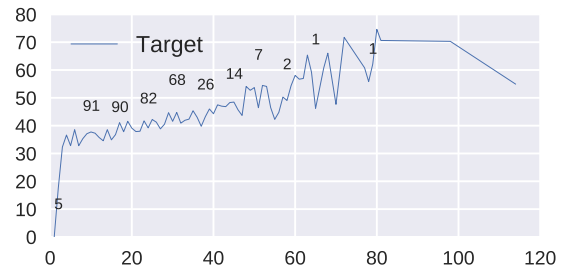
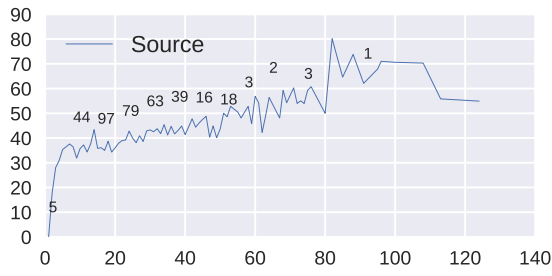


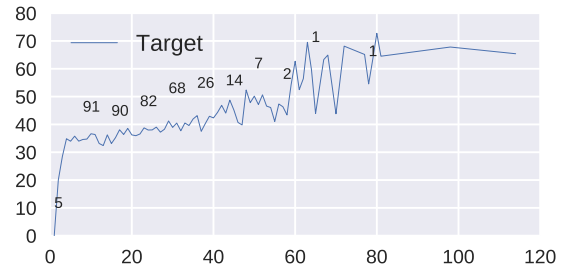
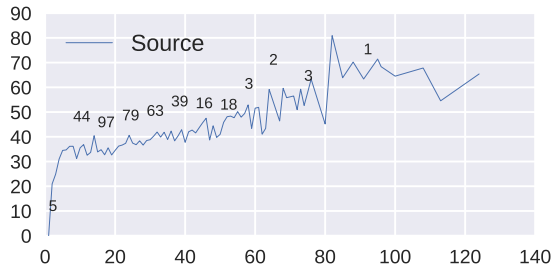
Figure D.58: Variational models for the direction German-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



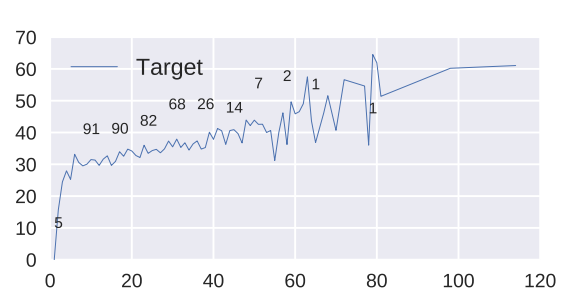
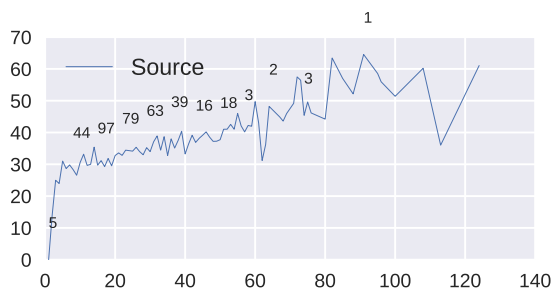
IBM-1+BPE+VAE



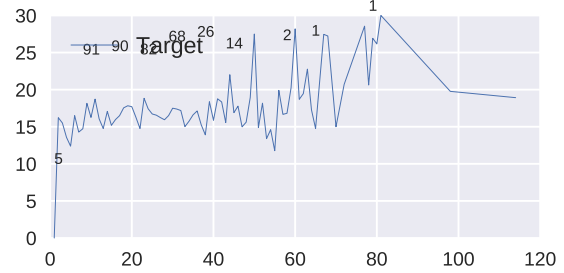
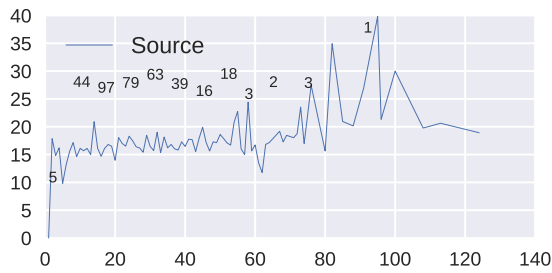
IBM-1+BPE+VAE+SP



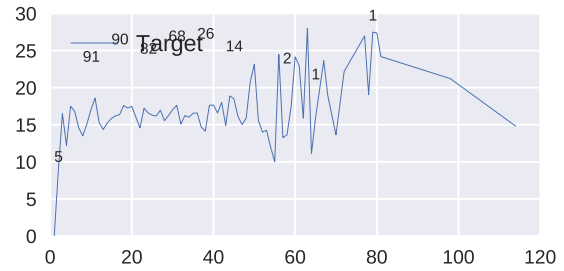
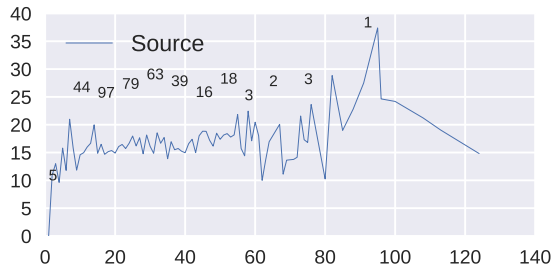
IBM-1+BPE+VAE+SP+AC



HMM+BPE



HMM+BPE+B



7

7

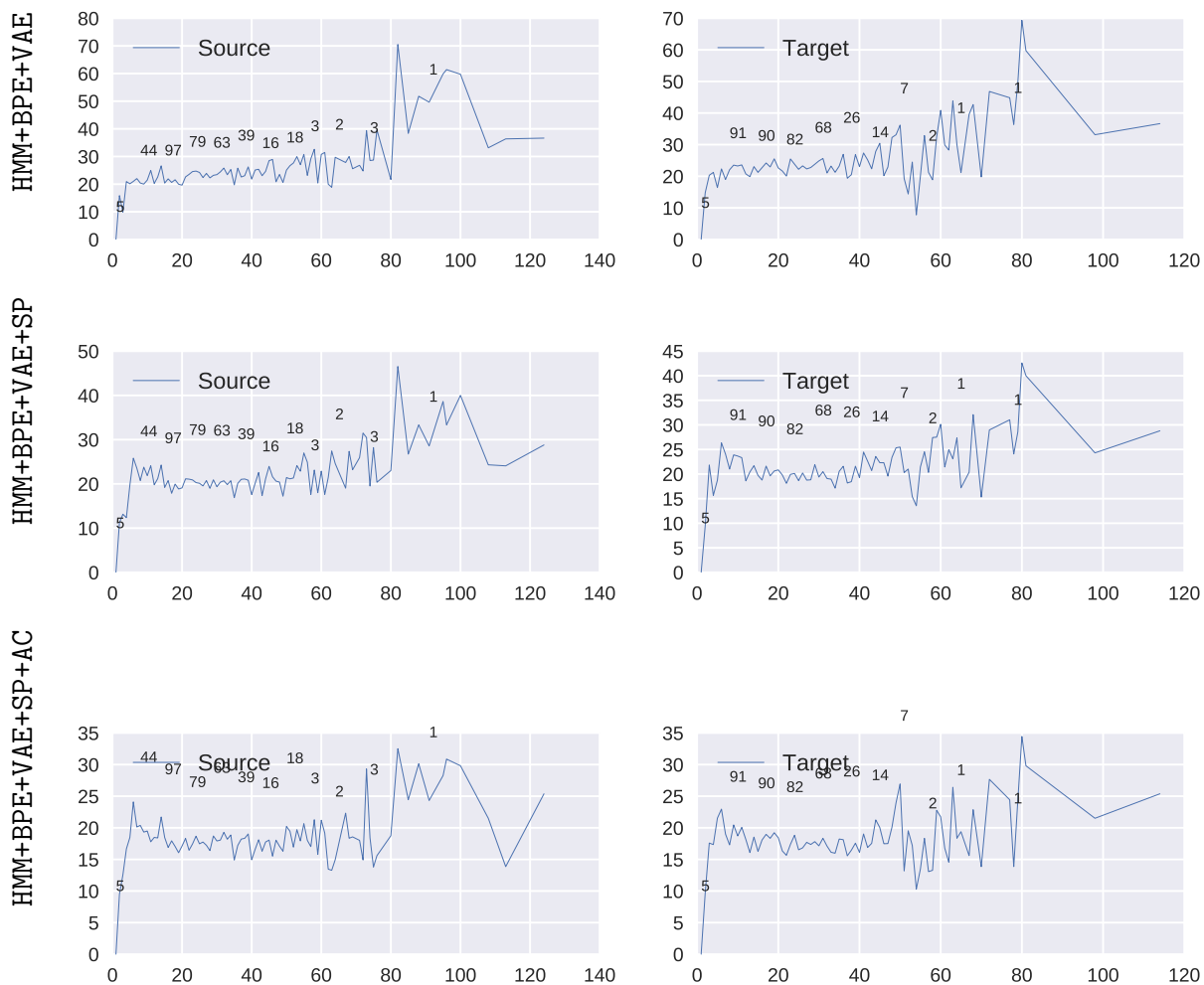
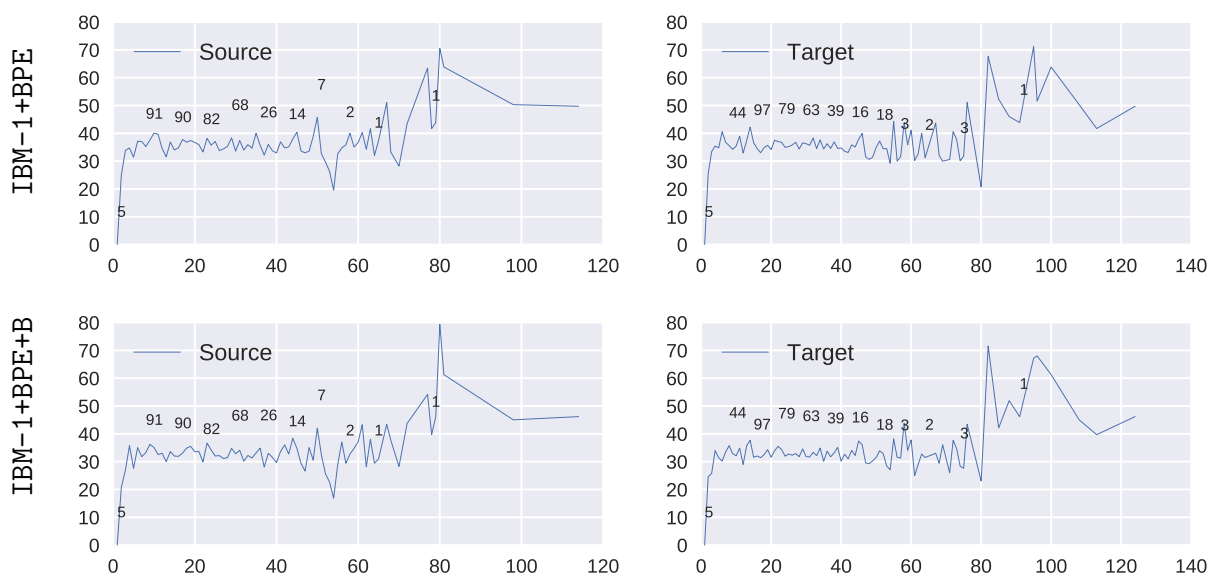
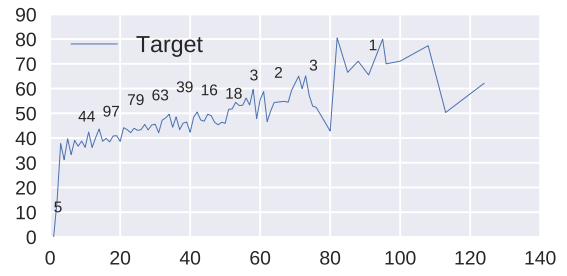
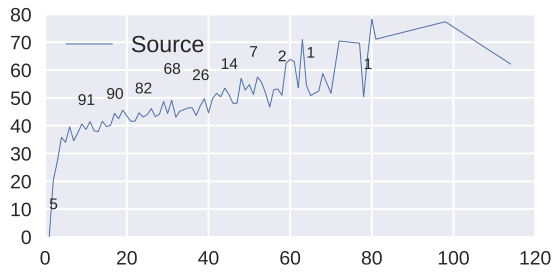


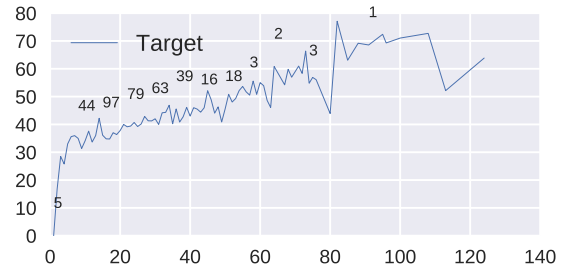
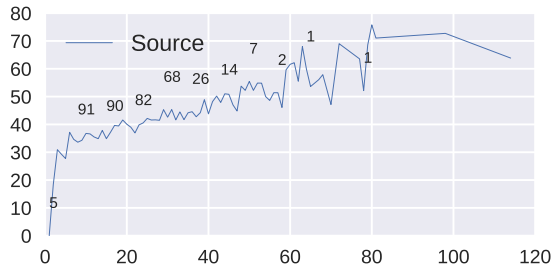
Figure D.59: Variational models for the direction English-Czech: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



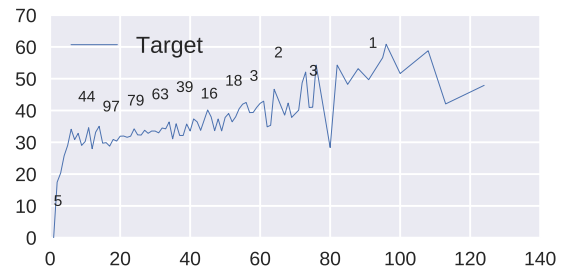
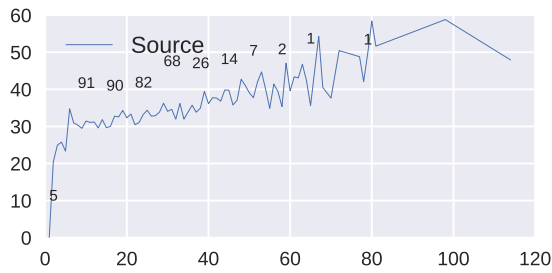
IBM-1+BPE+VAE



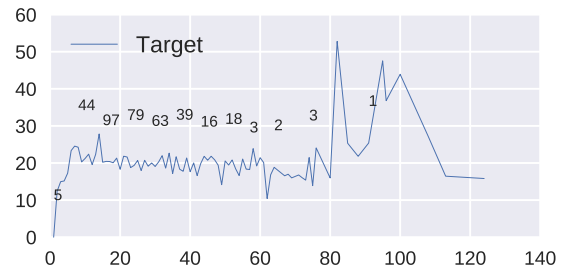
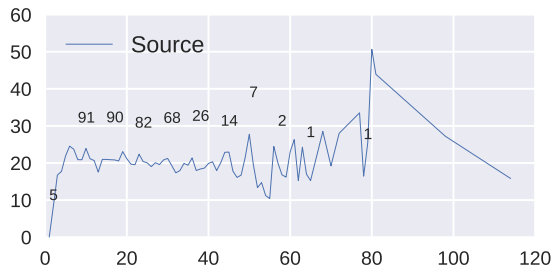
IBM-1+BPE+VAE+SP



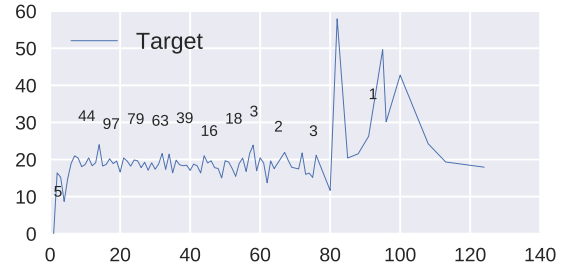
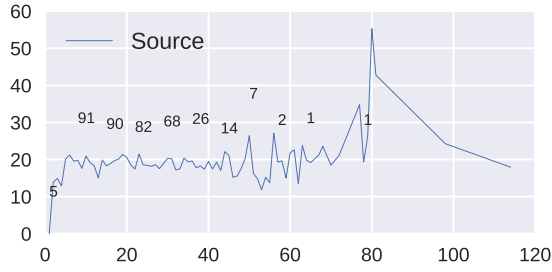
IBM-1+BPE+VAE+SP+AC



HMM+BPE



HMM+BPE+B



HMM+BPE+VAE

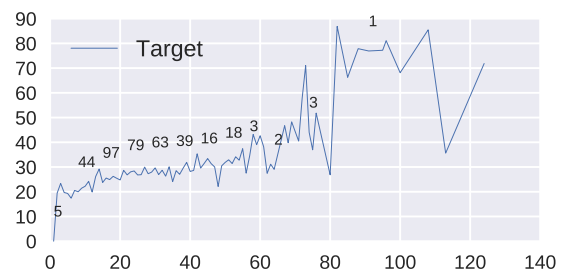
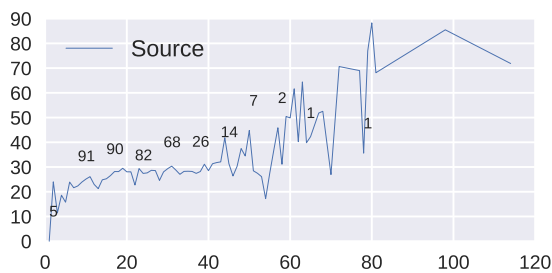
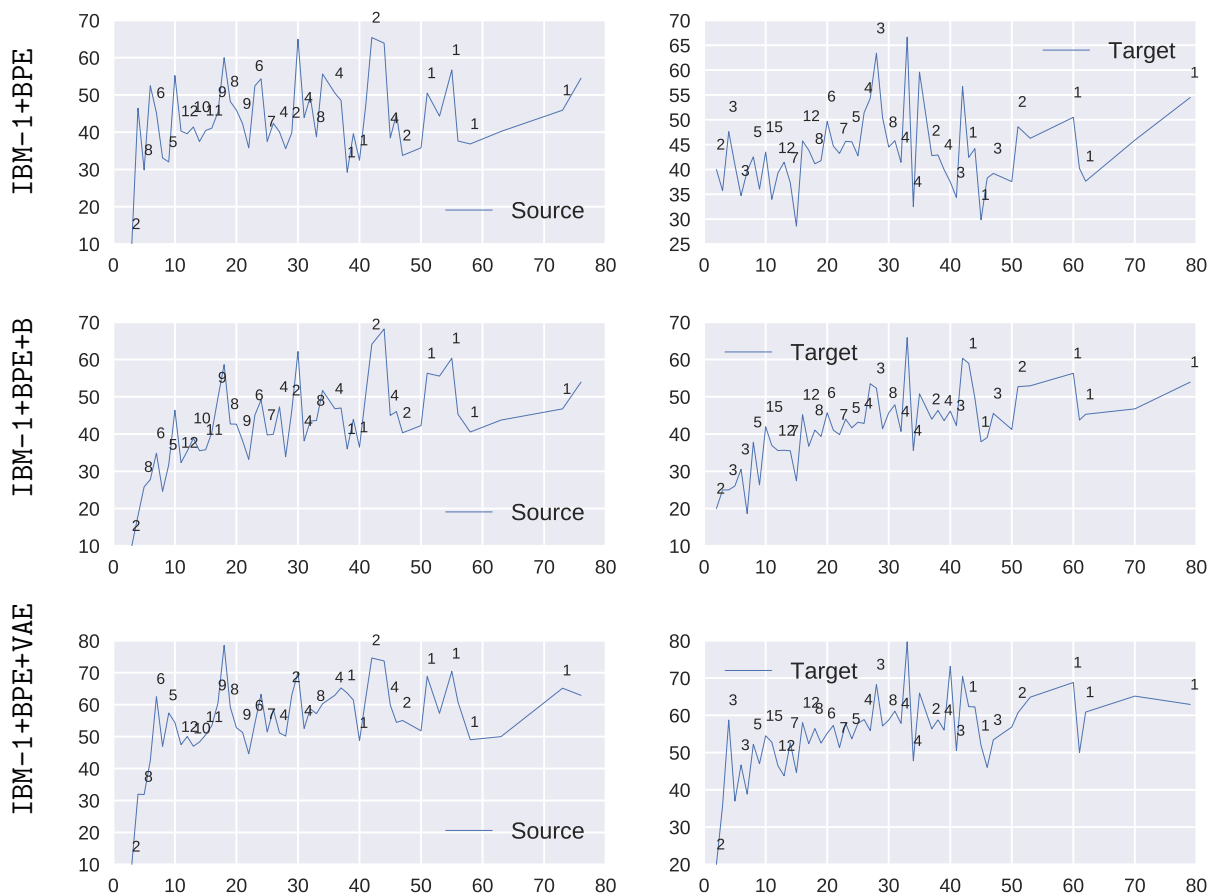
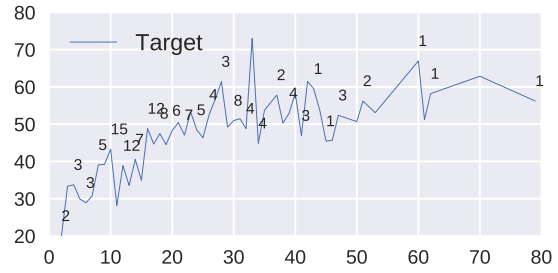
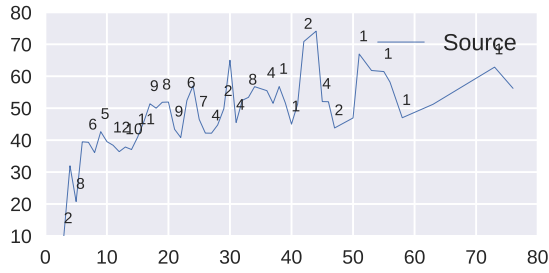




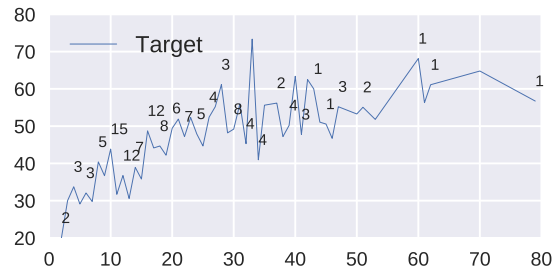
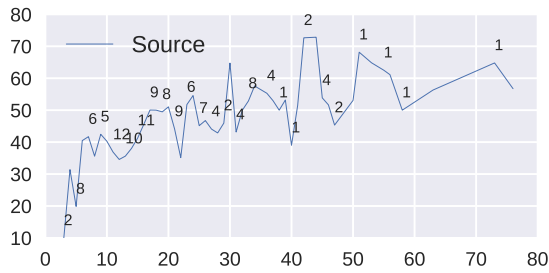
Figure D.60: Variational models for the direction Czech-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



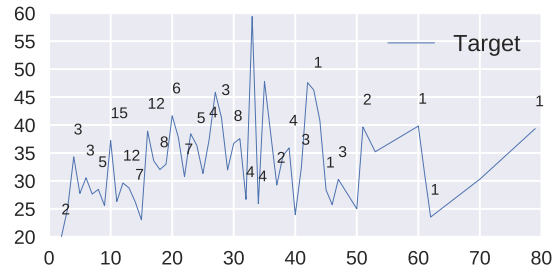
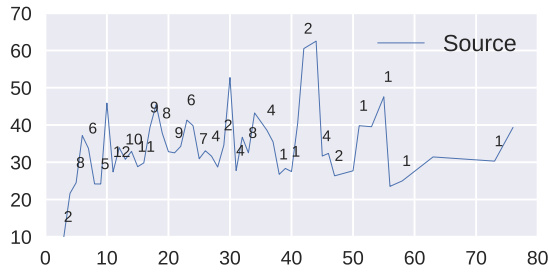
IBM-1+BPE+VAE+SP



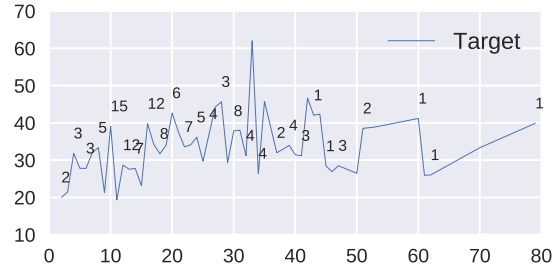
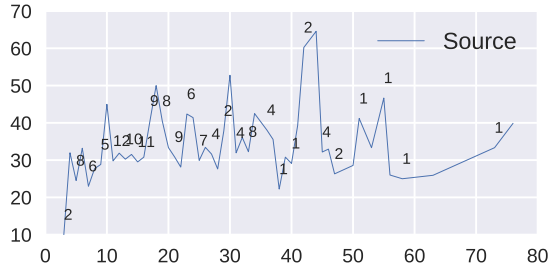
IBM-1+BPE+VAE+SP+AC



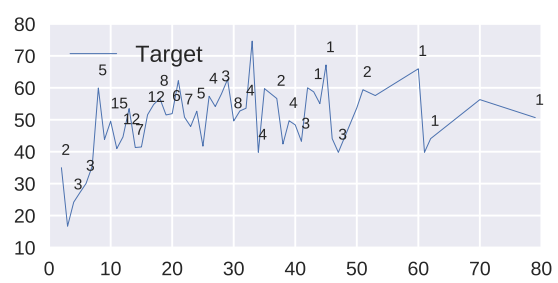
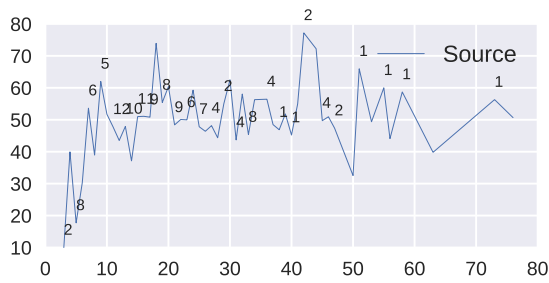
HMM+BPE



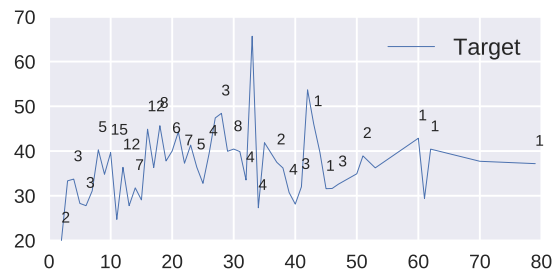
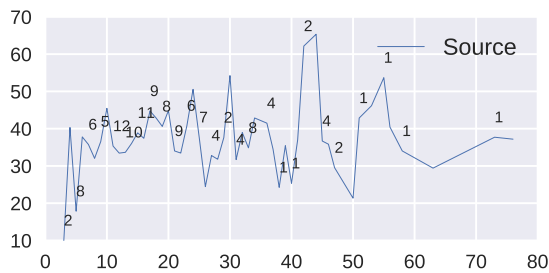
HMM+BPE+B



HMM+BPE+VAE



HMM+BPE+VAE+SP



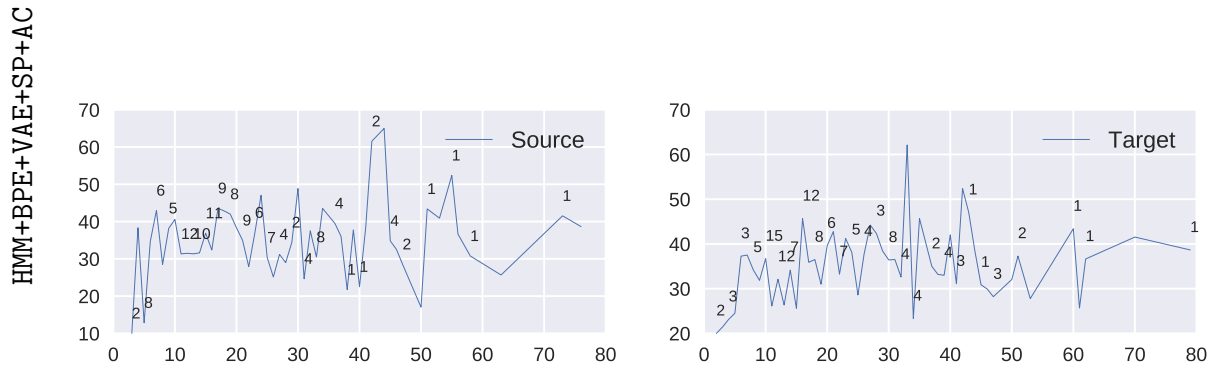
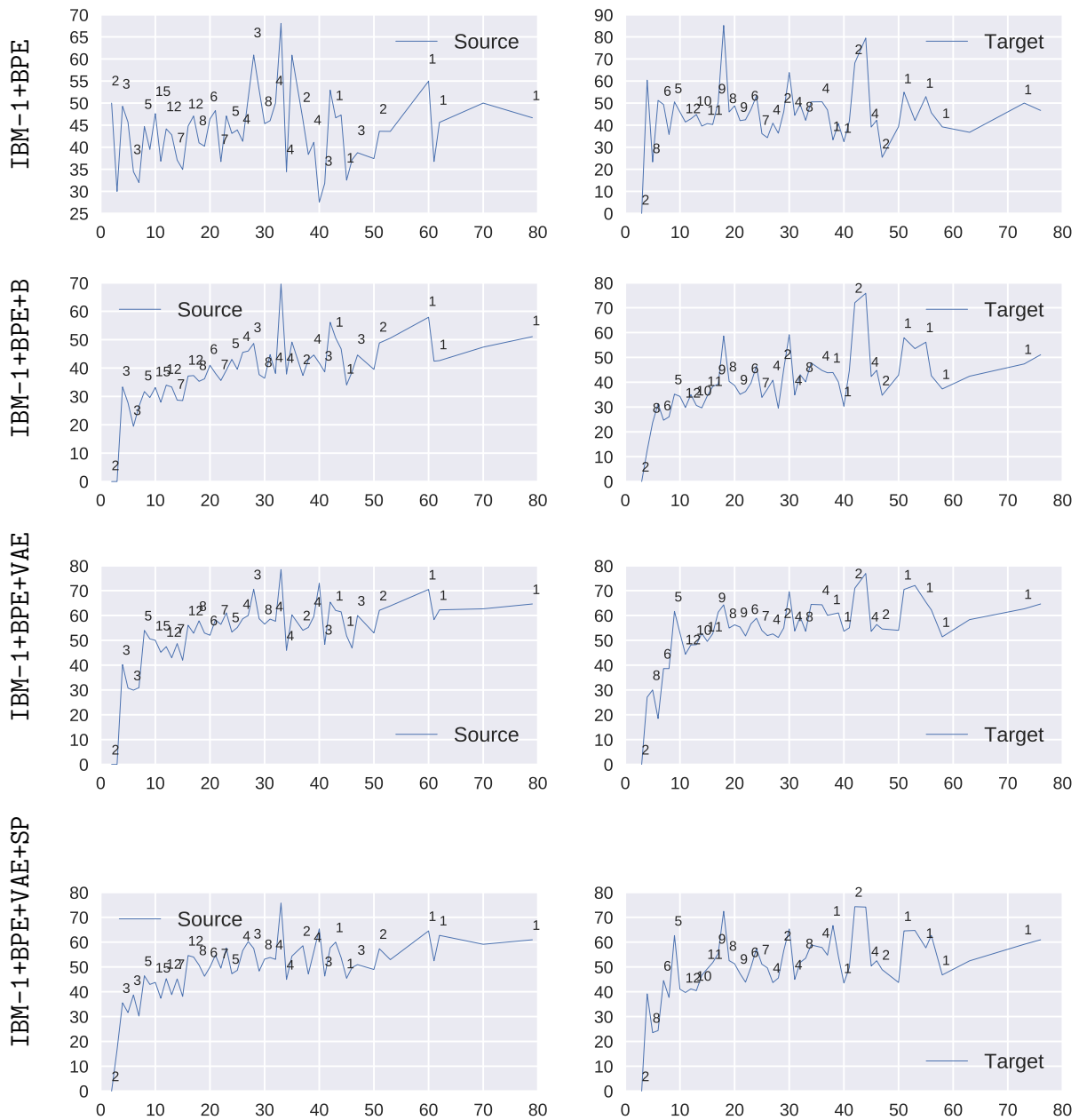
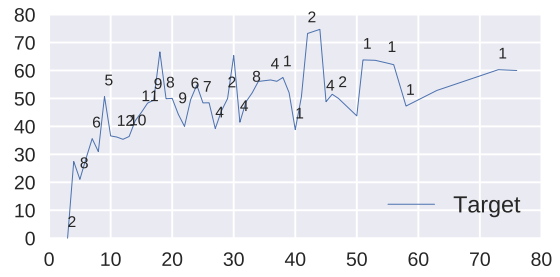
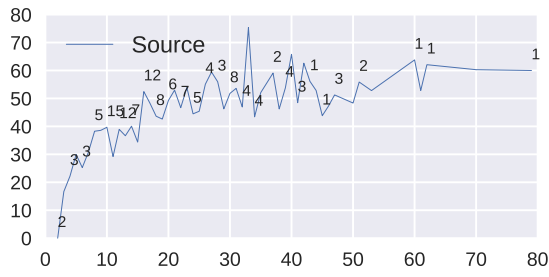


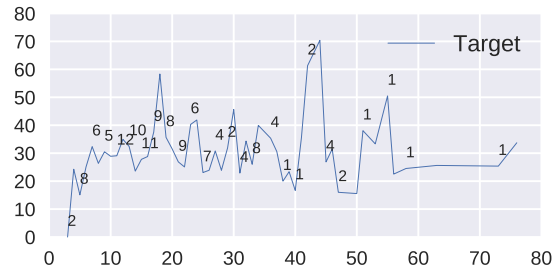
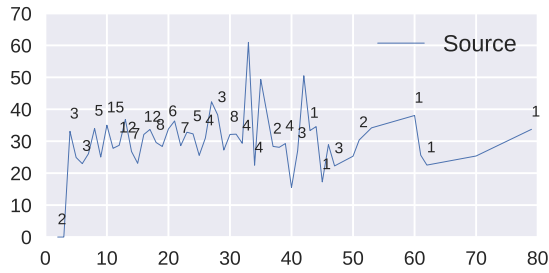
Figure D.61: Variational models for the direction English-Romanian: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



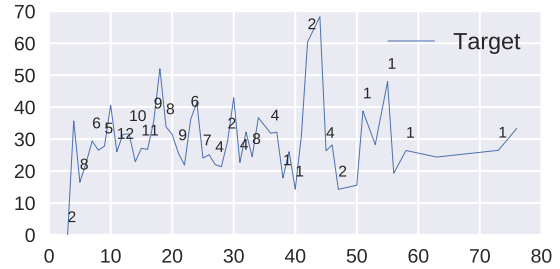
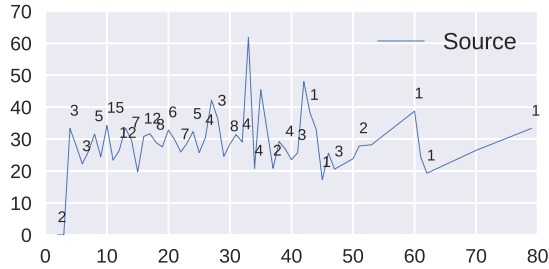
IBM-1+BPE+VAE+SP+AC



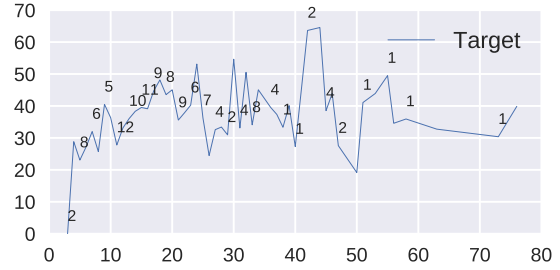
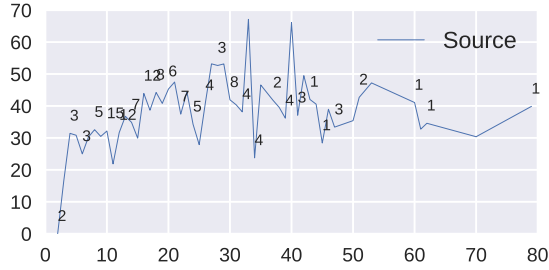
HMM+BPE



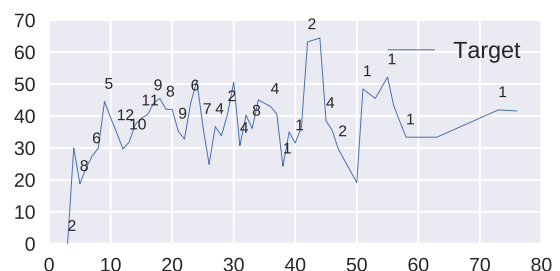
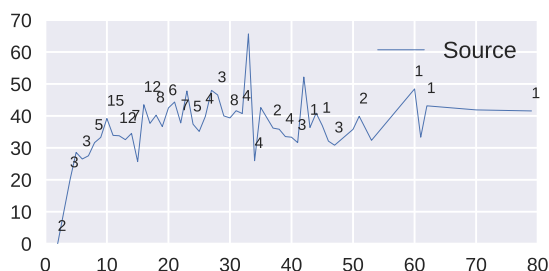
HMM+BPE+B



HMM+BPE+VAE



HMM+BPE+VAE+SP



HMM+BPE+VAE+SP+AC

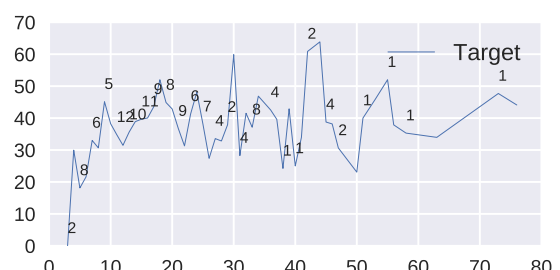
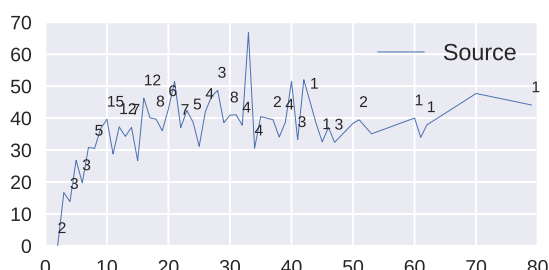
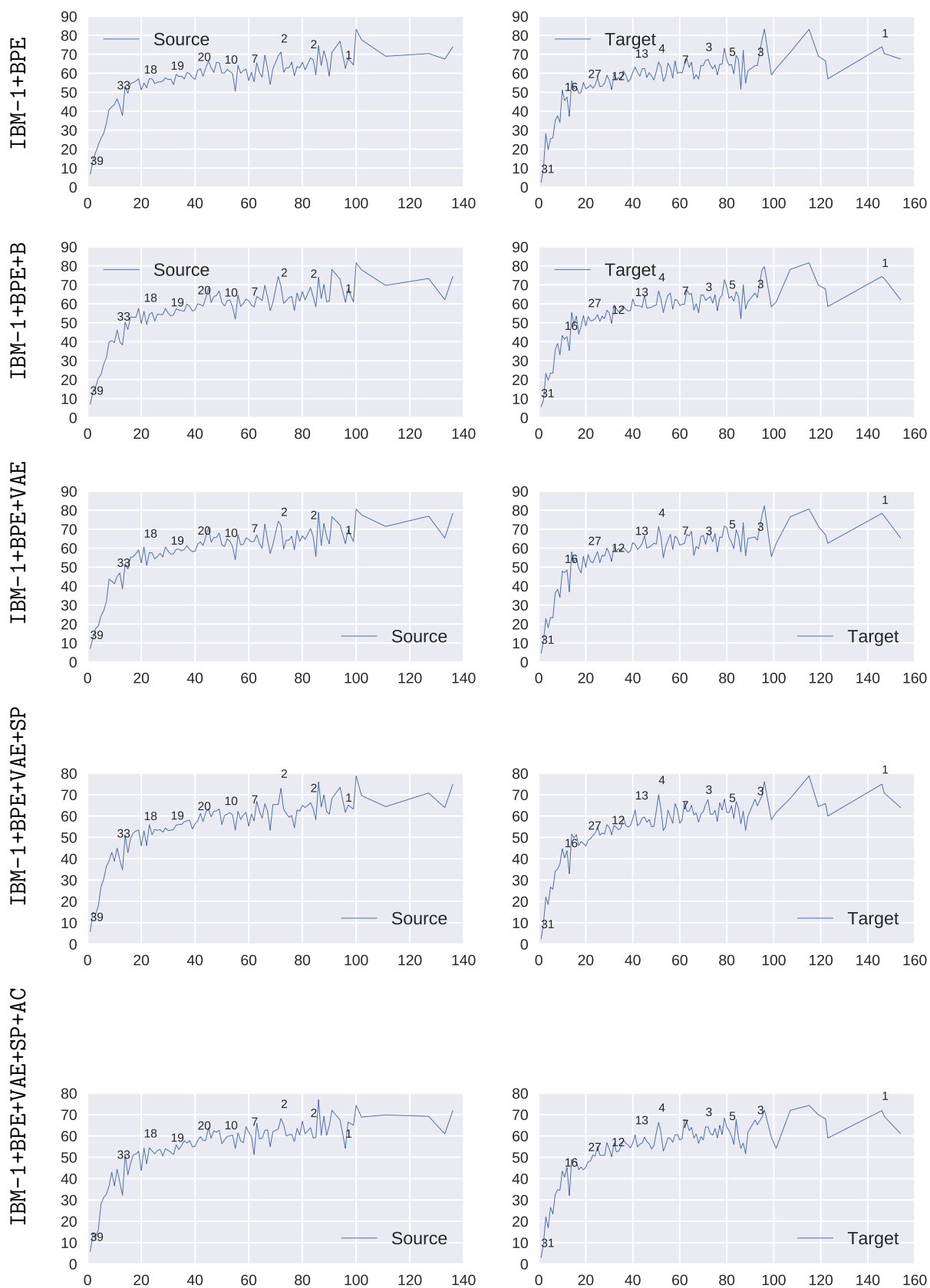


Figure D.62: Variational models for the direction Romanian-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



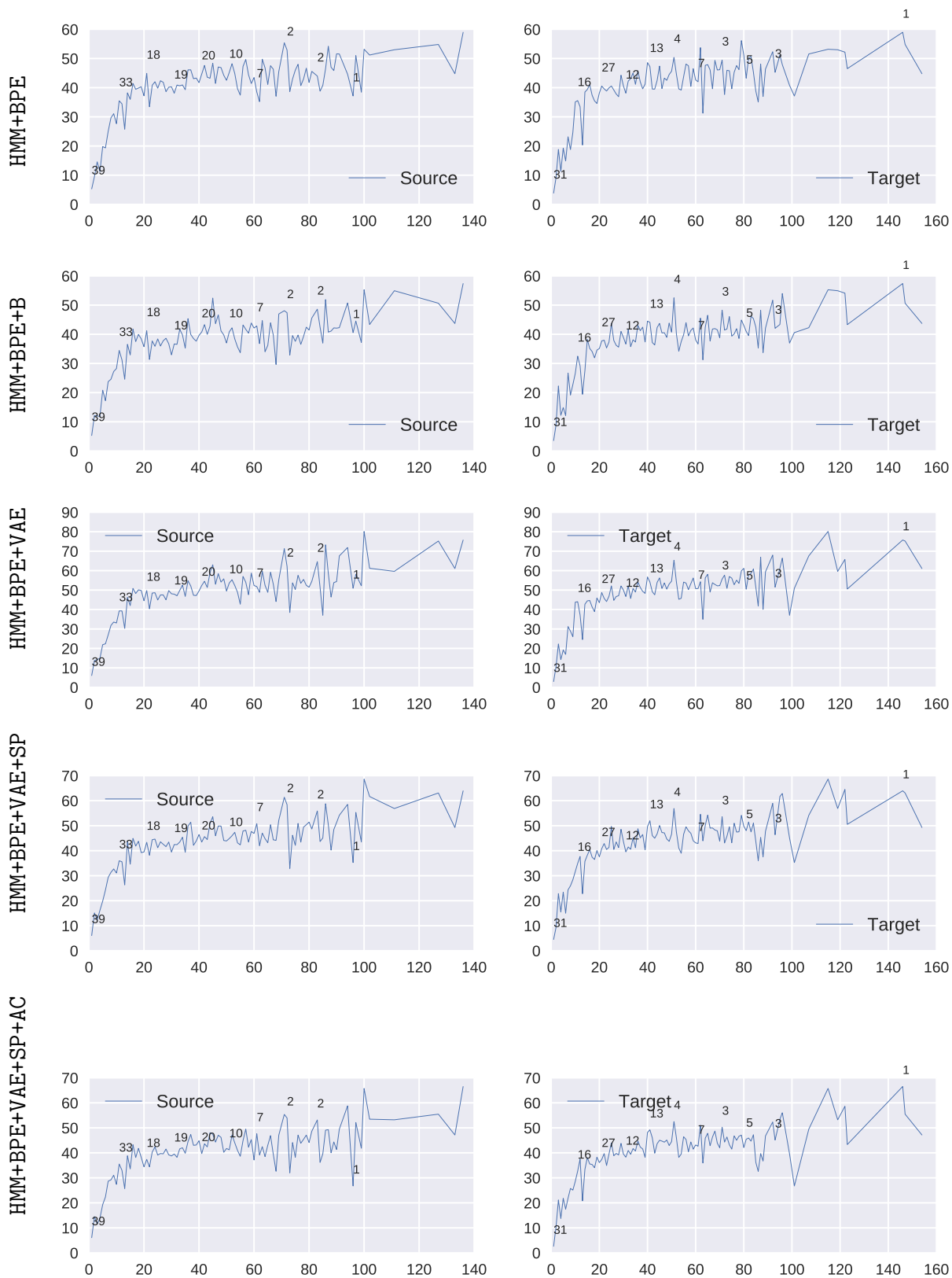
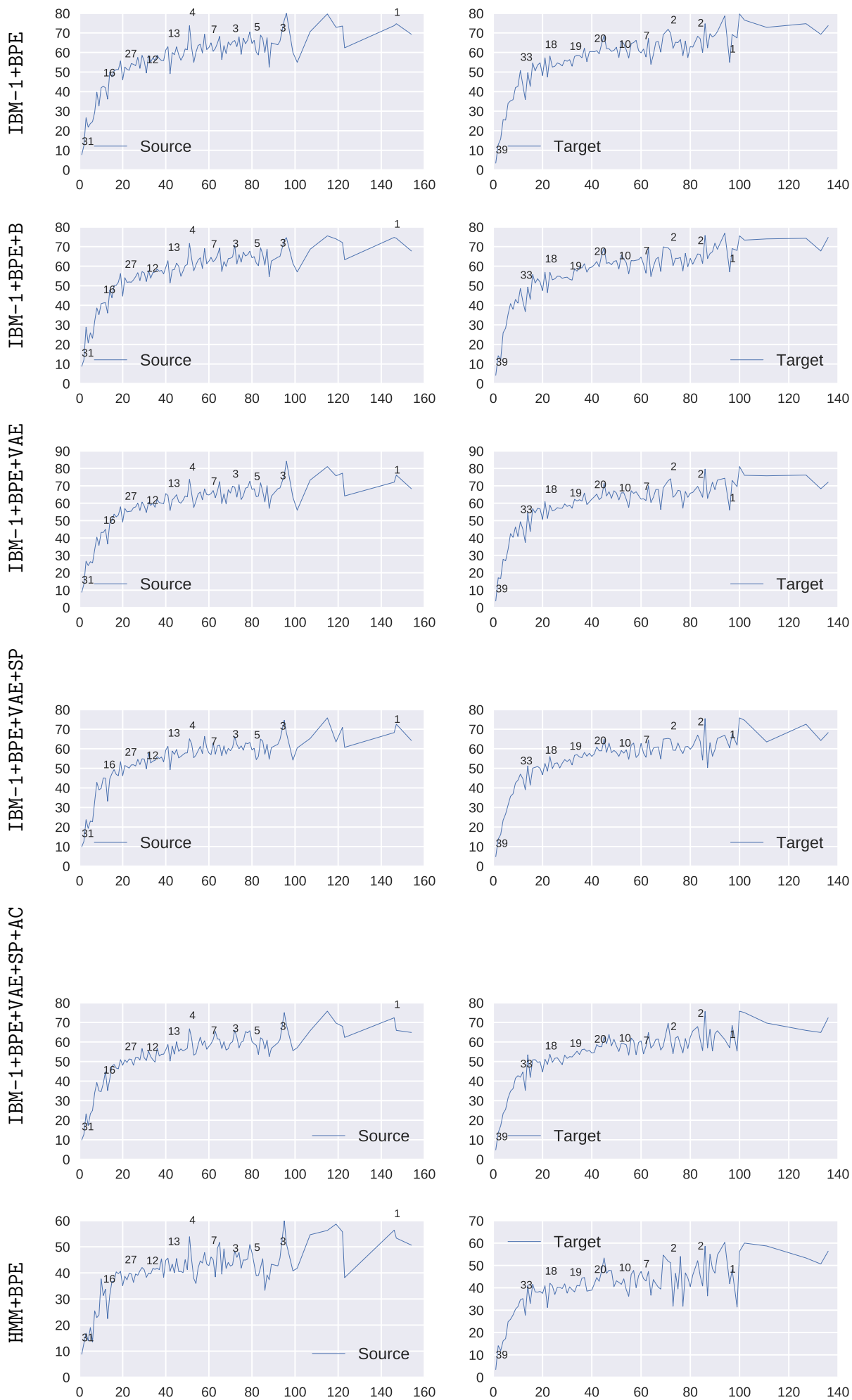


Figure D.63: Variational models for the direction English-Japanese: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



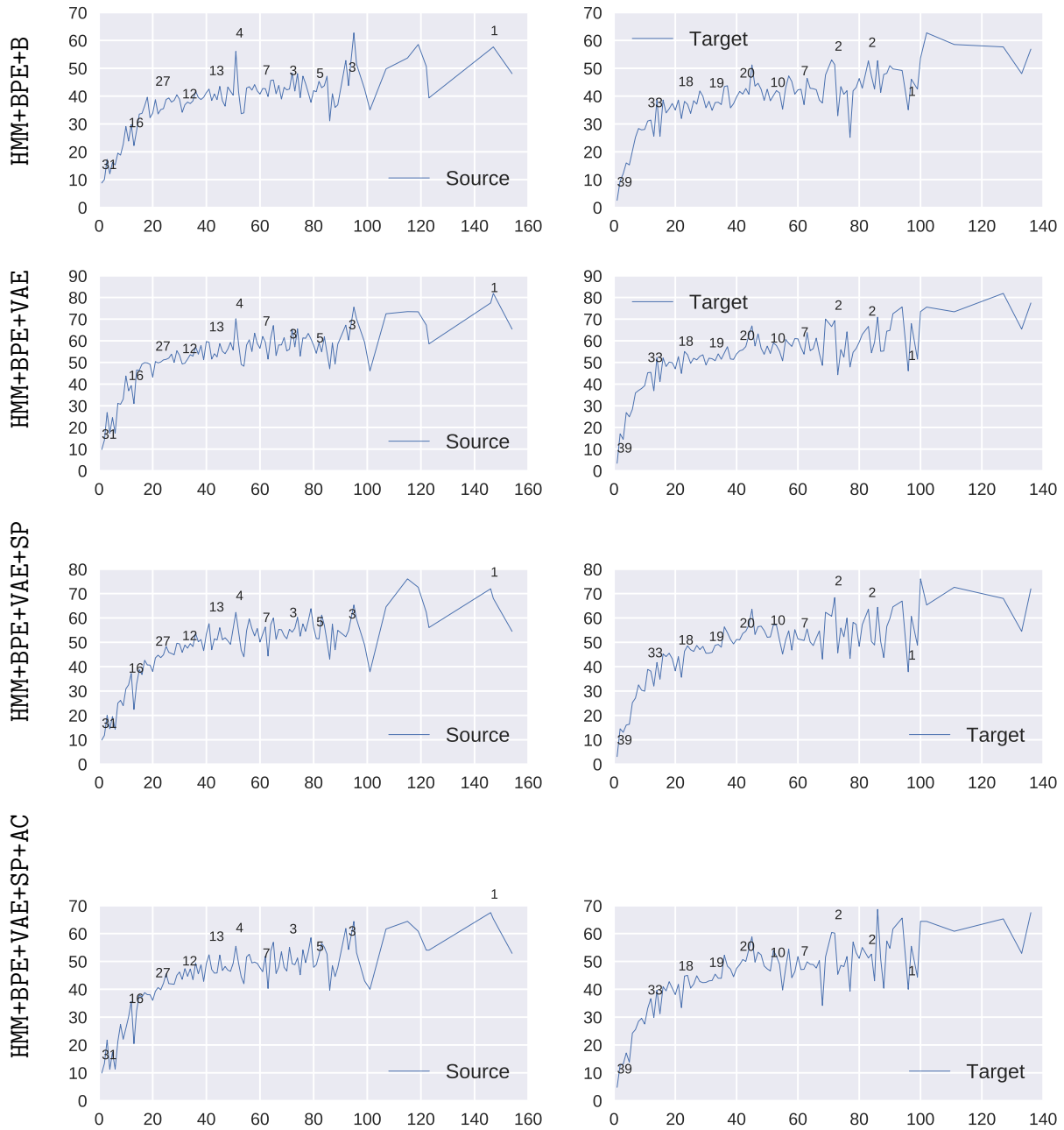
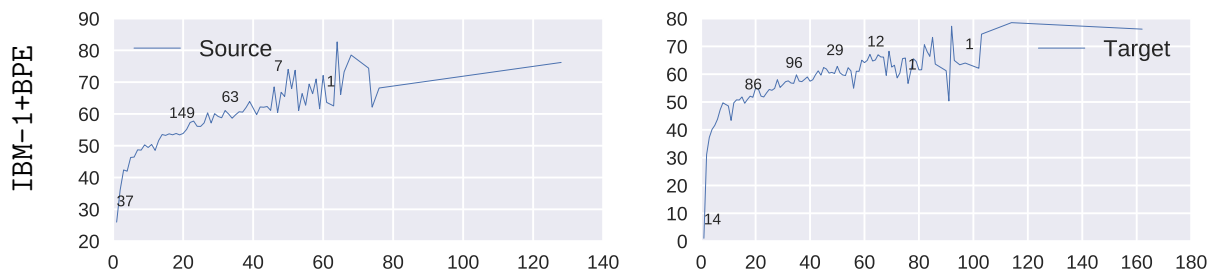
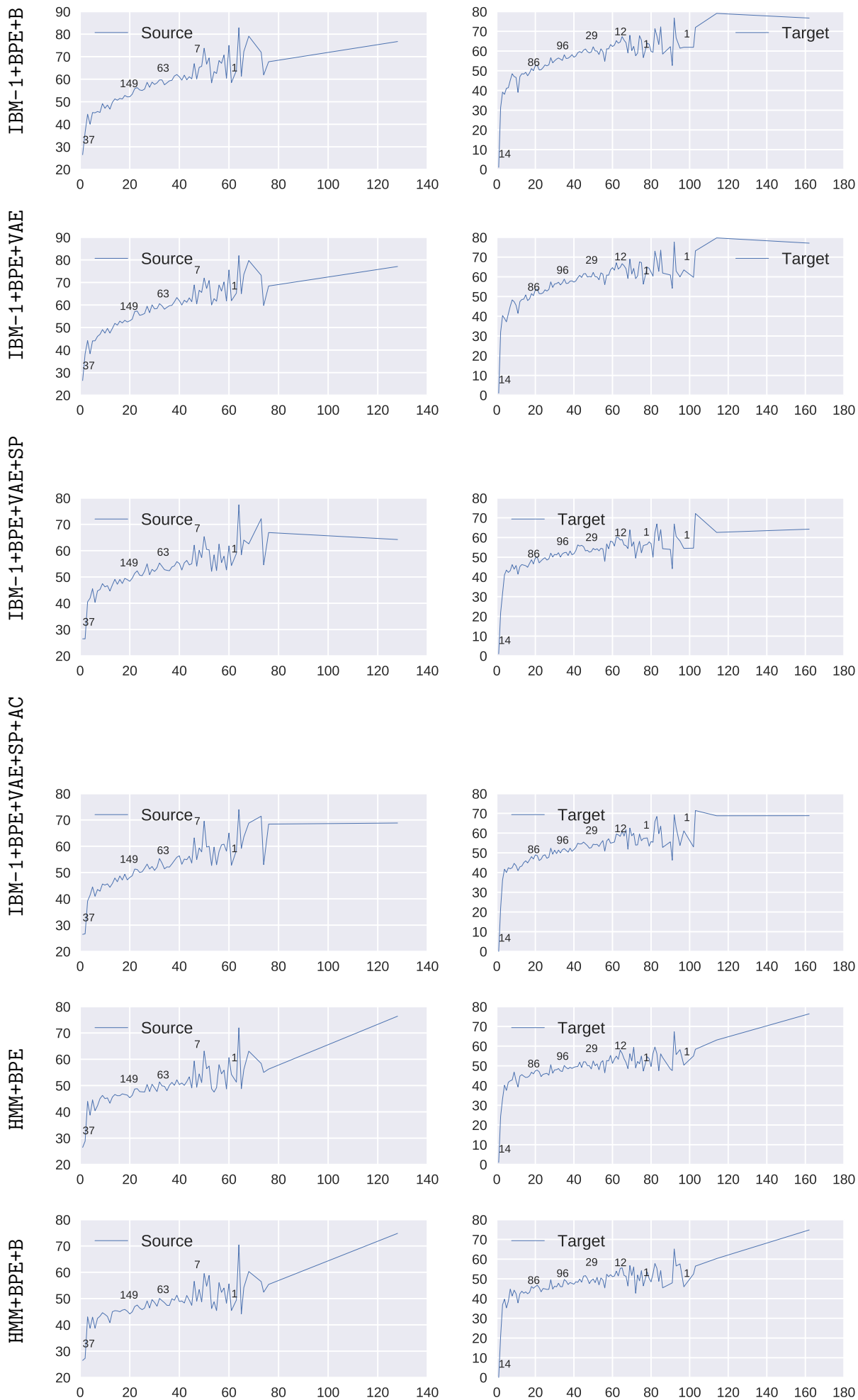


Figure D.64: Variational models for the direction Japanese-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.





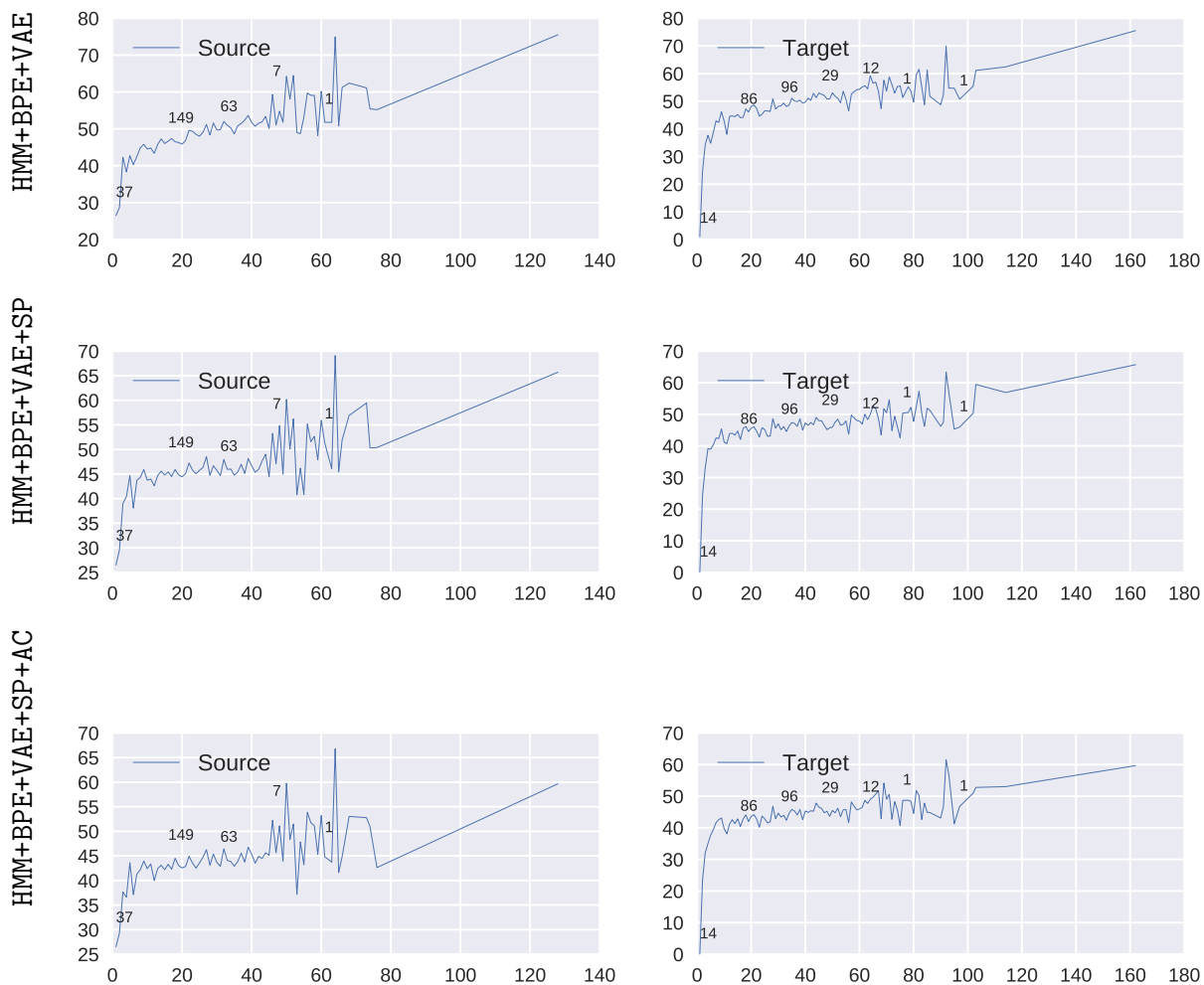
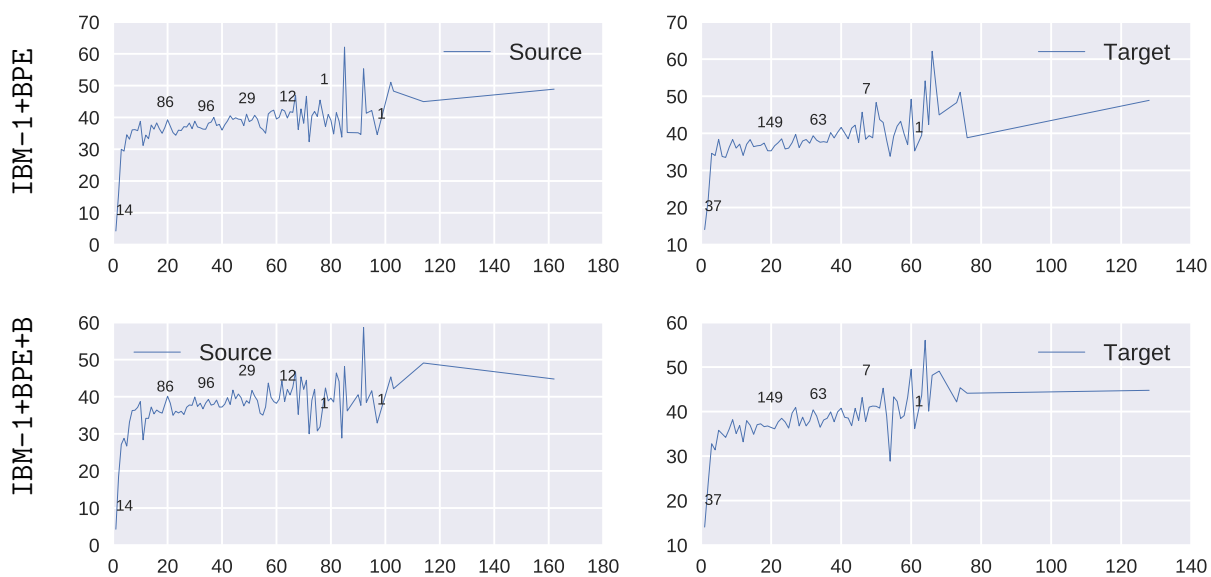
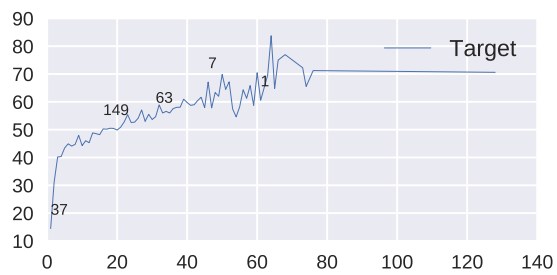
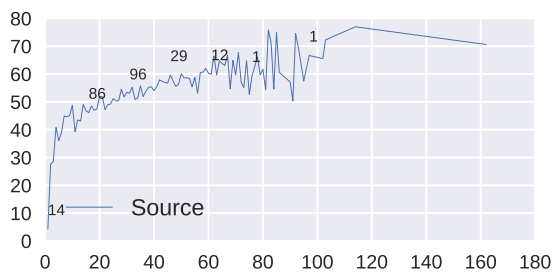


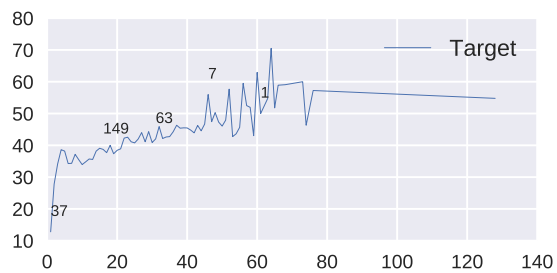
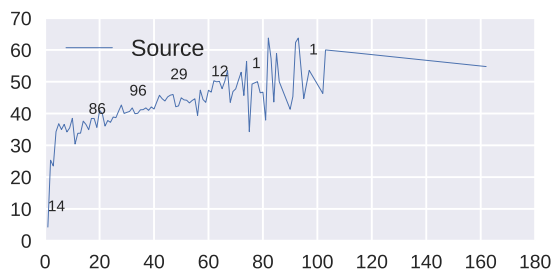
Figure D.65: Variational models for the direction English-Vietnamese: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.



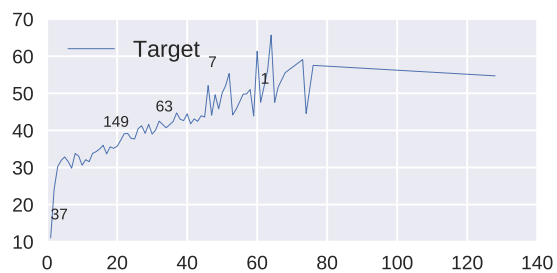
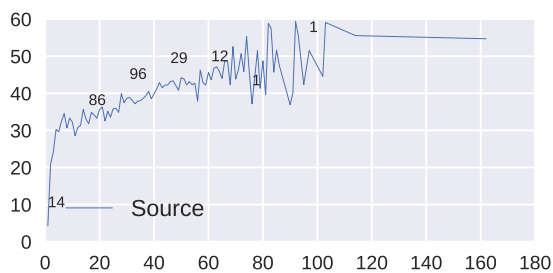
IBM-1+BPE+VAE



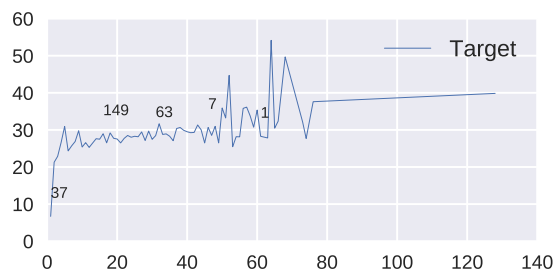
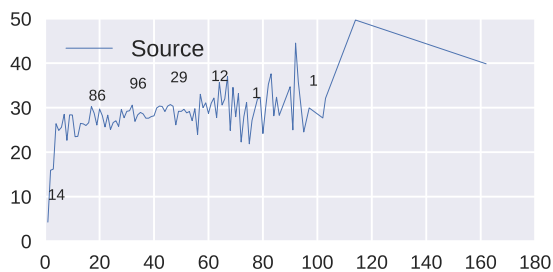
IBM-1+BPE+VAE+SP



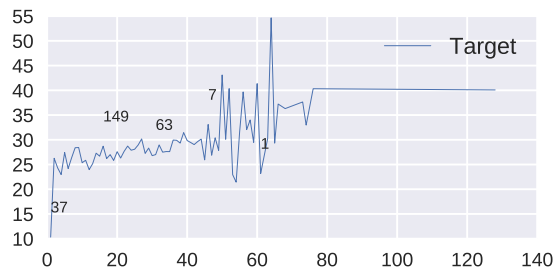
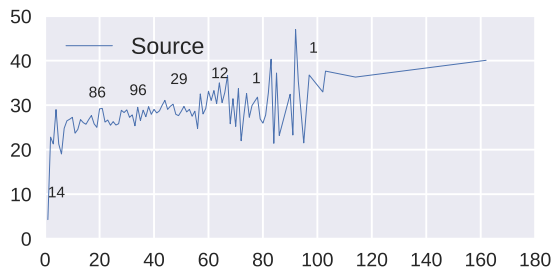
IBM-1+BPE+VAE+SP+AC



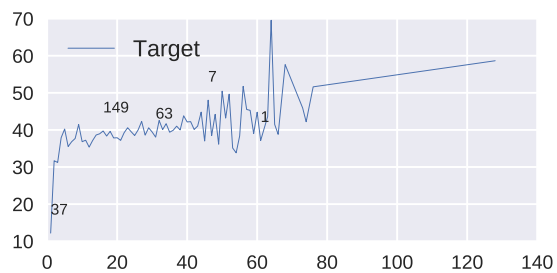
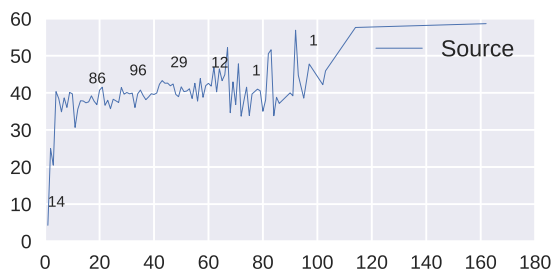
HMM+BPE



HMM+BPE+B



HMM+BPE+VAE



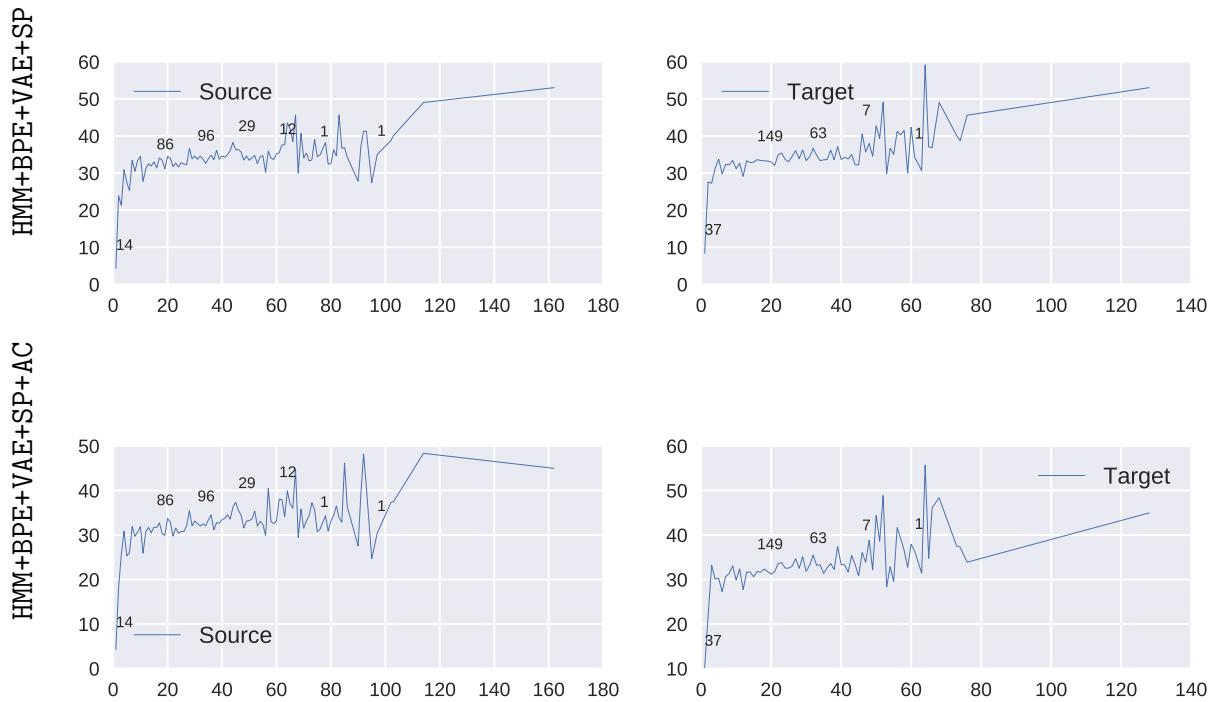
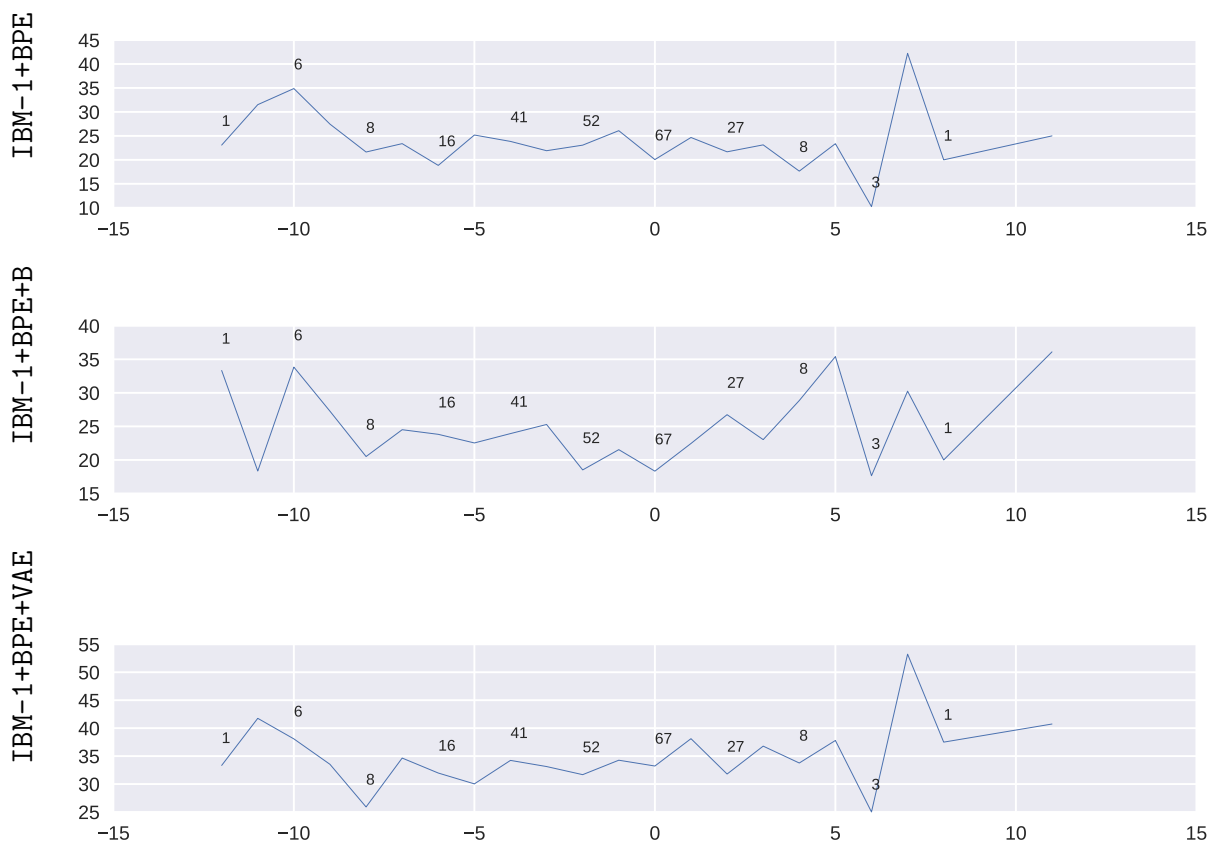
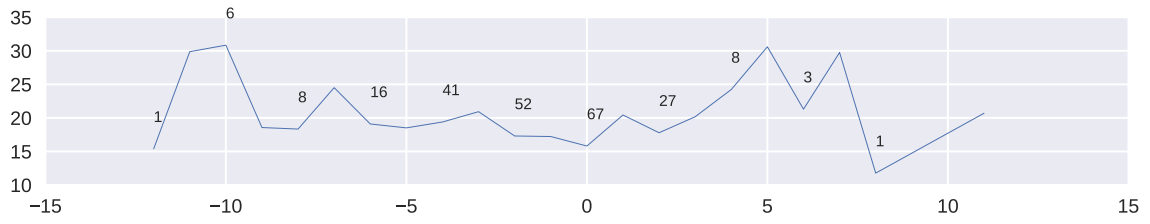


Figure D.66: Variational models for the direction Vietnamese-English: AER score as a function of sentence length. The x-axis shows the sentence length. The y-axis represents the AER. The annotation displays the number of sentences.

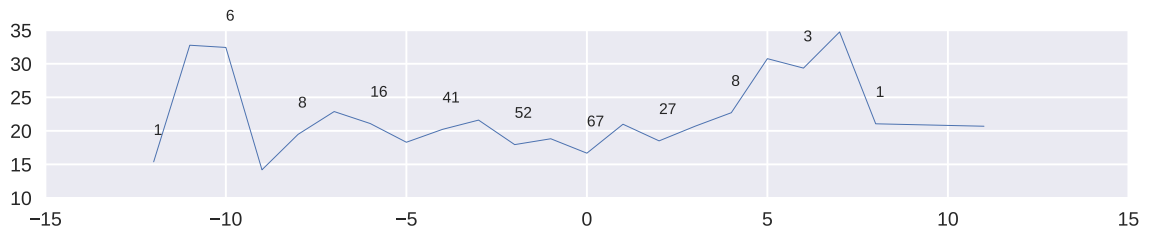
D.12 AER as a function of sentence length difference



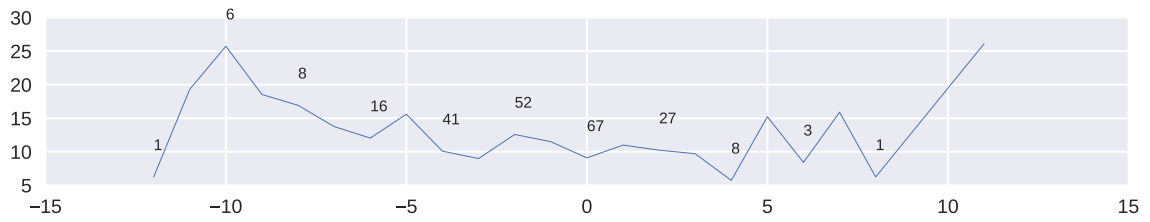
IBM-1+BPE+VAE+SP



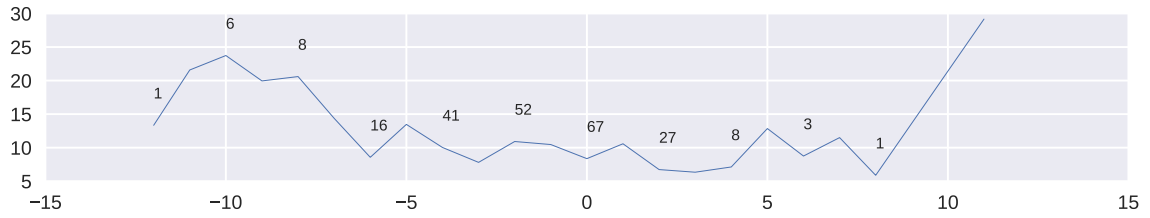
IBM-1+BPE+VAE+SP+AC



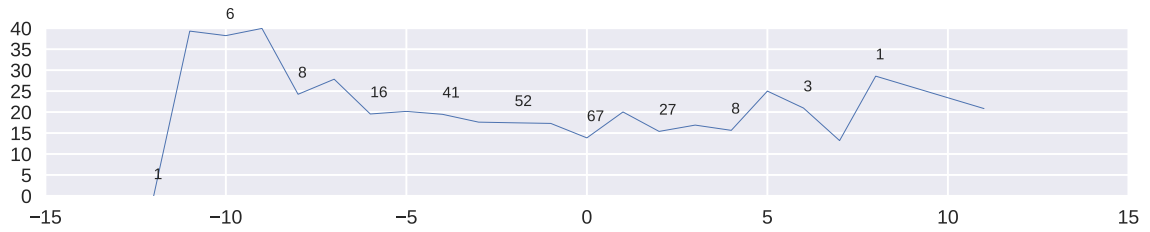
HMM+BPE



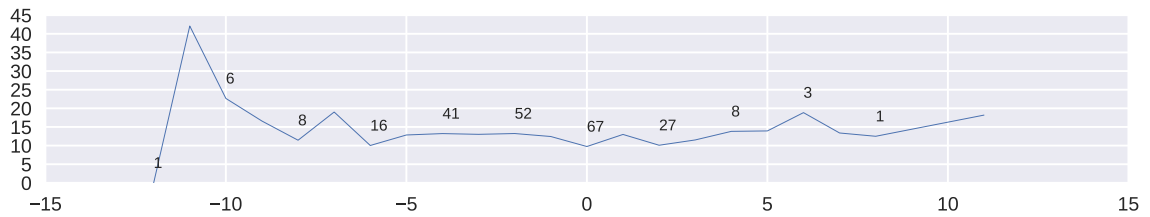
HMM+BPE+B



HMM+BPE+VAE



HMM+BPE+VAE+SP



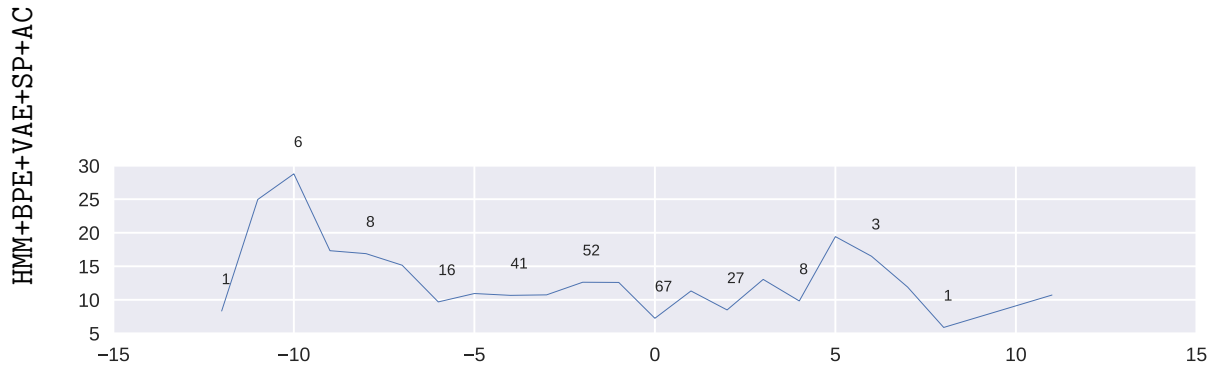
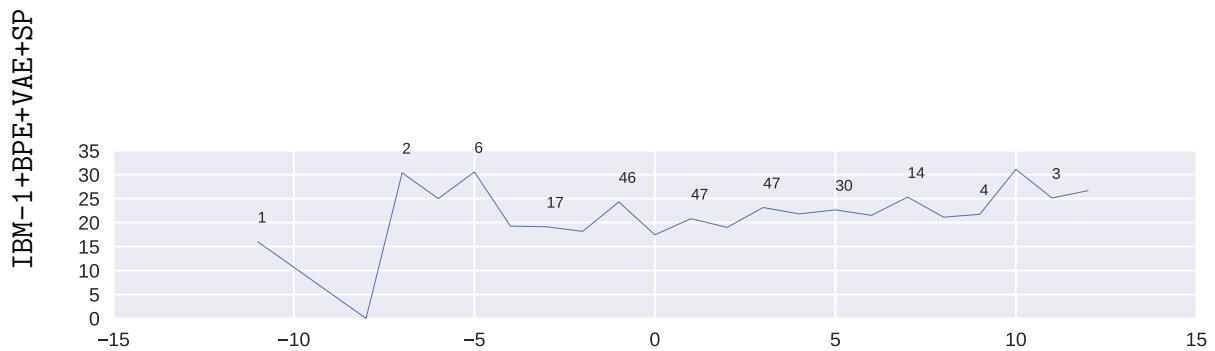
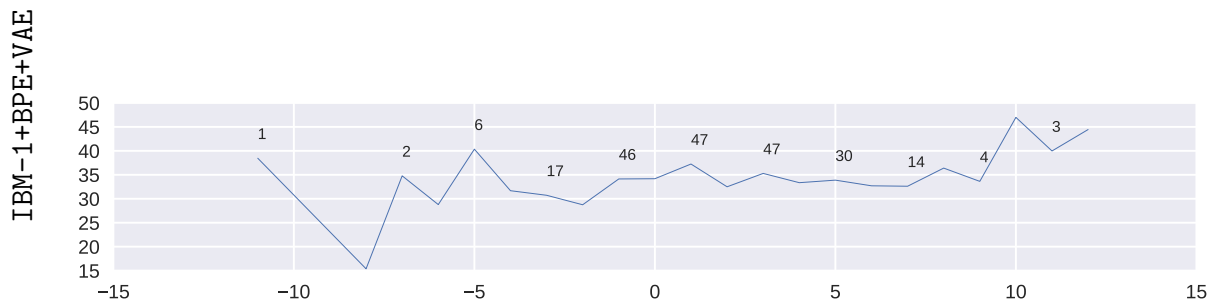
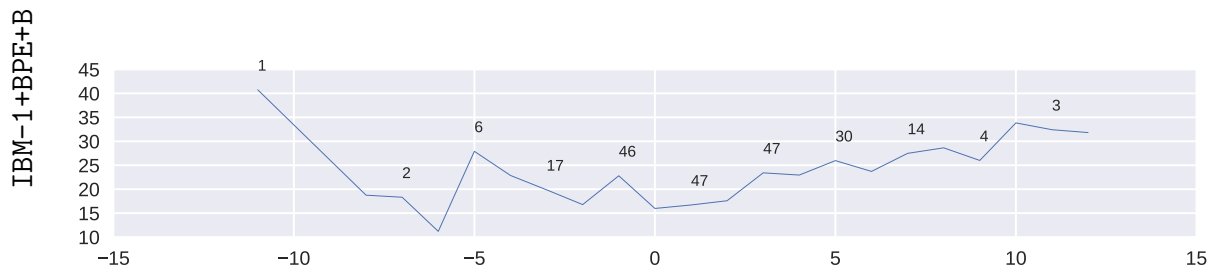
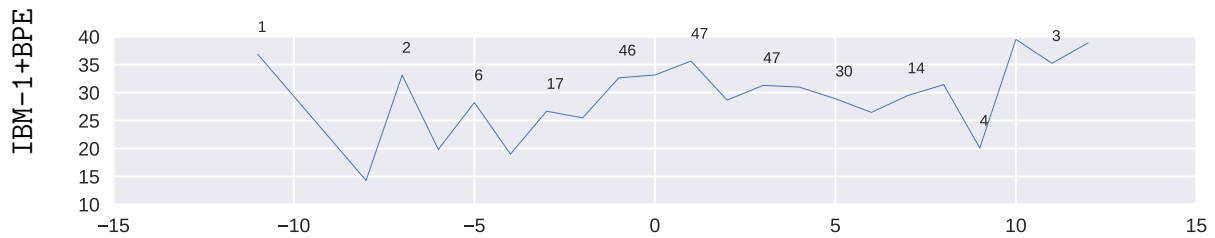


Figure D.67: Models for the direction English-French: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



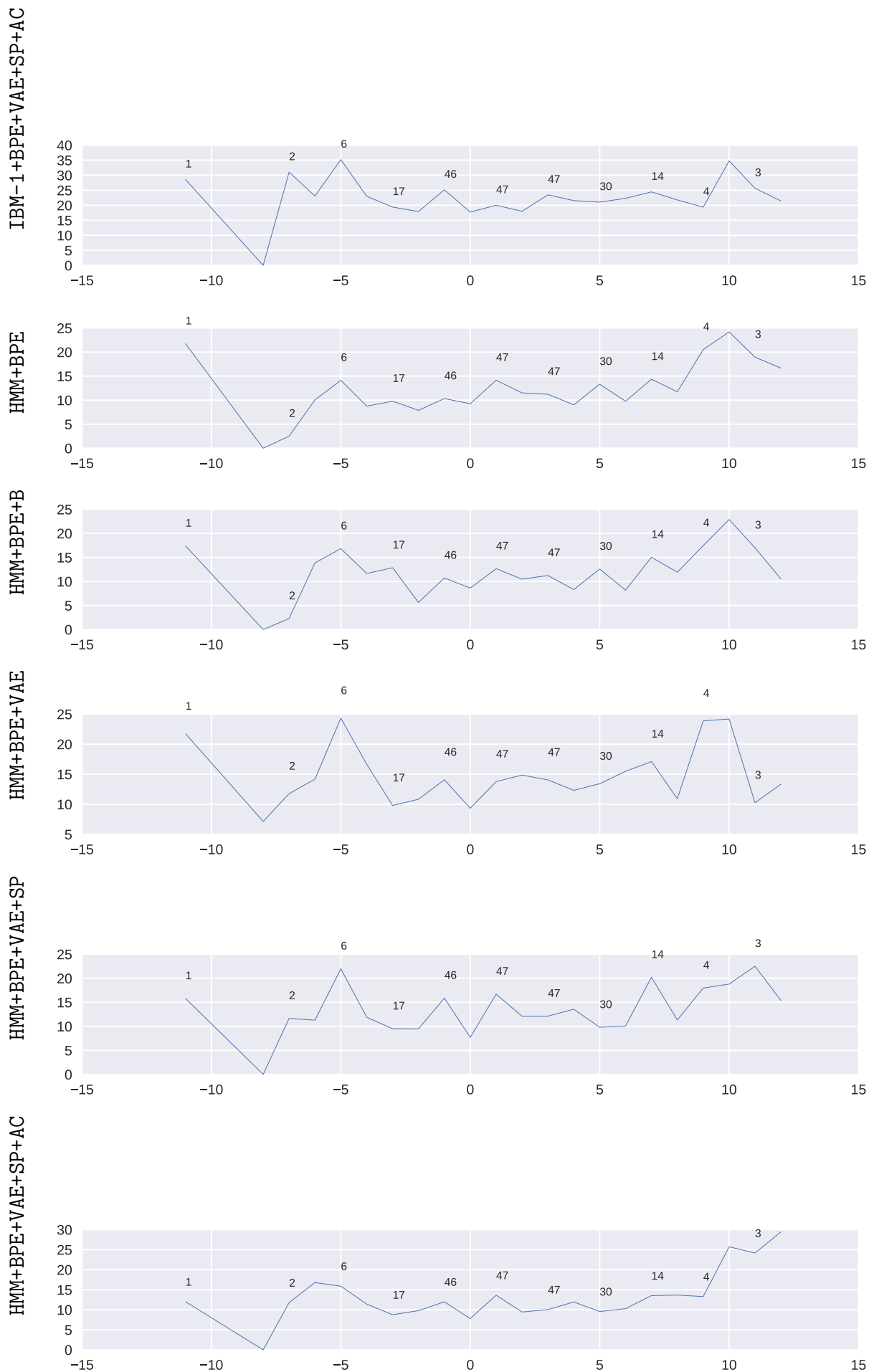
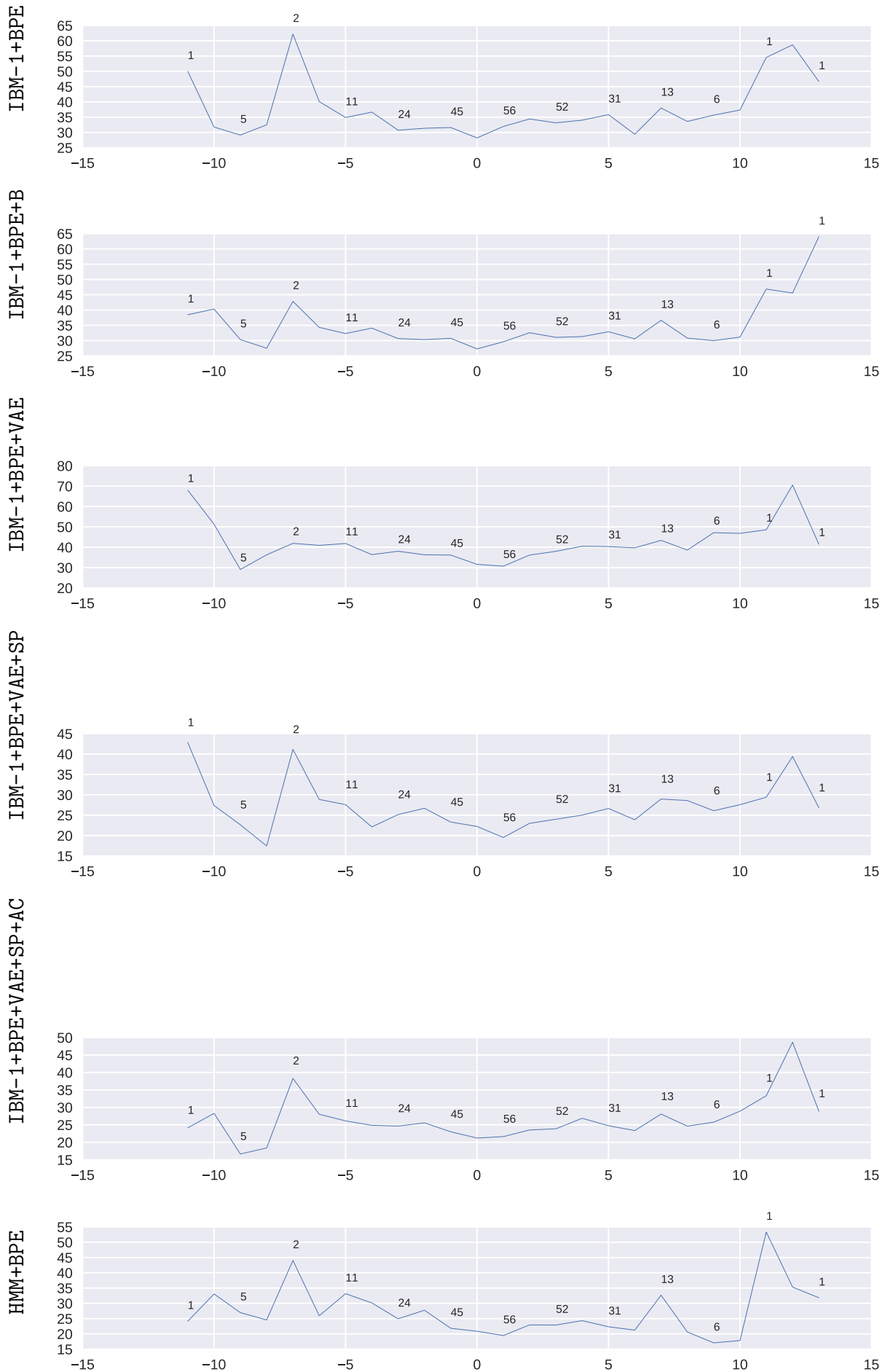


Figure D.68: Models for the direction French-English: AER score as a function of sentence

length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



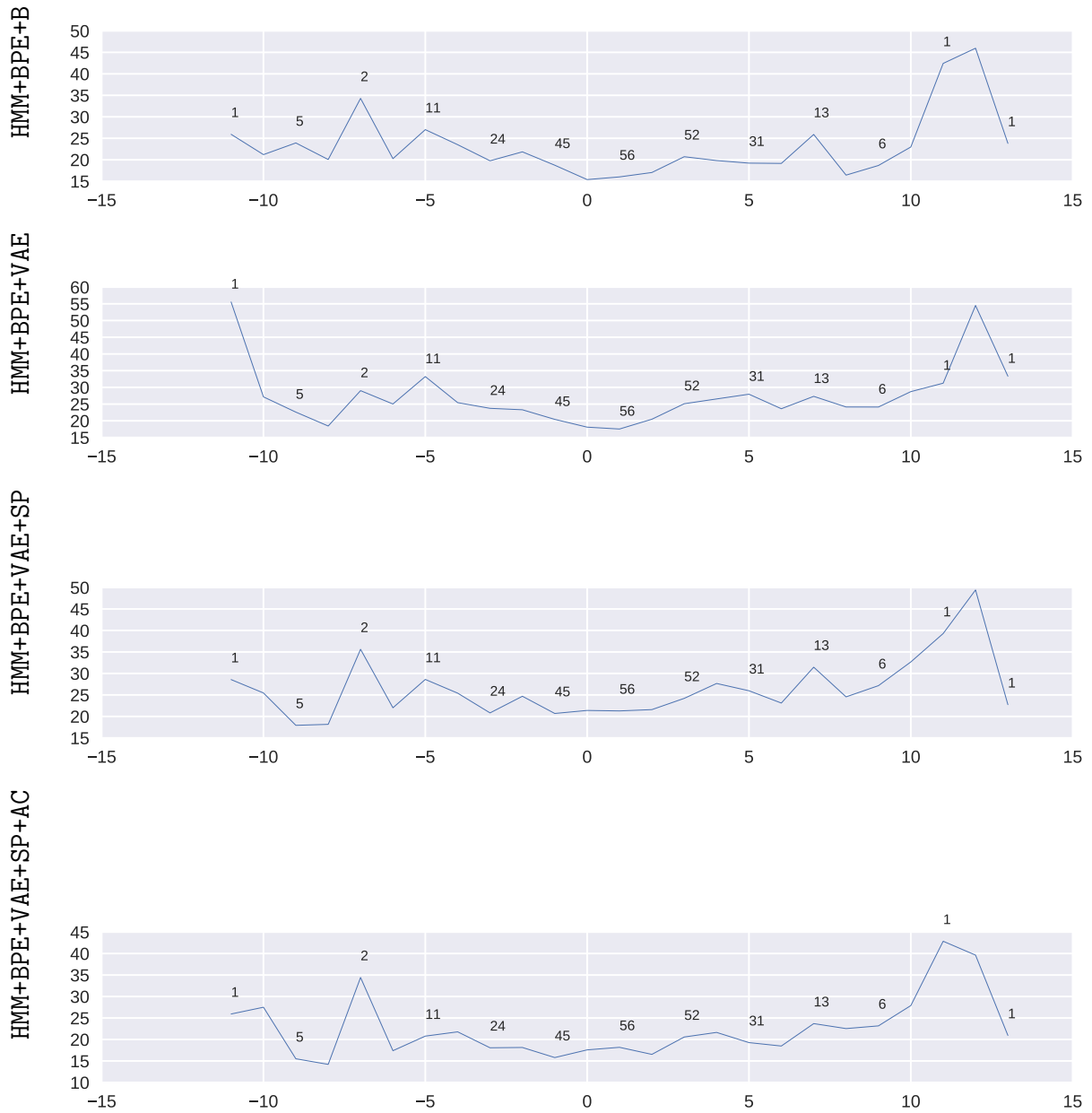
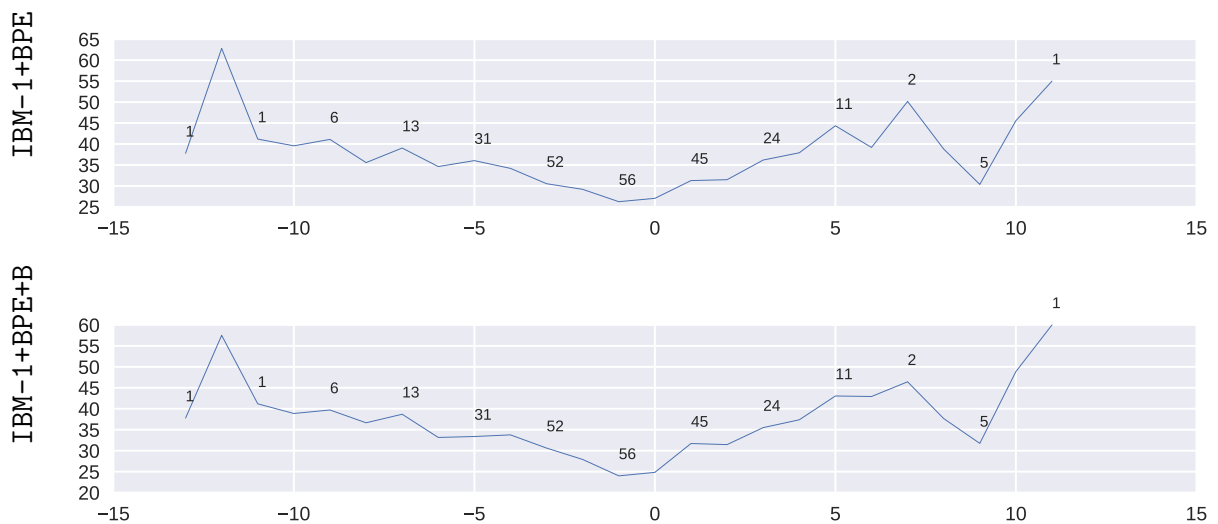
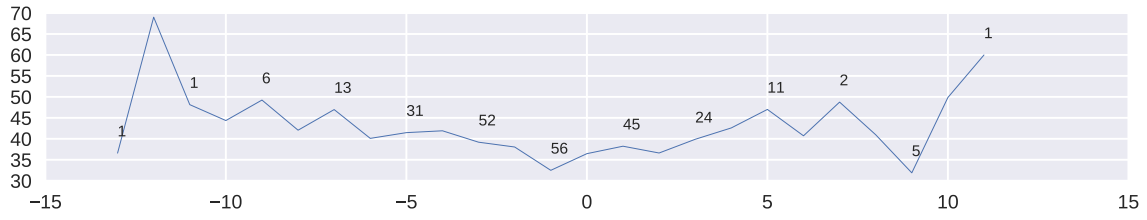


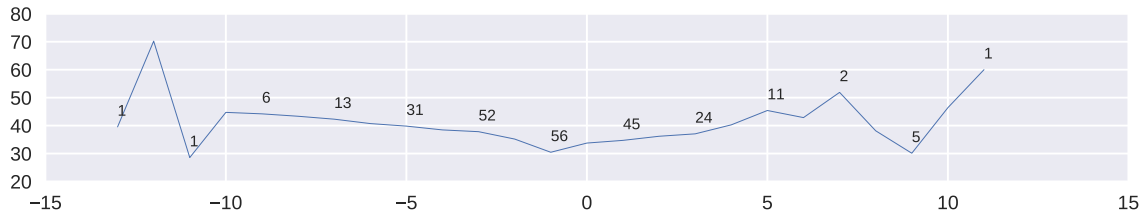
Figure D.69: Models for the direction English-German: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



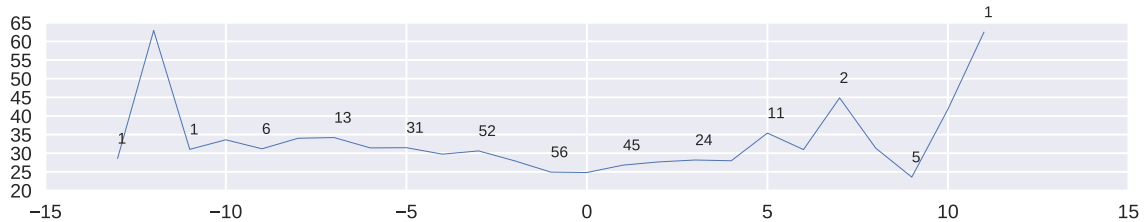
IBM-1+BPE+VAE



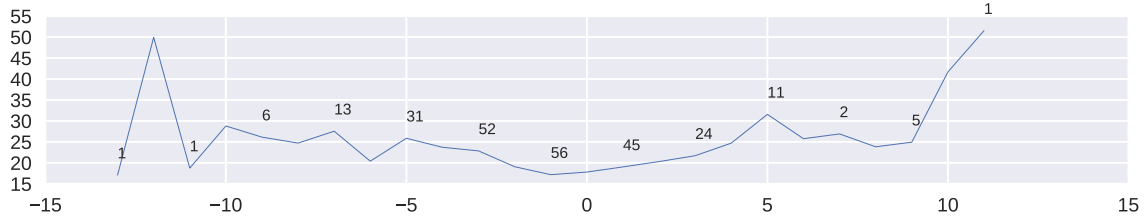
IBM-1+BPE+VAE+SP



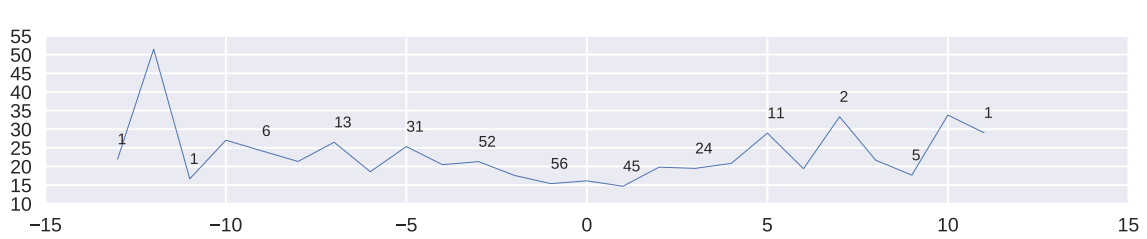
IBM-1+BPE+VAE+SP+AC



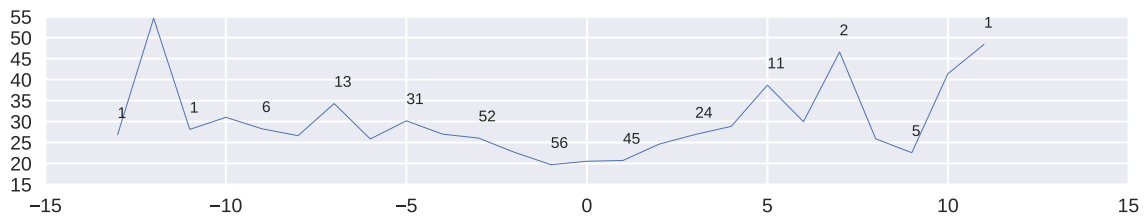
HMM+BPE



HMM+BPE+B



HMM+BPE+VAE



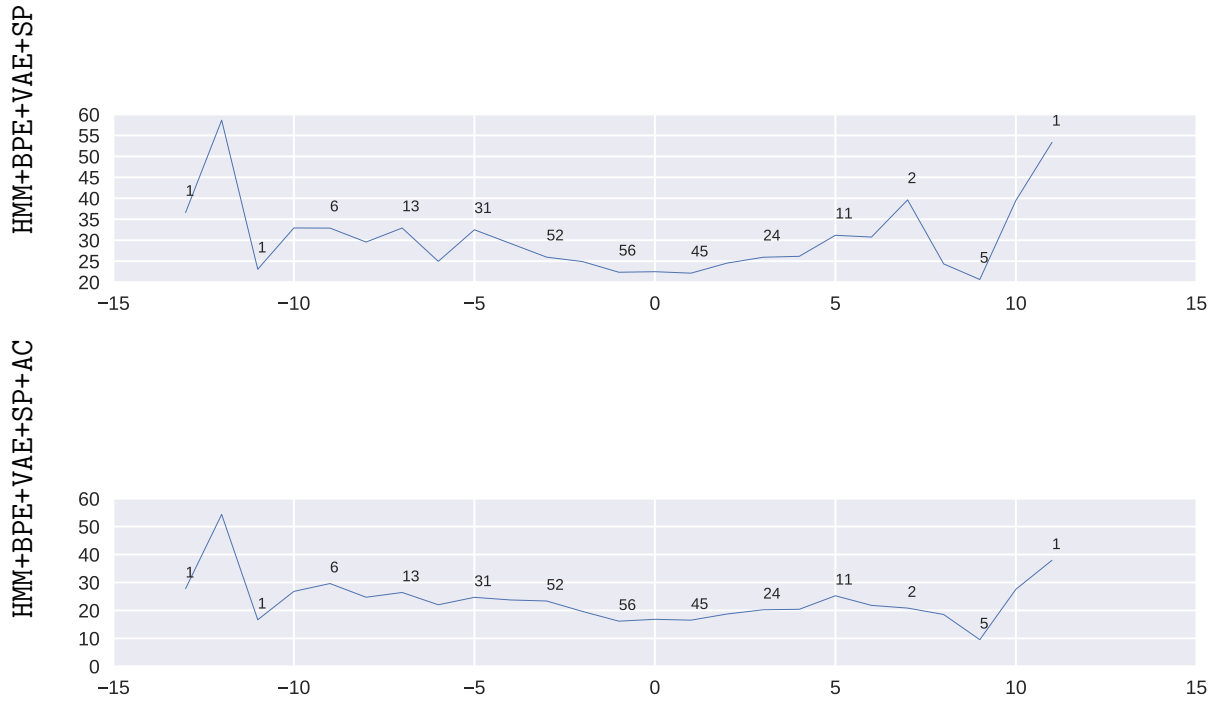
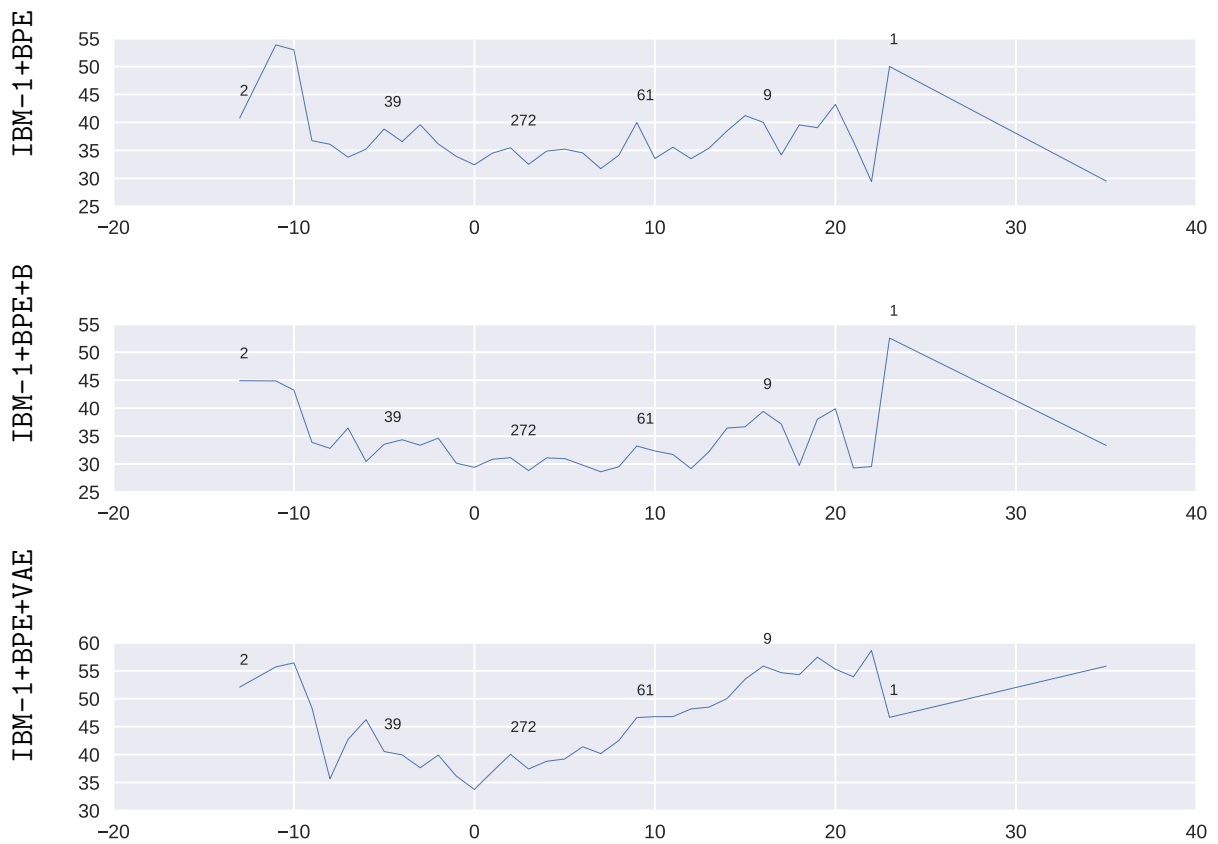
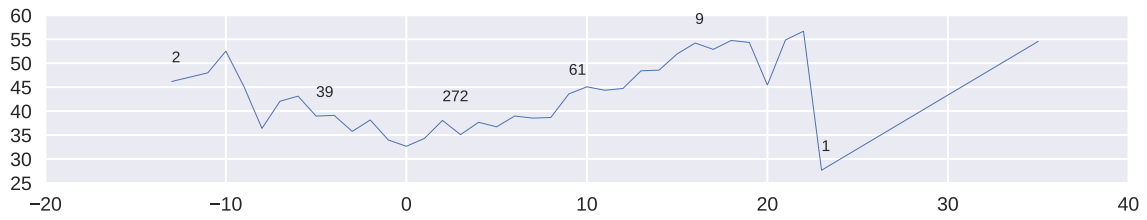


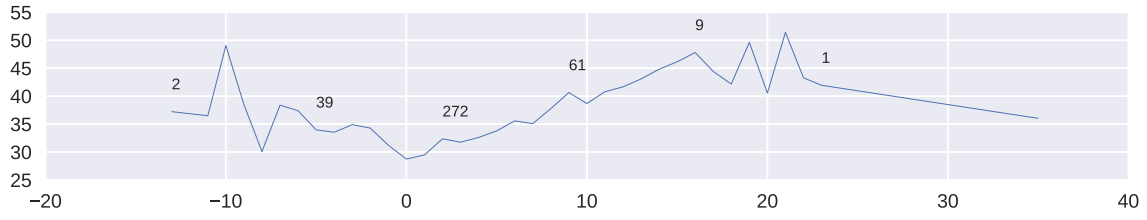
Figure D.70: Models for the direction German-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



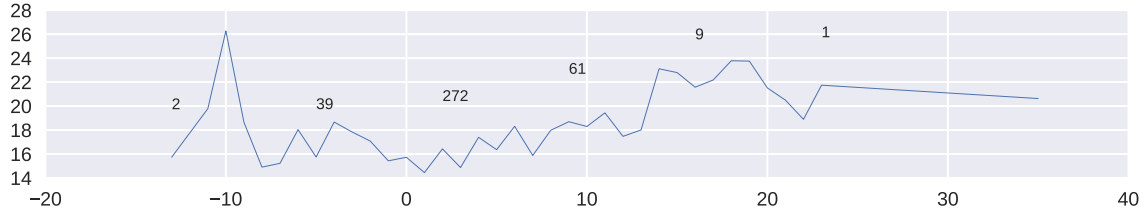
IBM-1+BPE+VAE+SP



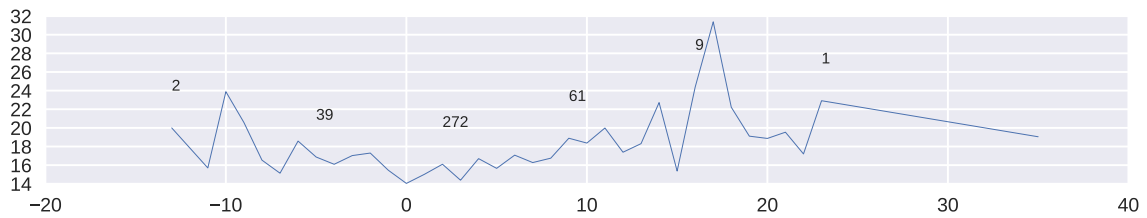
IBM-1+BPE+VAE+SP+AC



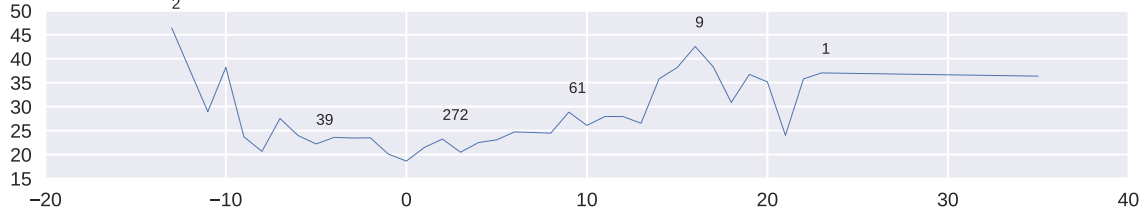
HMM+BPE



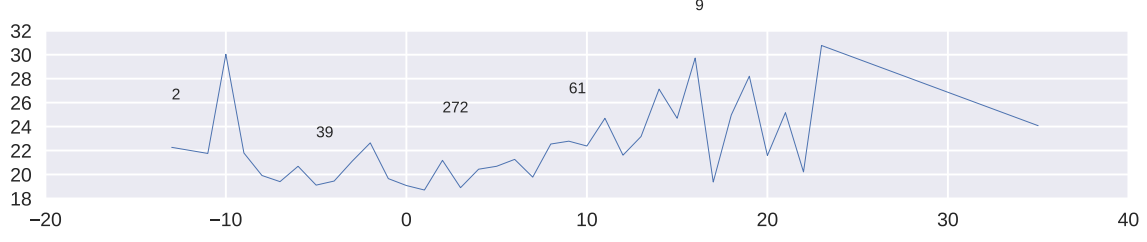
HMM+BPE+B



HMM+BPE+VAE



HMM+BPE+VAE+SP



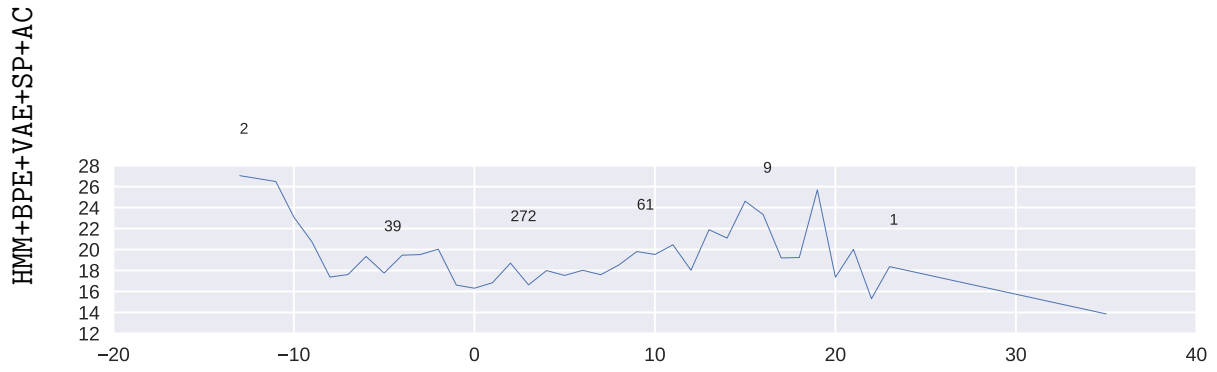
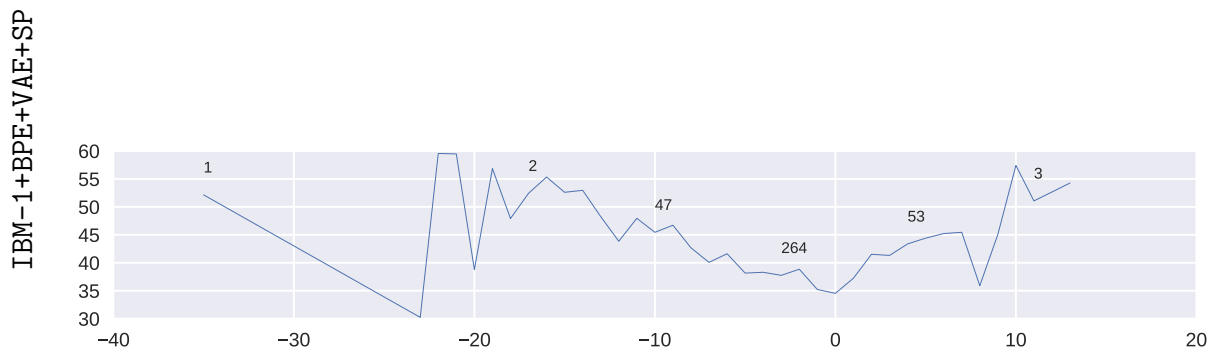
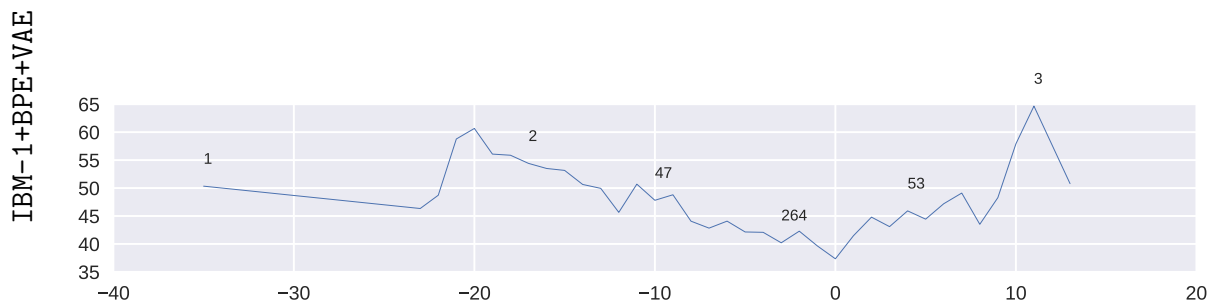
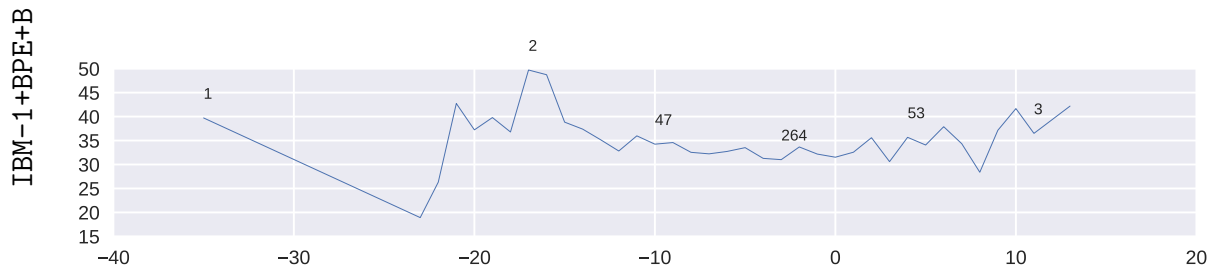
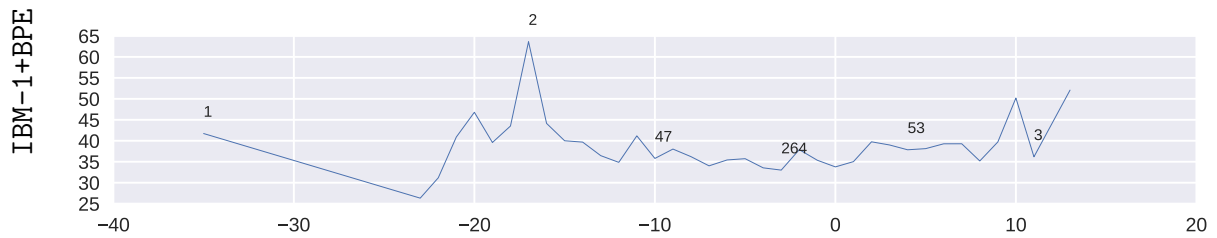


Figure D.71: Models for the direction English-Czech: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



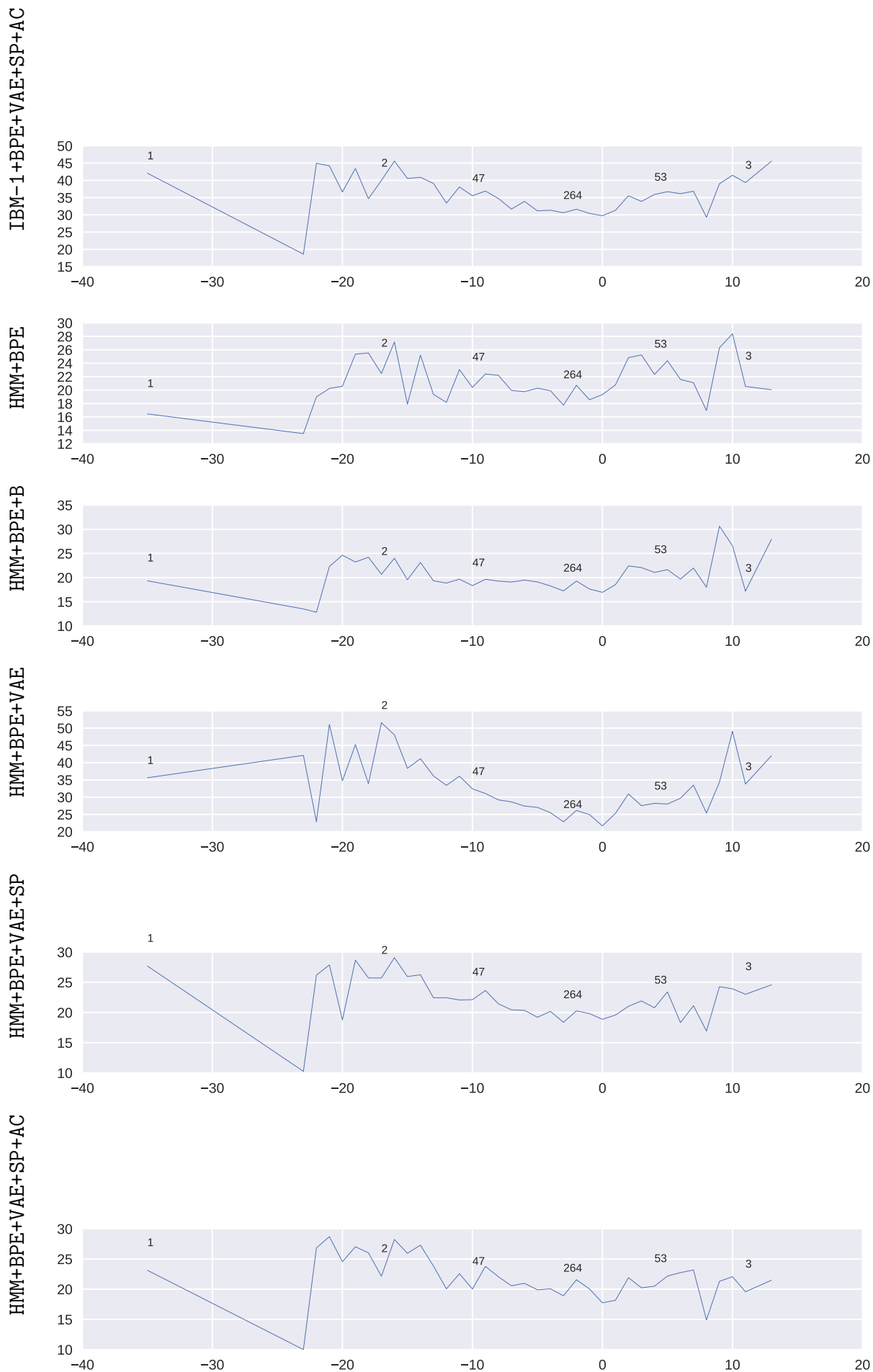
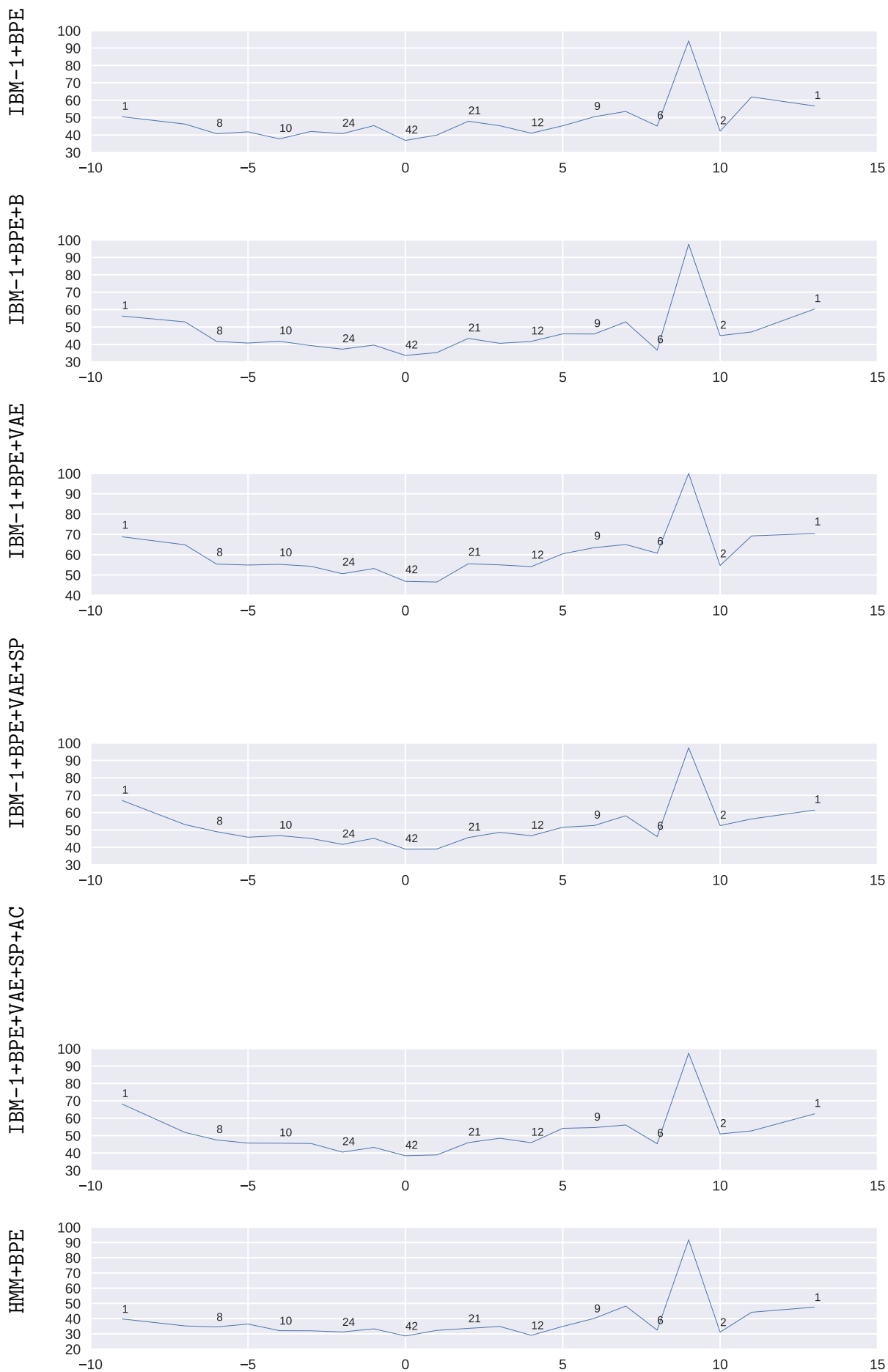


Figure D.72: Models for the direction Czech-English: AER score as a function of sentence

length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



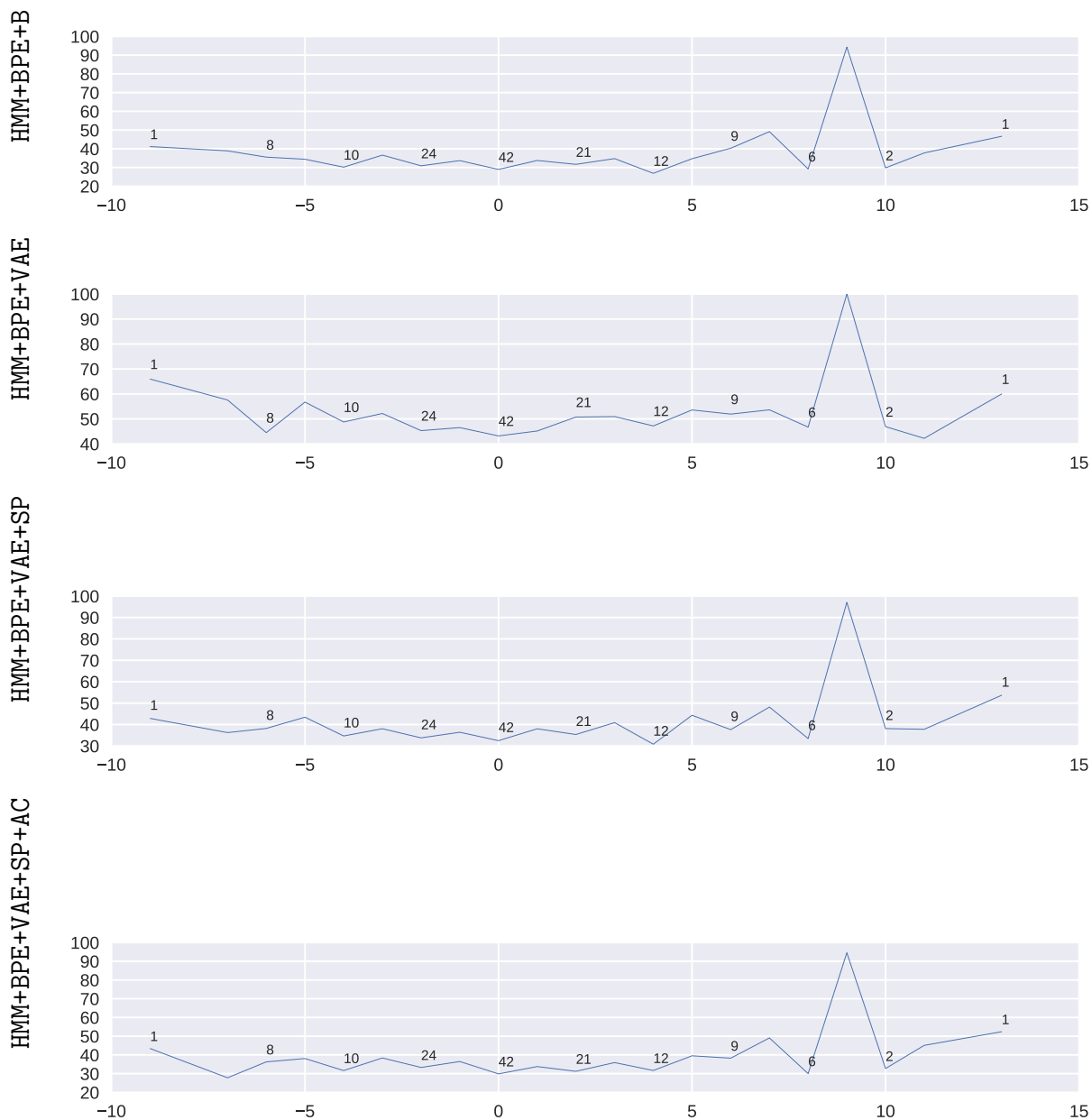
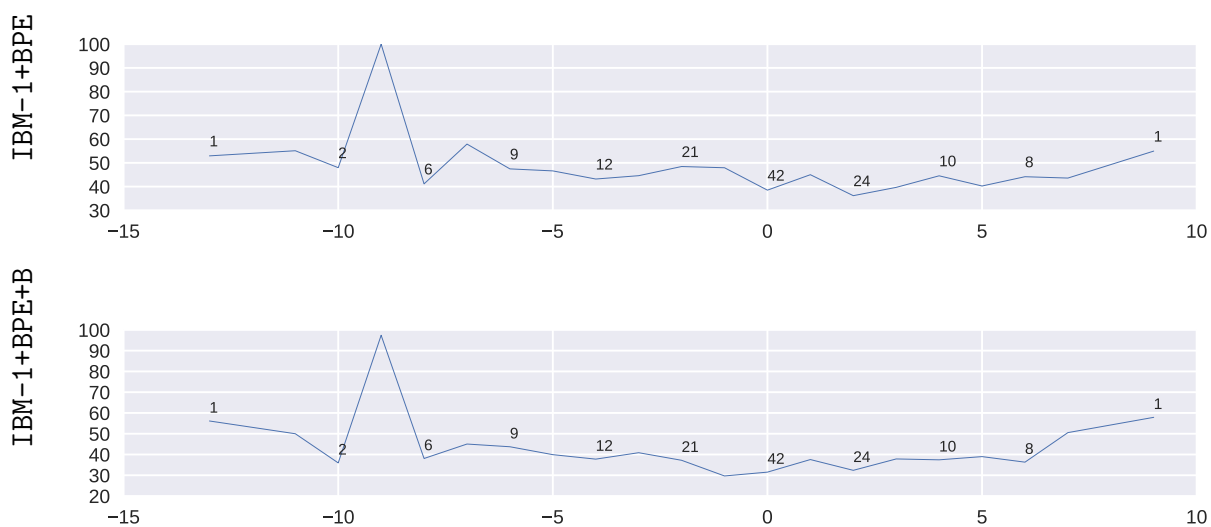
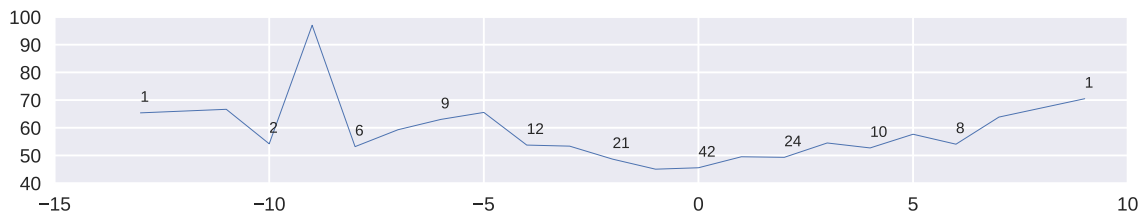


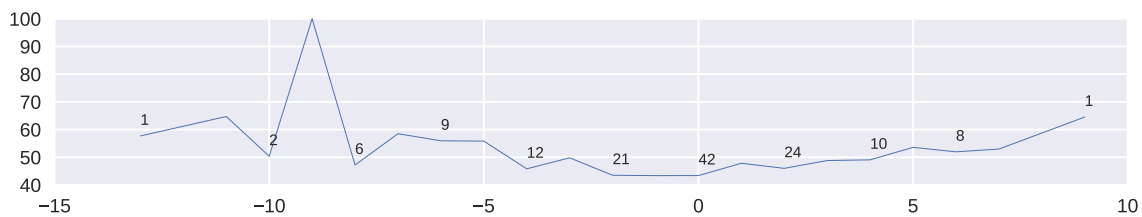
Figure D.73: Models for the direction English-Romanian: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



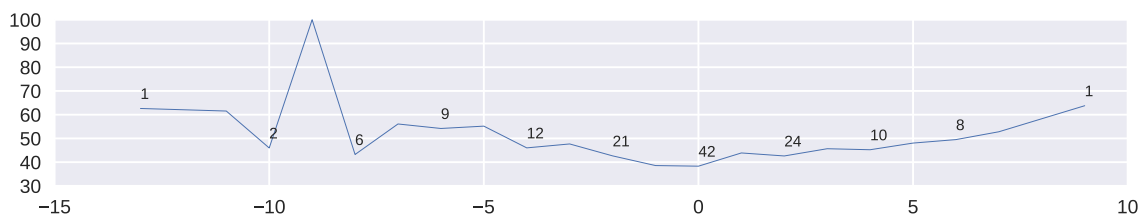
IBM-1+BPE+VAE



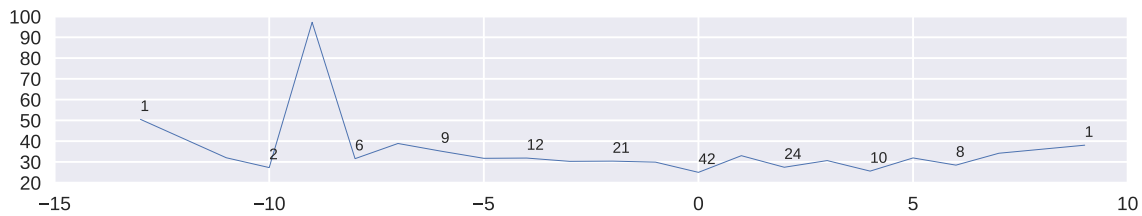
IBM-1+BPE+VAE+SP



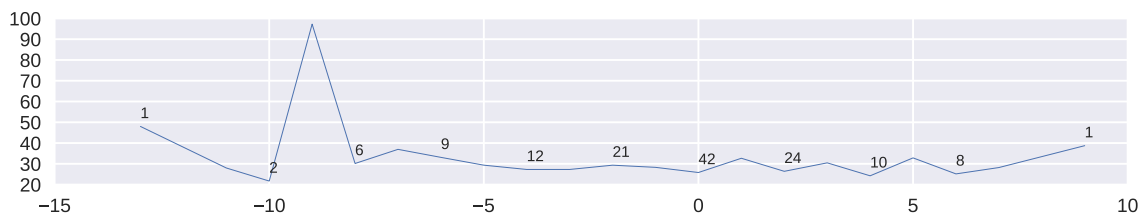
IBM-1+BPE+VAE+SP+AC



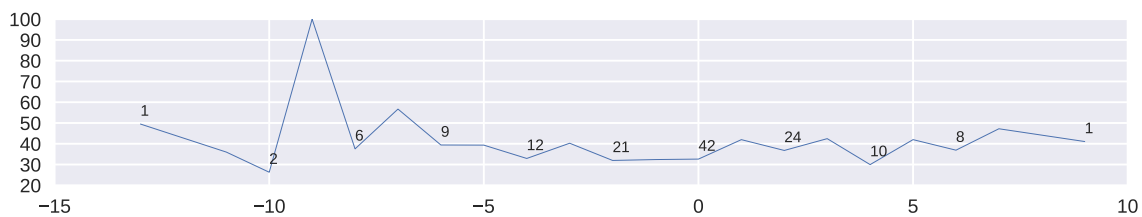
HMM+BPE



HMM+BPE+B



HMM+BPE+VAE



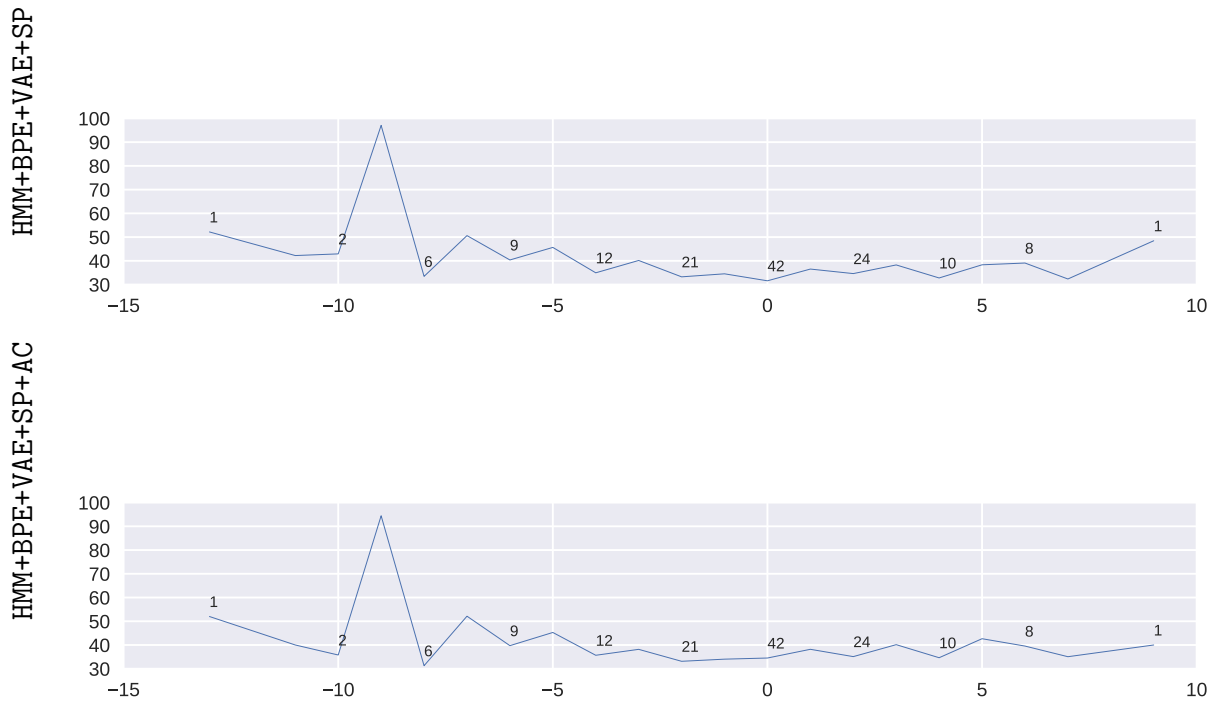
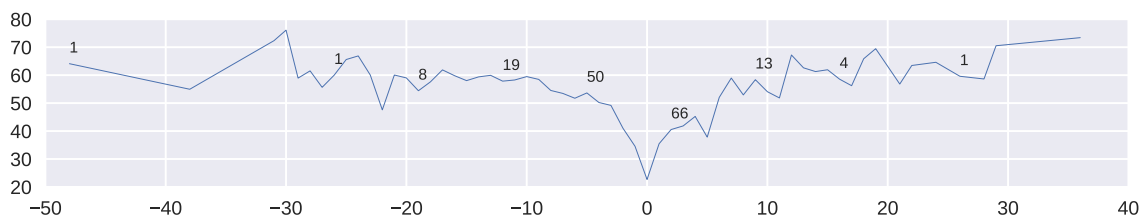


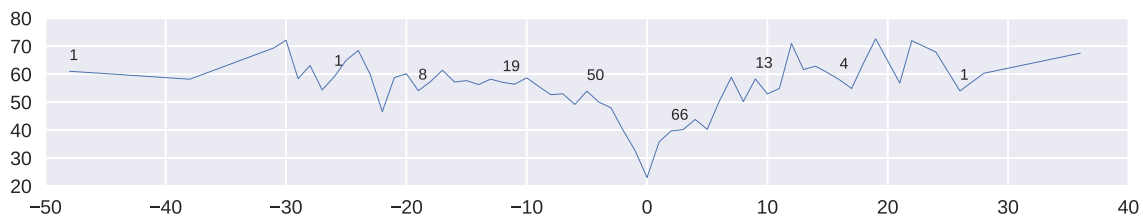
Figure D.74: Models for the direction Romanian-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



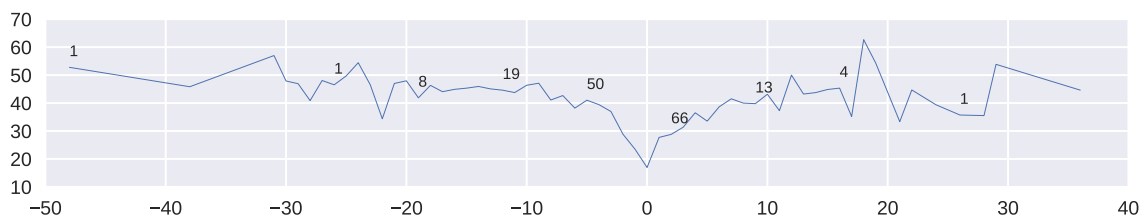
IBM-1+BPE+VAE+SP



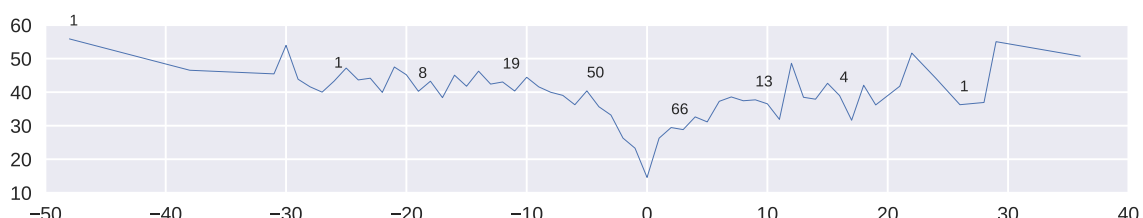
IBM-1+BPE+VAE+SP+AC



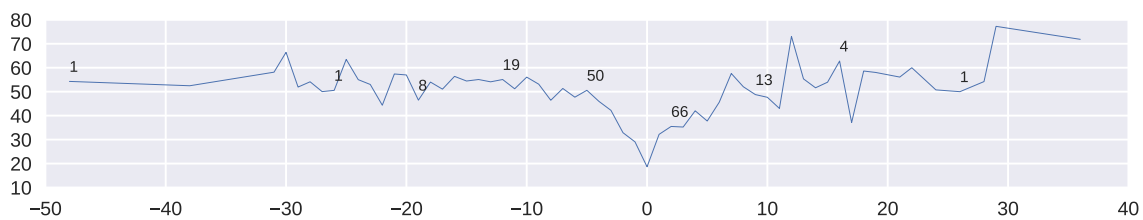
HMM+BPE



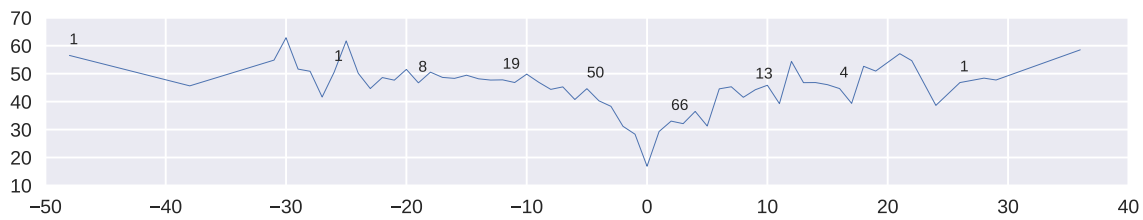
HMM+BPE+B



HMM+BPE+VAE



HMM+BPE+VAE+SP



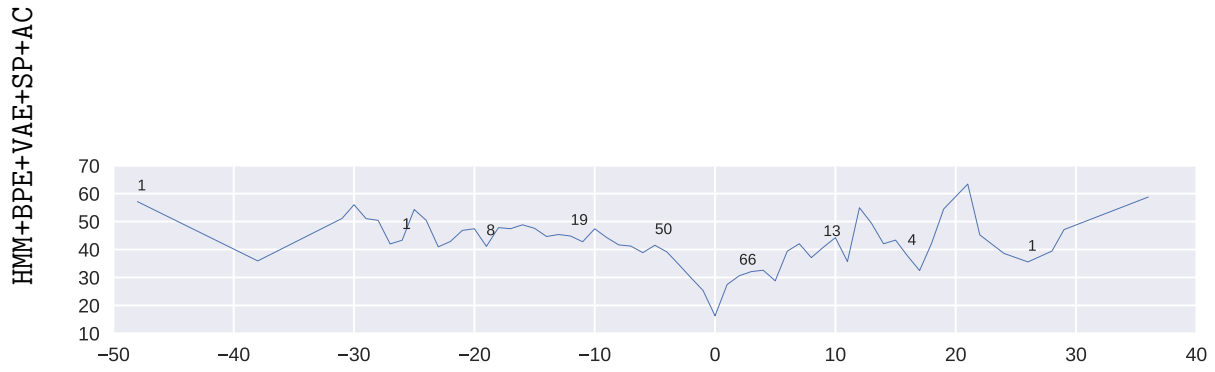
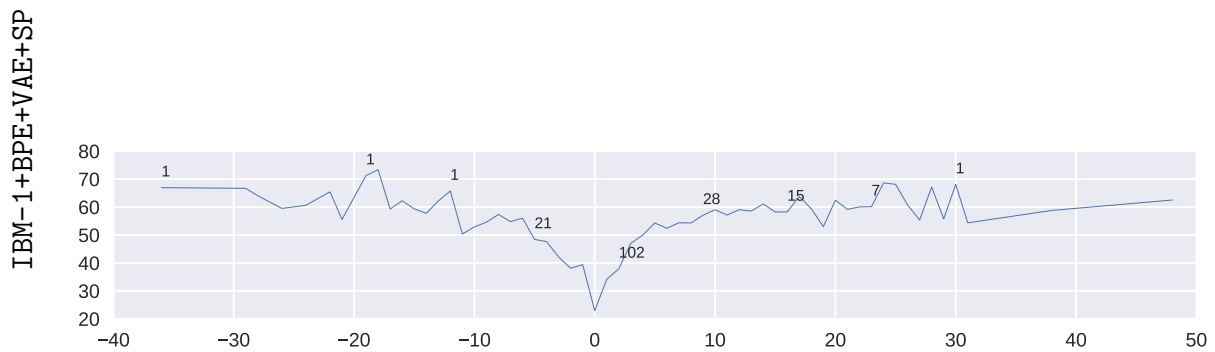
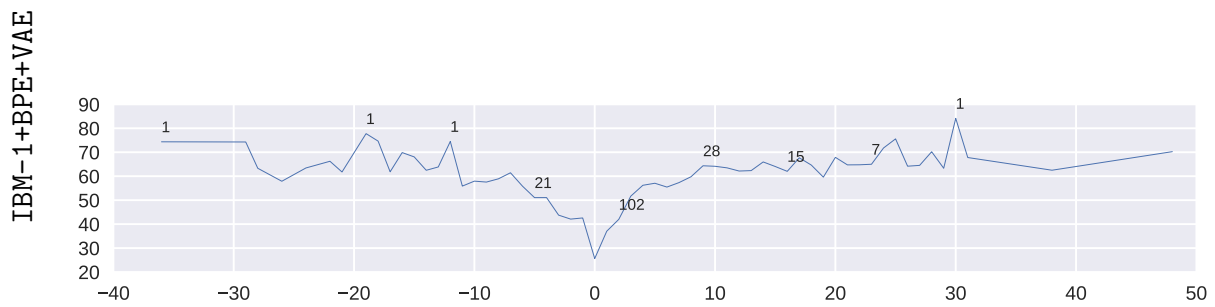
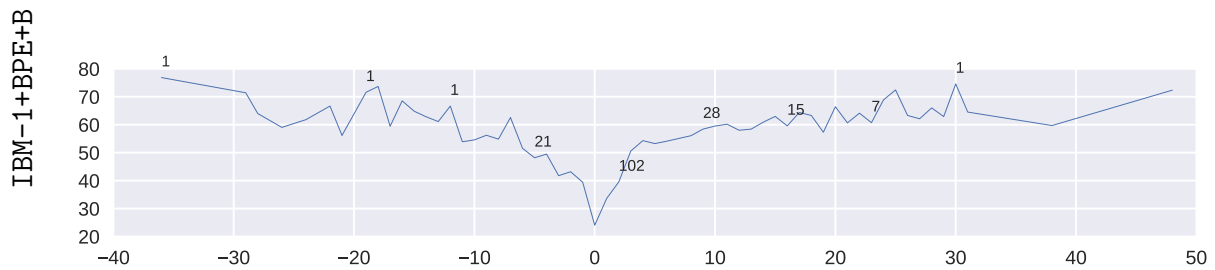
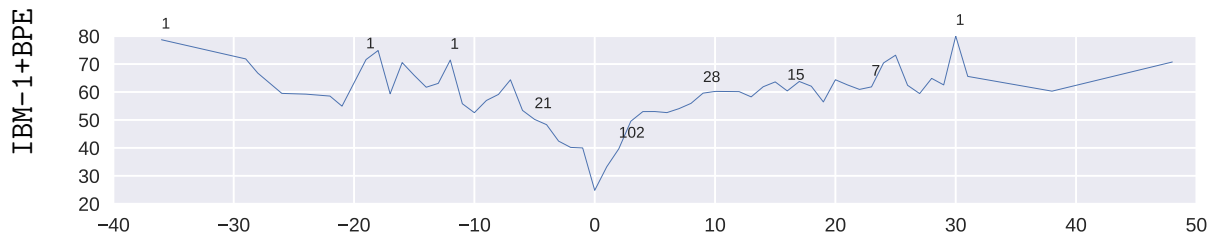


Figure D.75: Models for the direction English-Japanese: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



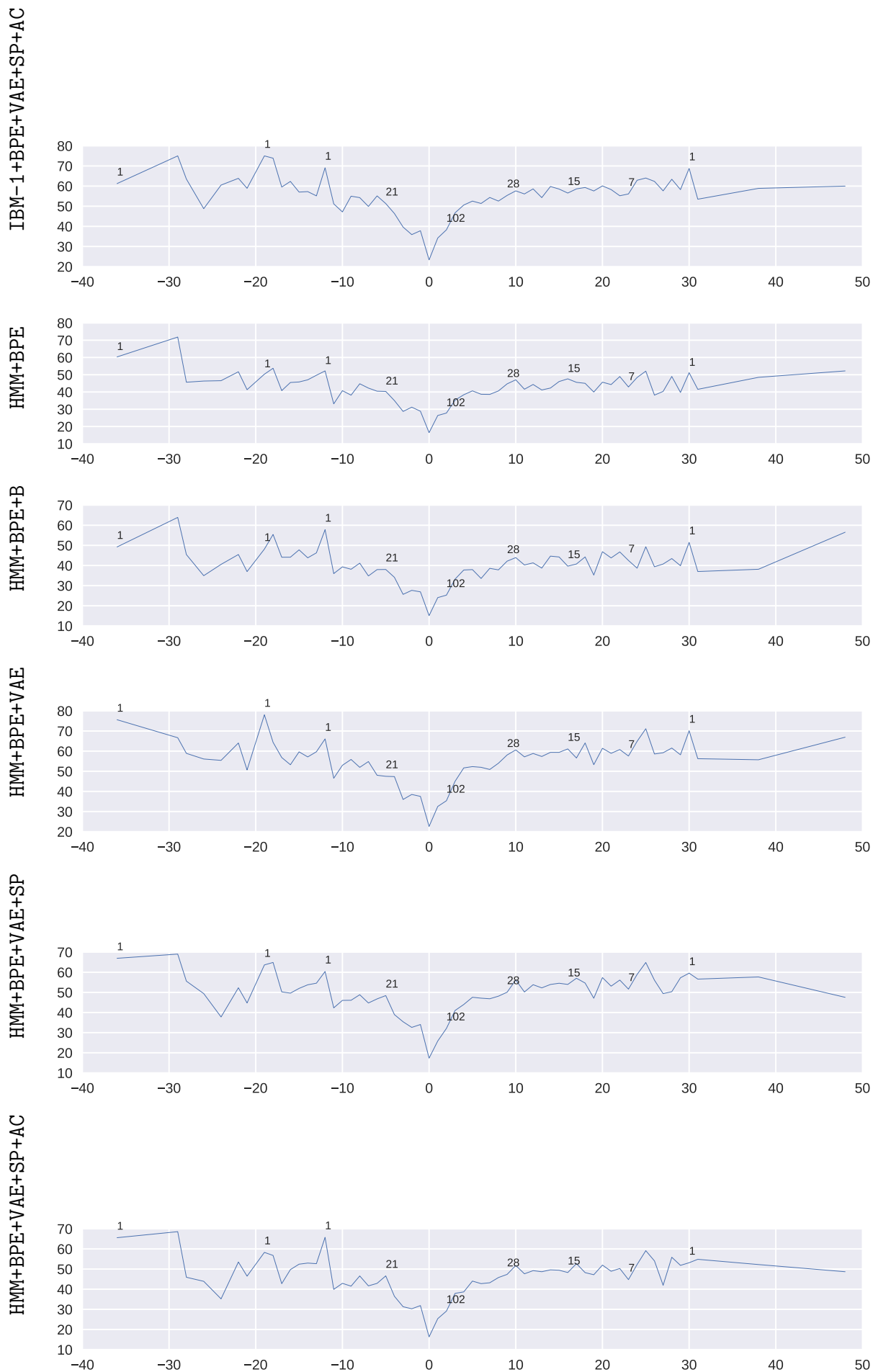
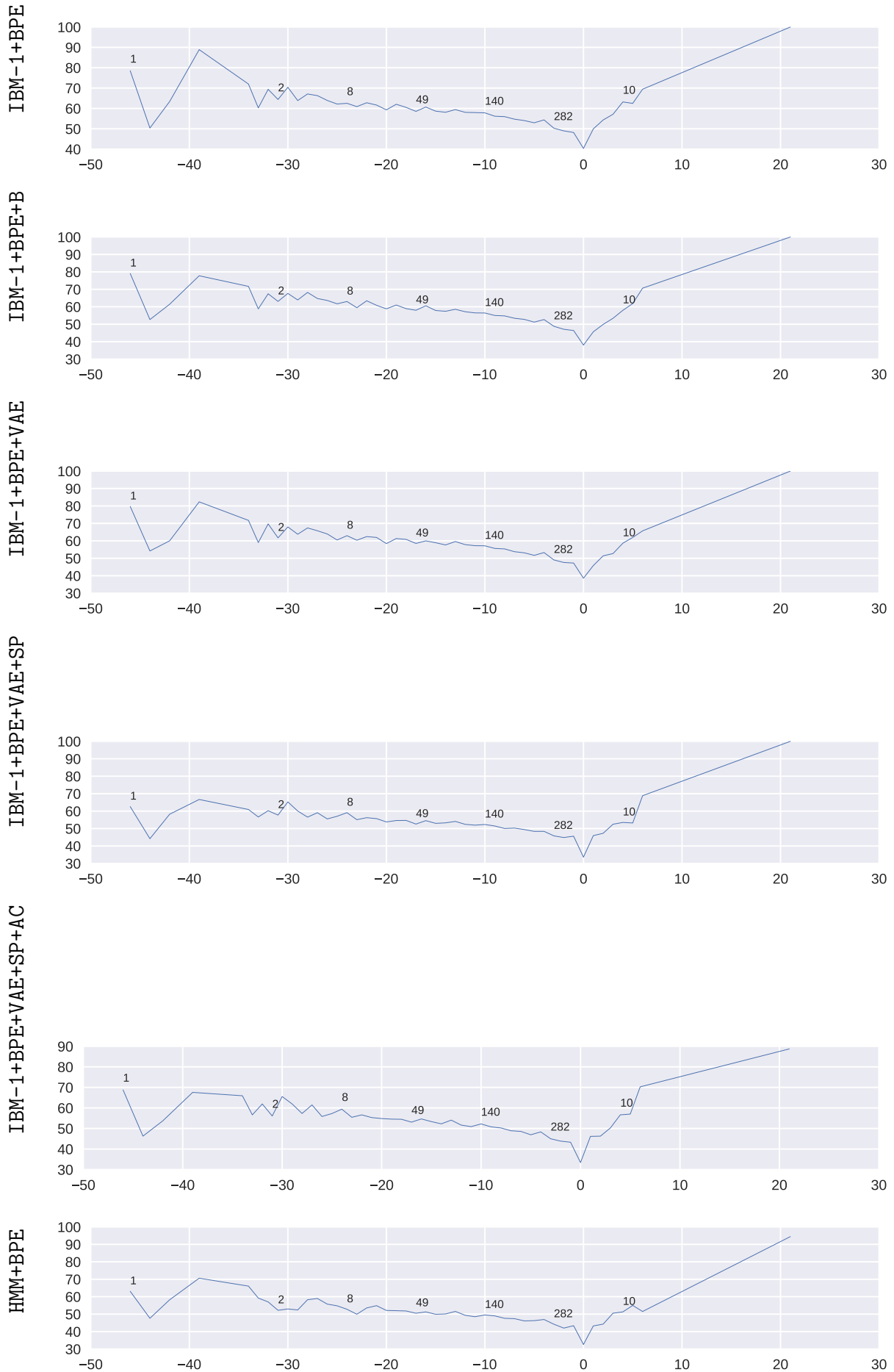


Figure D.76: Models for the direction Japanese-English: AER score as a function of sentence

length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



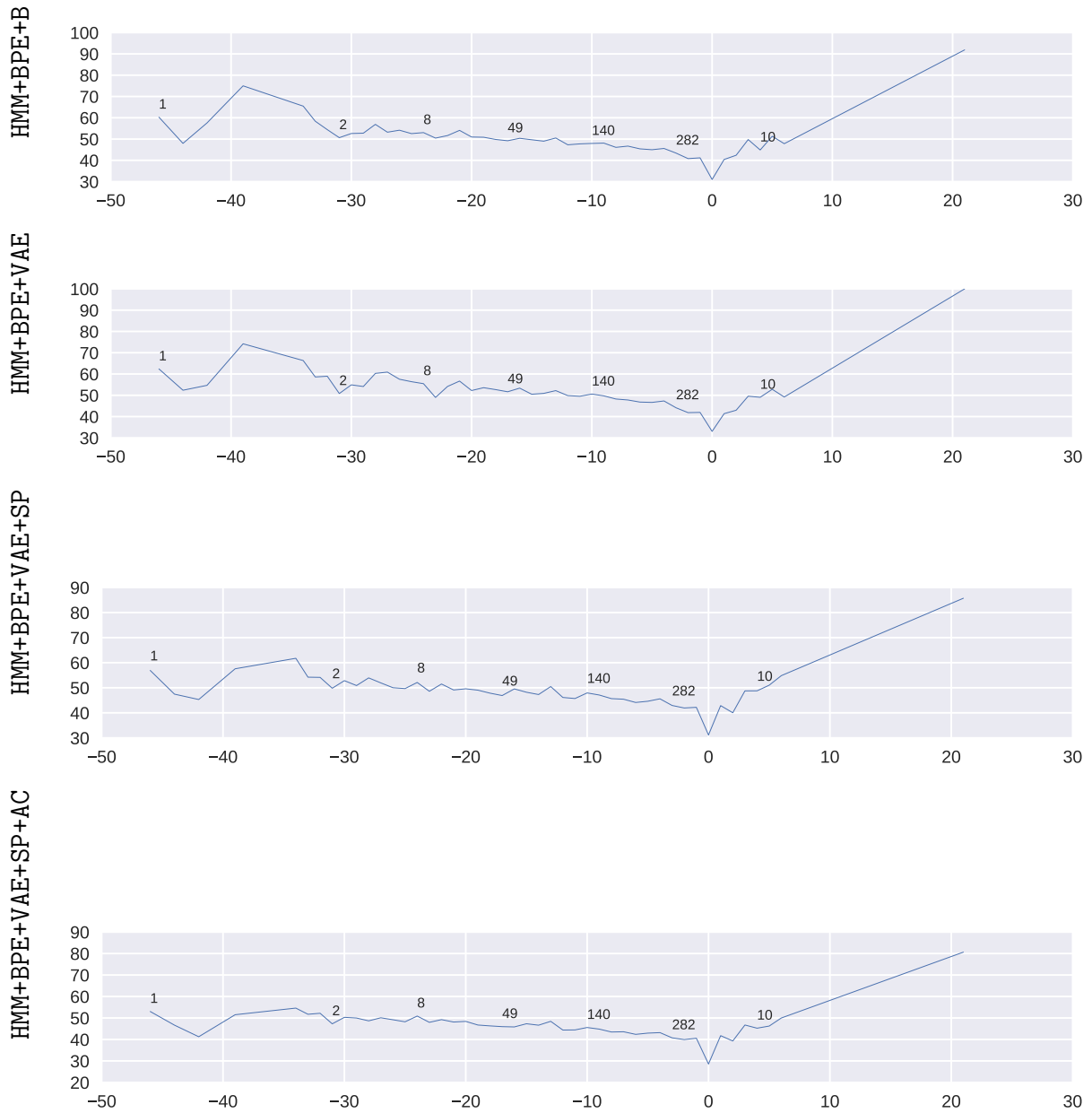
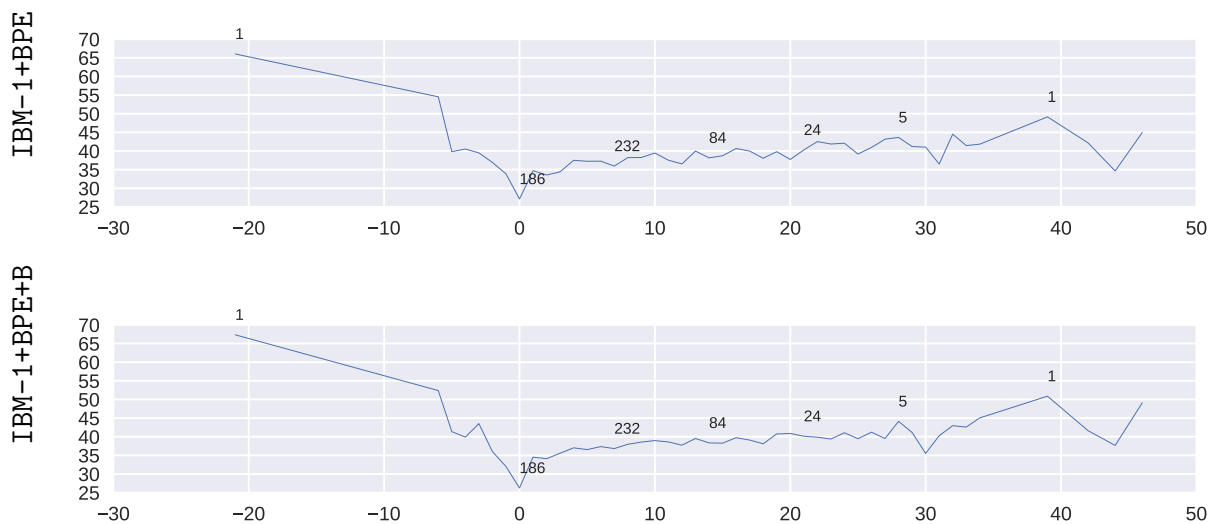
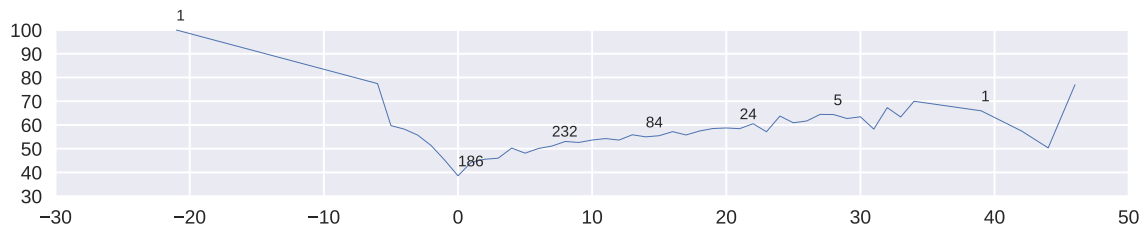


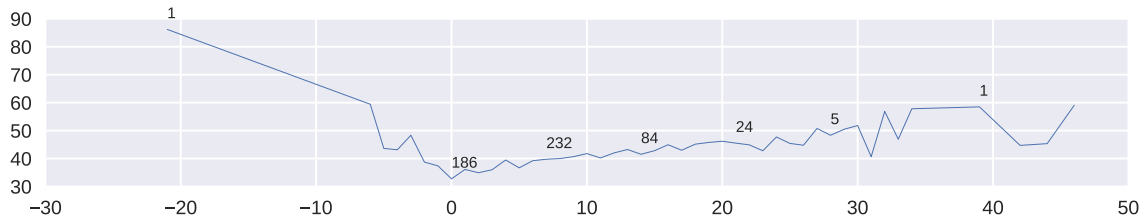
Figure D.77: Models for the direction English-Vietnamese: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.



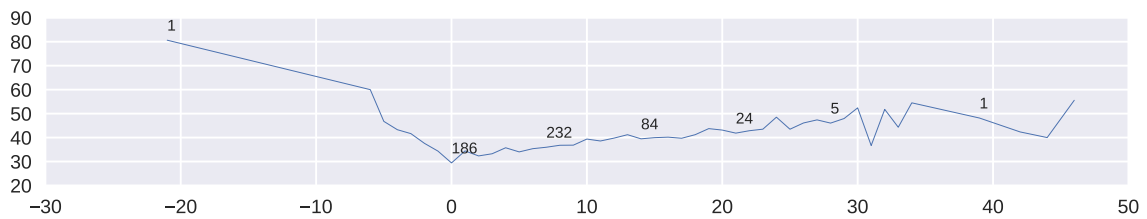
IBM-1+BPE+VAE



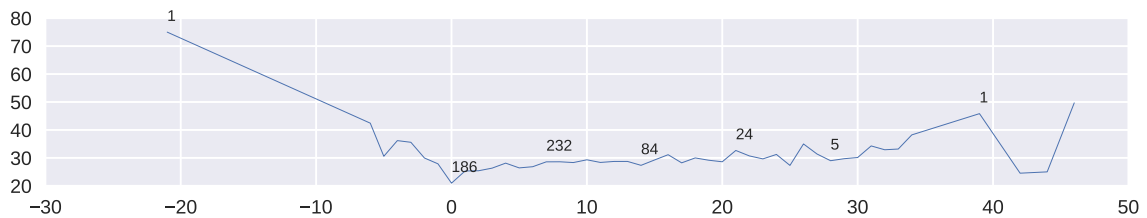
IBM-1+BPE+VAE+SP



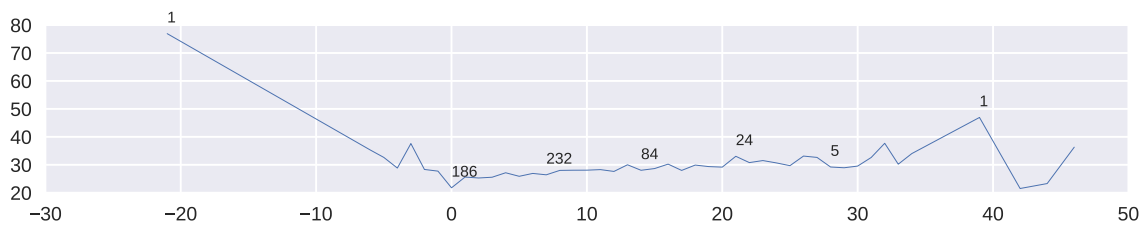
IBM-1+BPE+VAE+SP+AC



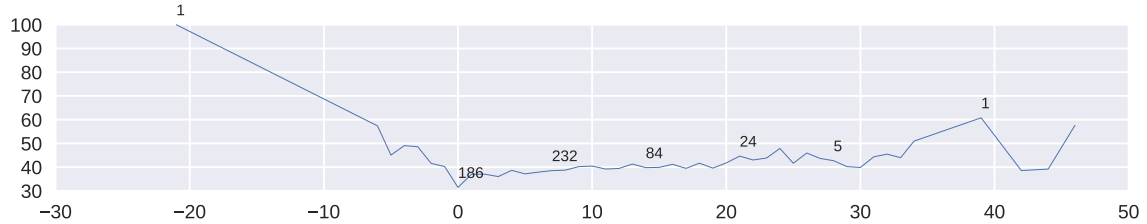
HMM+BPE



HMM+BPE+B



HMM+BPE+VAE



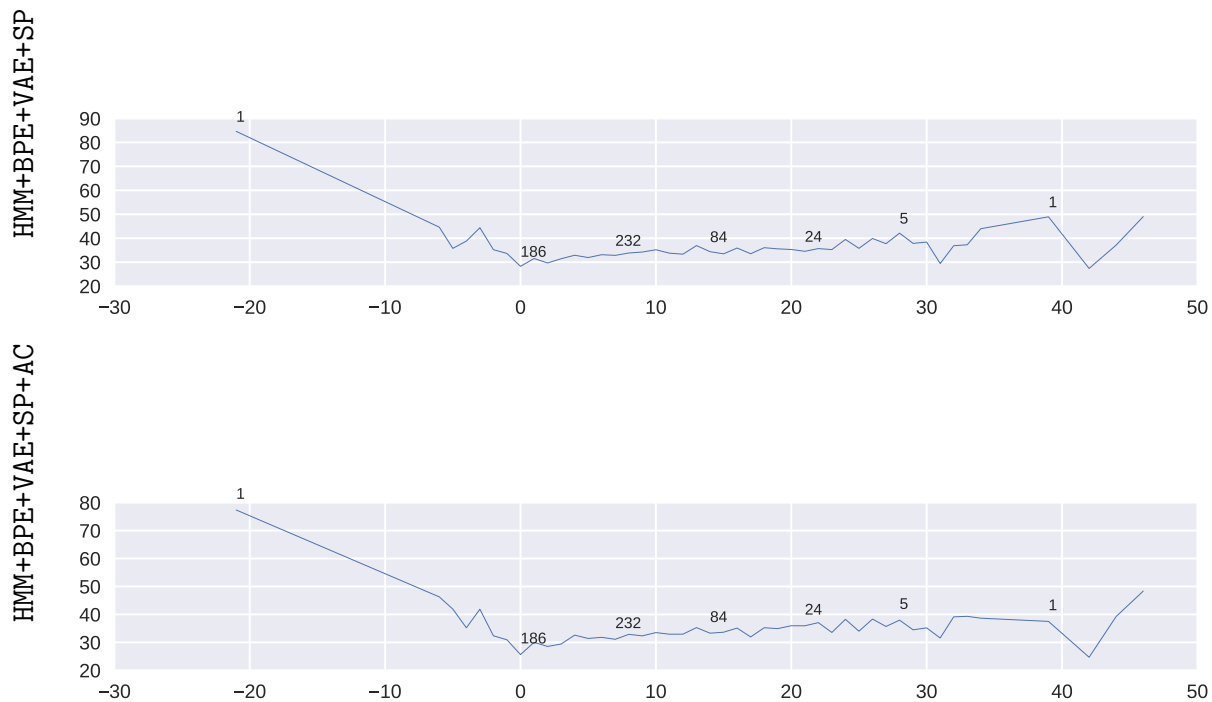


Figure D.78: Models for the direction Vietnamese-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The annotation displays the number of sentences.

Appendix E

Appendix: Subwords for word alignment task

Contents

E.1 Scores	405
E.1.1 Fastalign: Matrix of scores	405
E.1.2 Fastalign: Scores of BPE-based vs word-based alignment	409
E.1.3 Fastalign with BPE-dropout: Matrix of scores	422
E.1.4 Eflomal: Matrix of scores	426
E.1.5 Neural models with recommended BPE configurations	430
E.2 Correct and incorrect alignment links	433
E.3 Unaligned words	436
E.4 Alignment types	442
E.5 Rare words	457
E.5.1 Matrix of scores	457
E.5.2 Scores of BPE-based vs word-based alignment	464
E.6 Number of fragments as a function of word occurrence	479
E.7 Number of one-to-many/many-to-one as a function of word occurrence	485
E.8 Part-of-Speech	546
E.8.1 Alignment links for content and function words	546
E.8.2 Unaligned words for content and function words	552
E.9 Symmetrization	558
E.9.1 GDF before the transformation step	558
E.9.2 GDF after the transformation step	561
E.10 Scores as a function of sentence length	585
E.11 AER as a function of sentence length difference	651

E.1 Scores

E.1.1 Fastalign: Matrix of scores

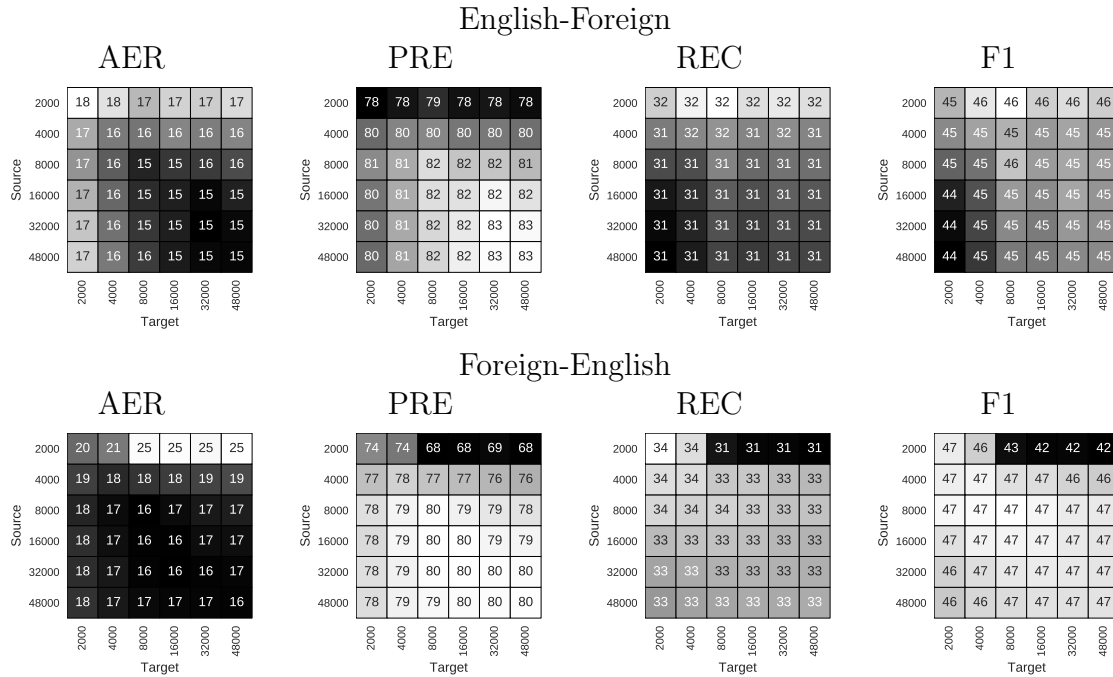


Figure E.1: BPE-based Fastalign for English-French: For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).

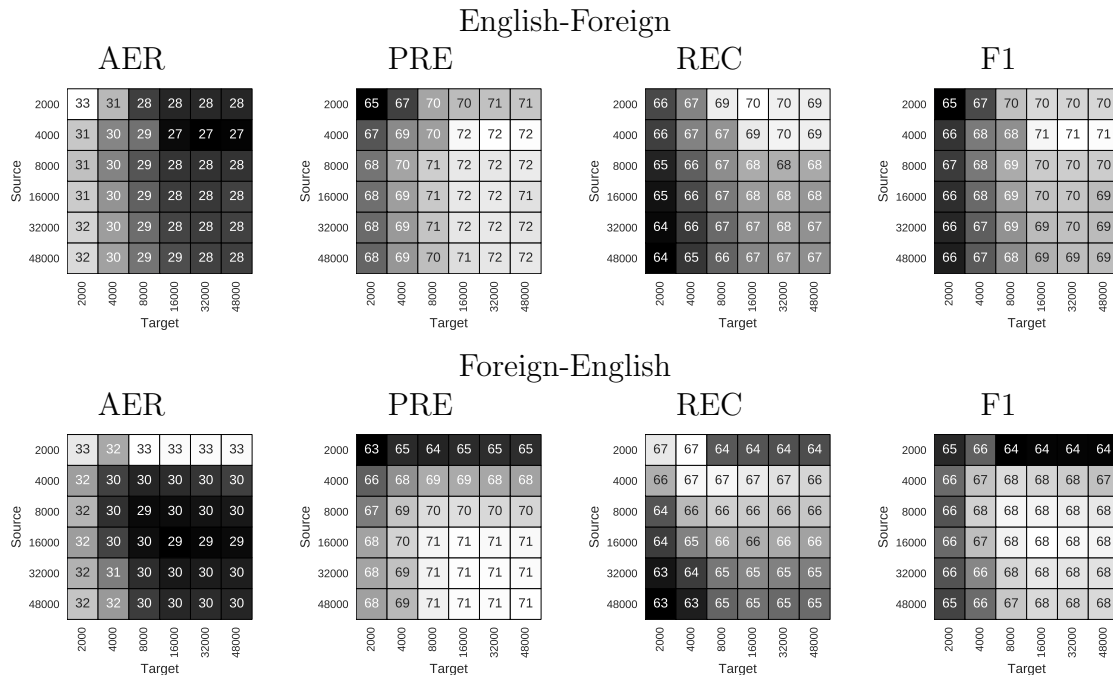


Figure E.2: BPE-based Fastalign for English-German: For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).

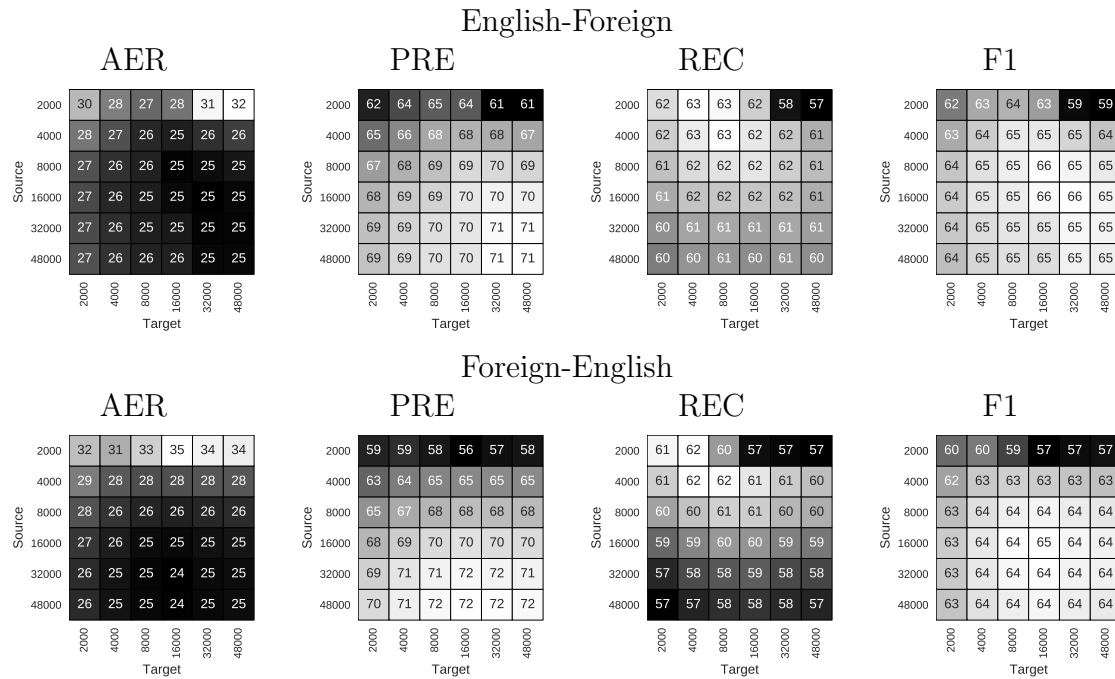


Figure E.3: BPE-based `Fastalign` for English-Czech: For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).

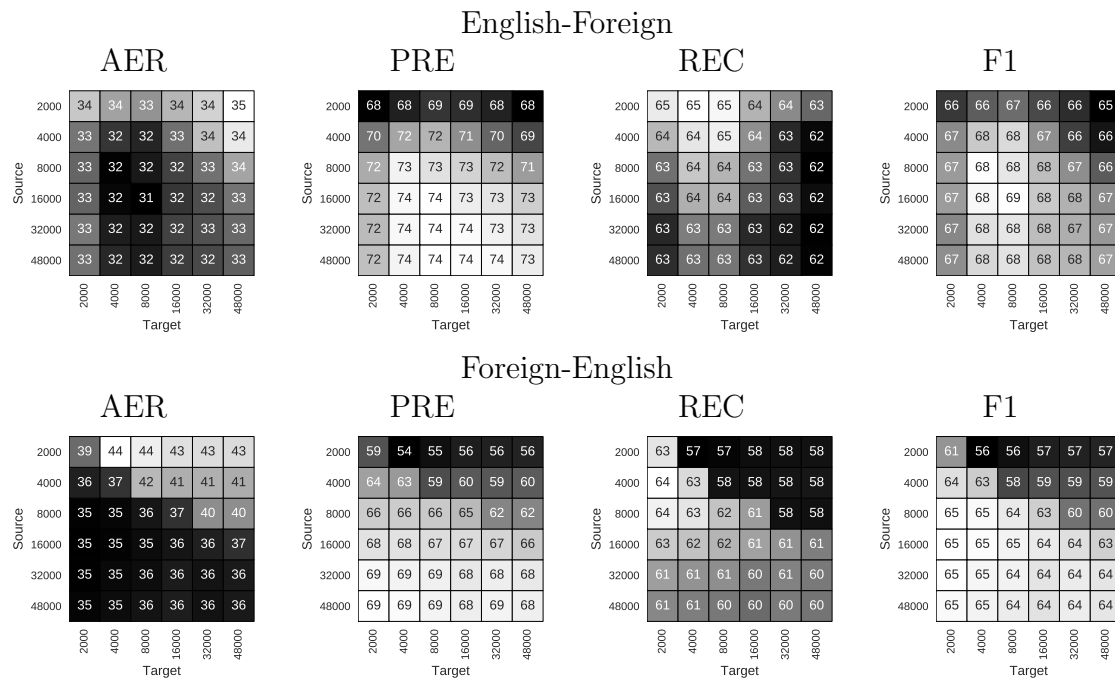


Figure E.4: BPE-based `Fastalign` for English-Romanian: For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).

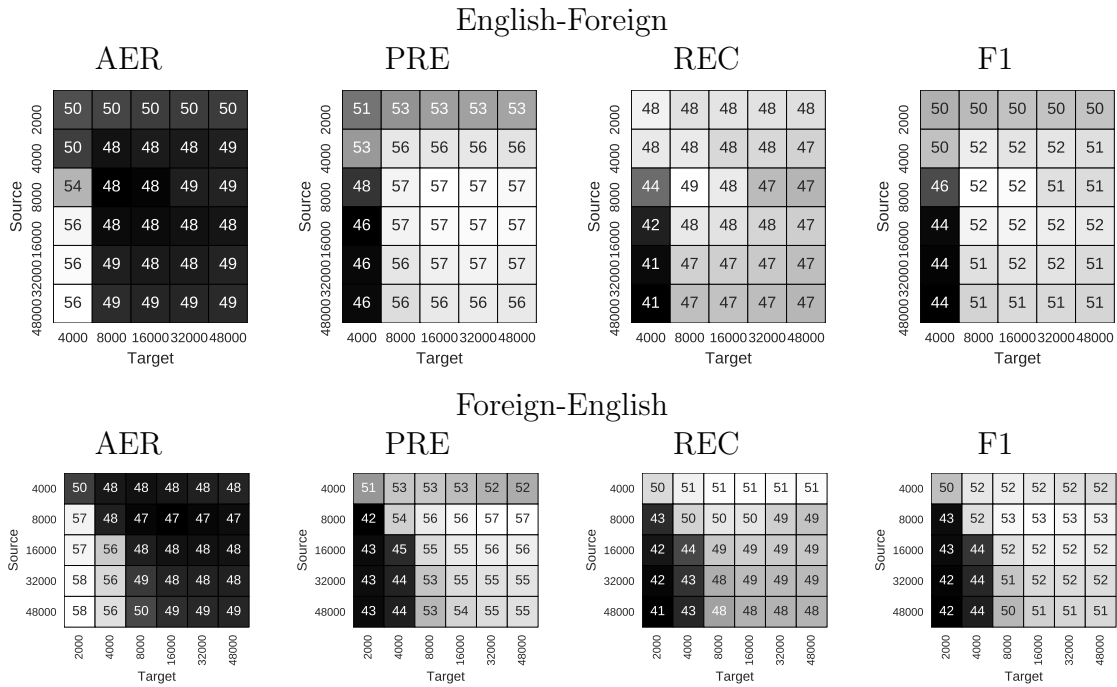


Figure E.5: BPE-based `Fastalign` for English-Japanese: For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).

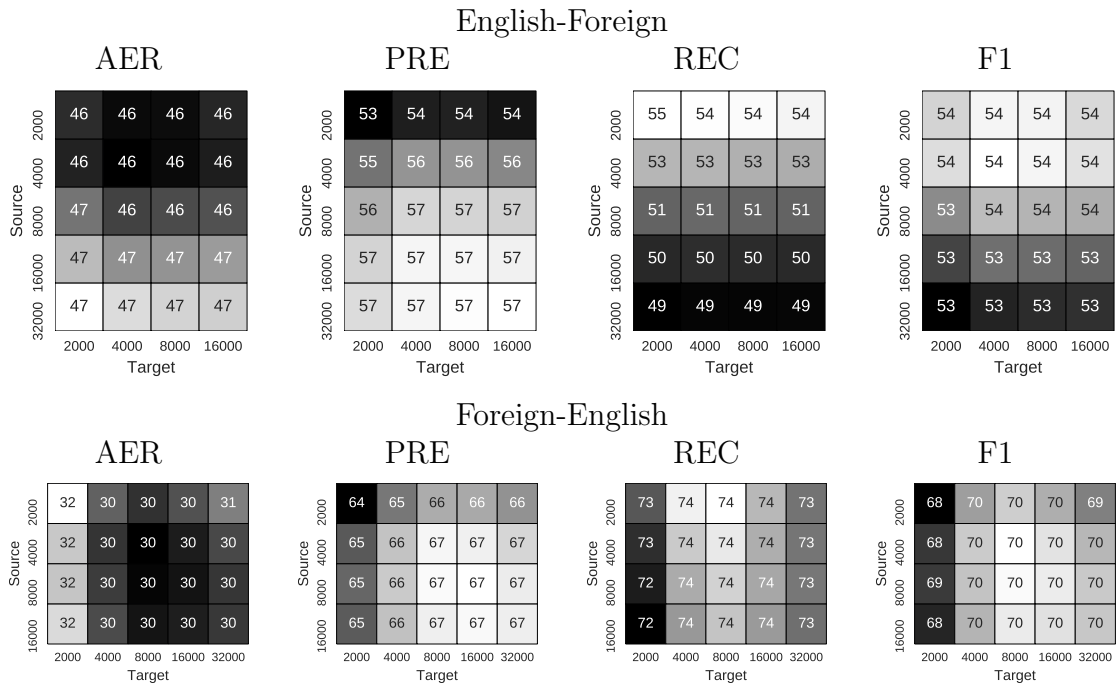


Figure E.6: BPE-based `Fastalign` for English-Vietnamese: For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).

E.1.2 Fastalign: Scores of BPE-based vs word-based alignment

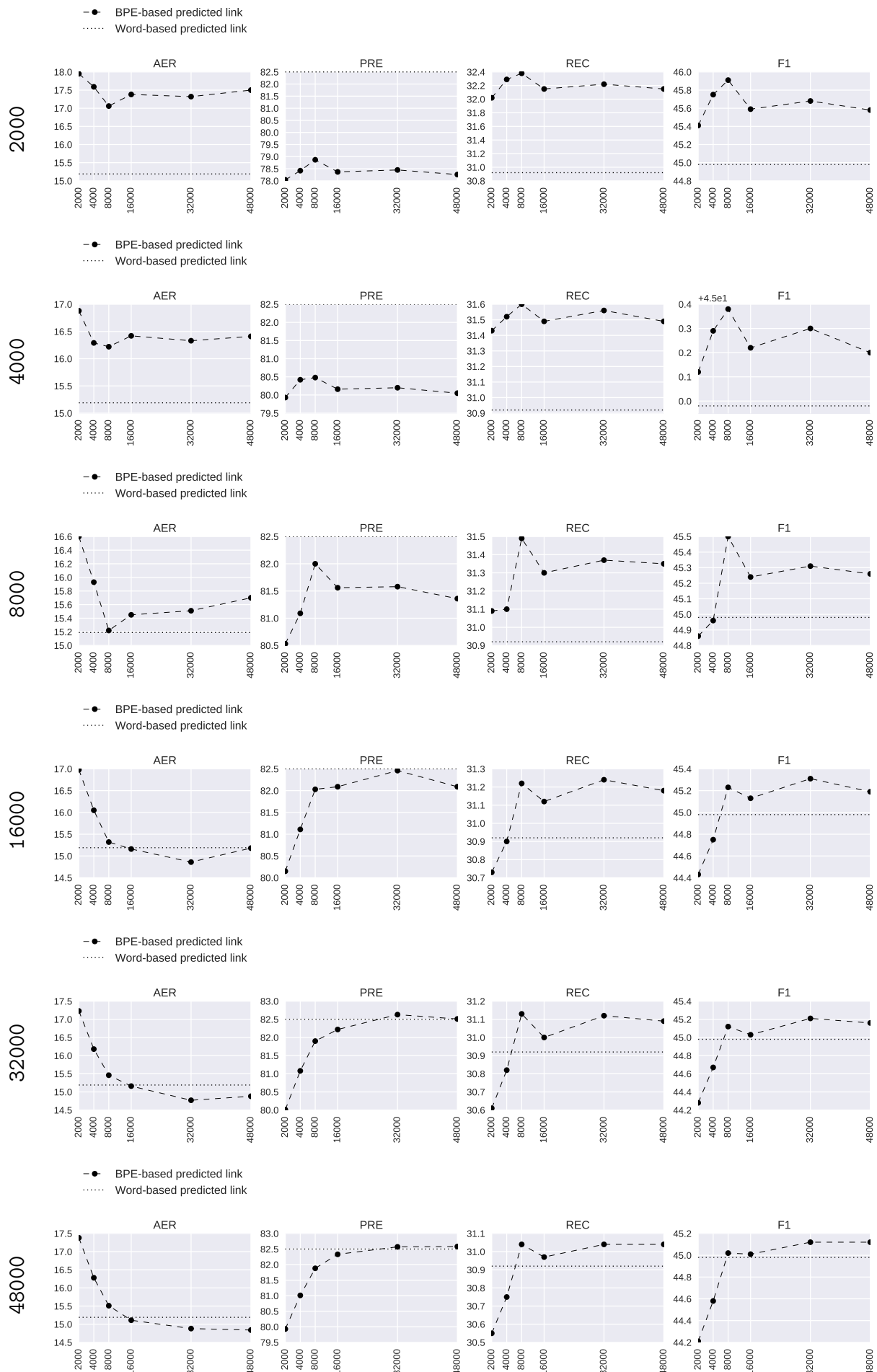
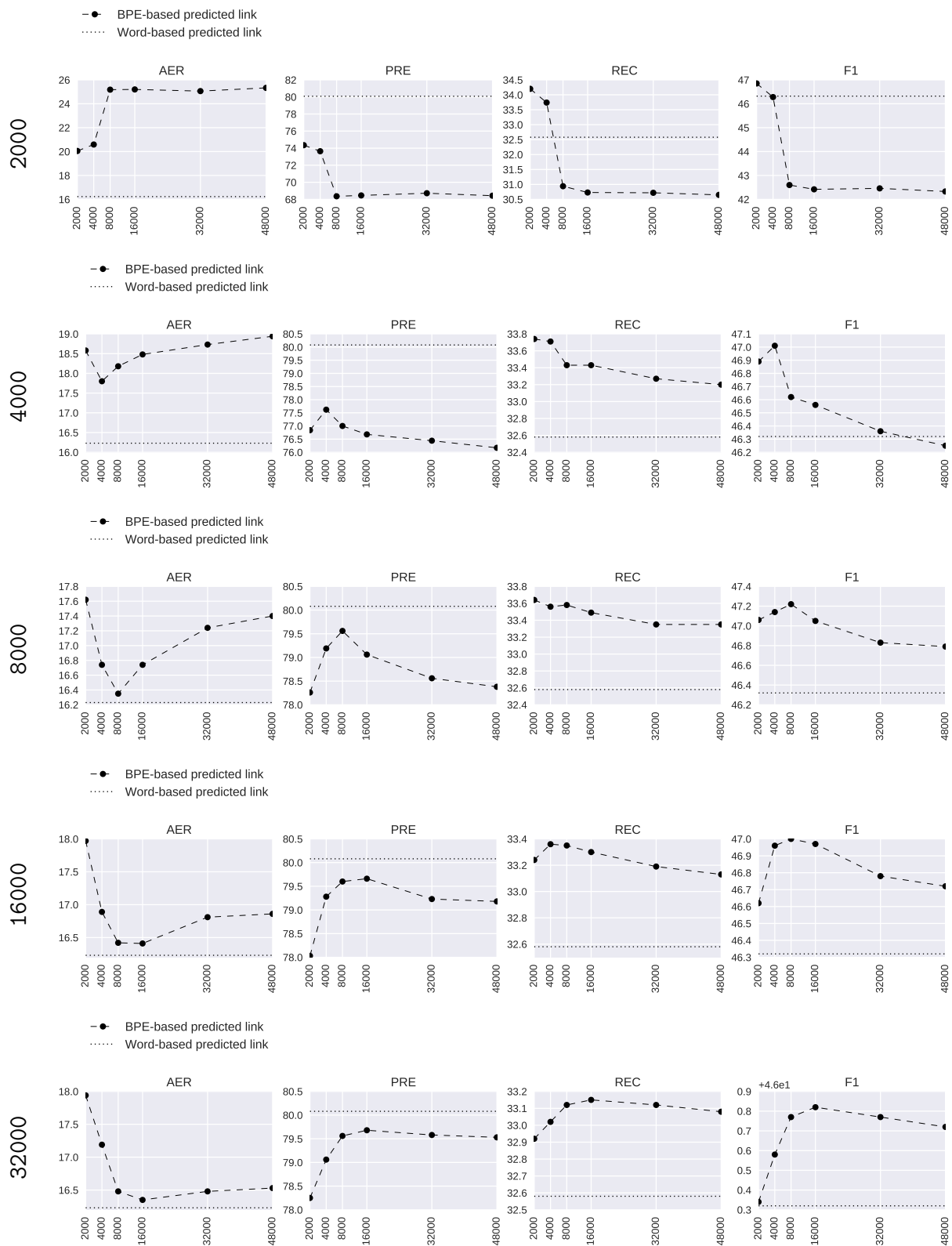


Figure E.7: BPE-based Fastalign for the direction English-French: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.



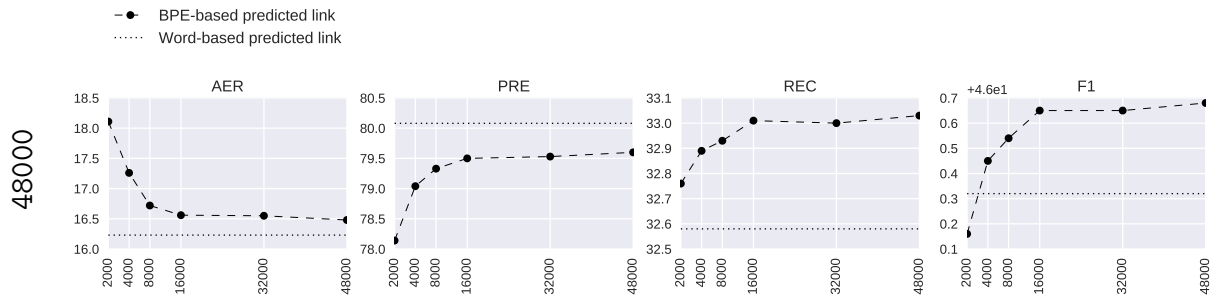
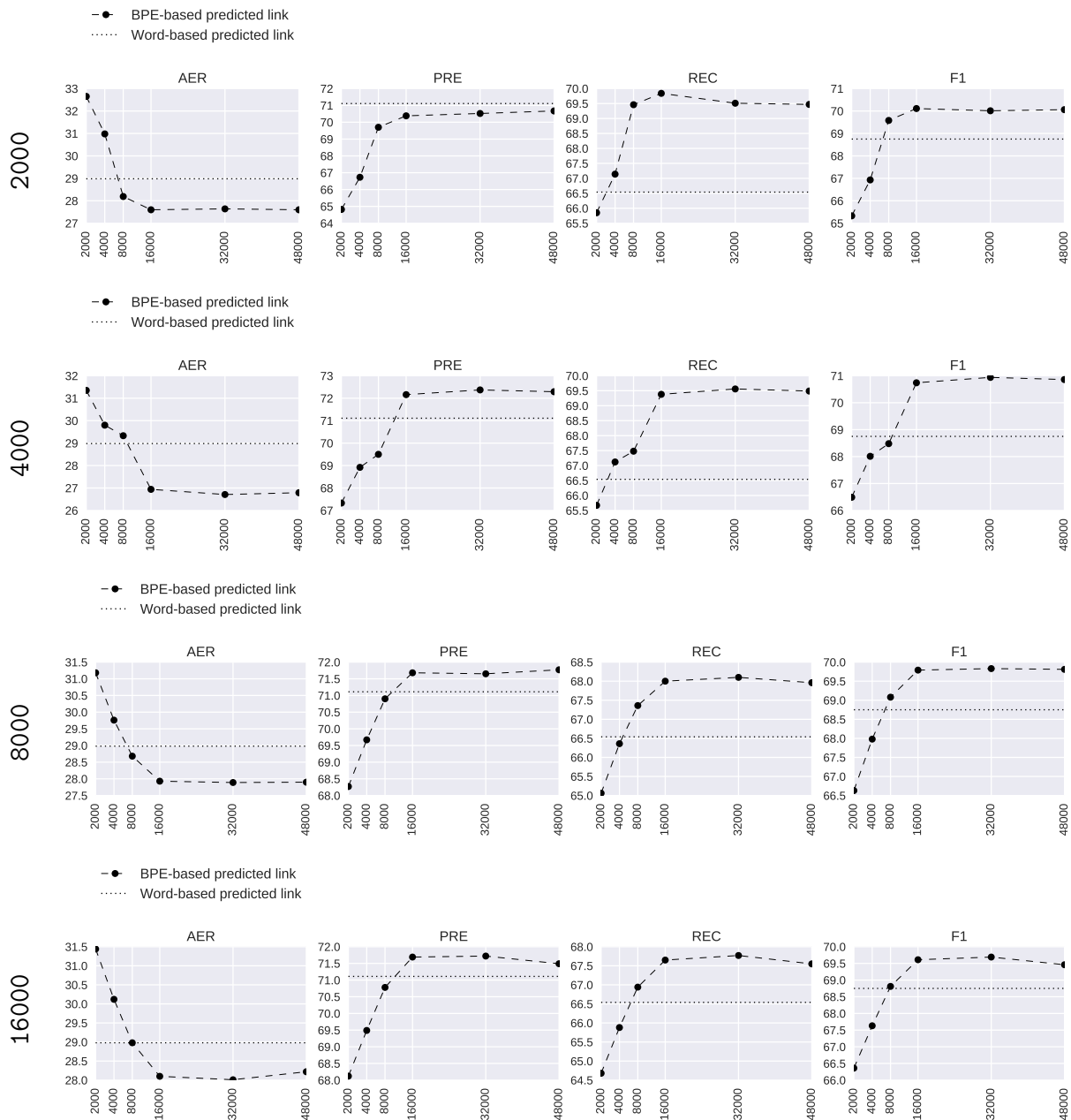


Figure E.8: BPE-based Fastalign for the direction French-English: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.



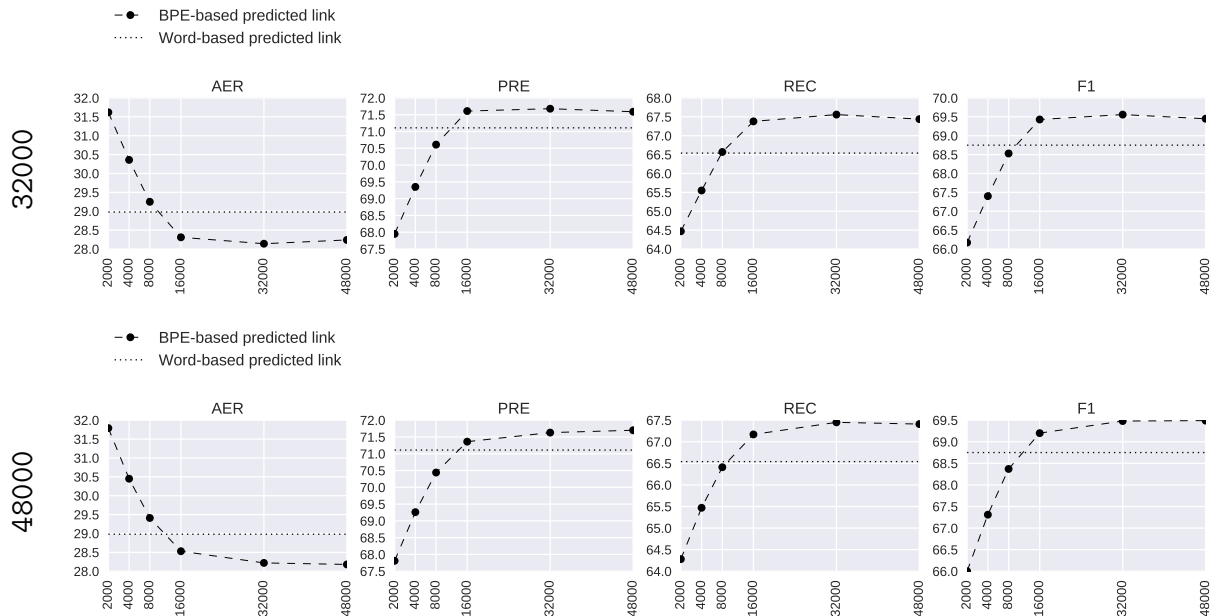
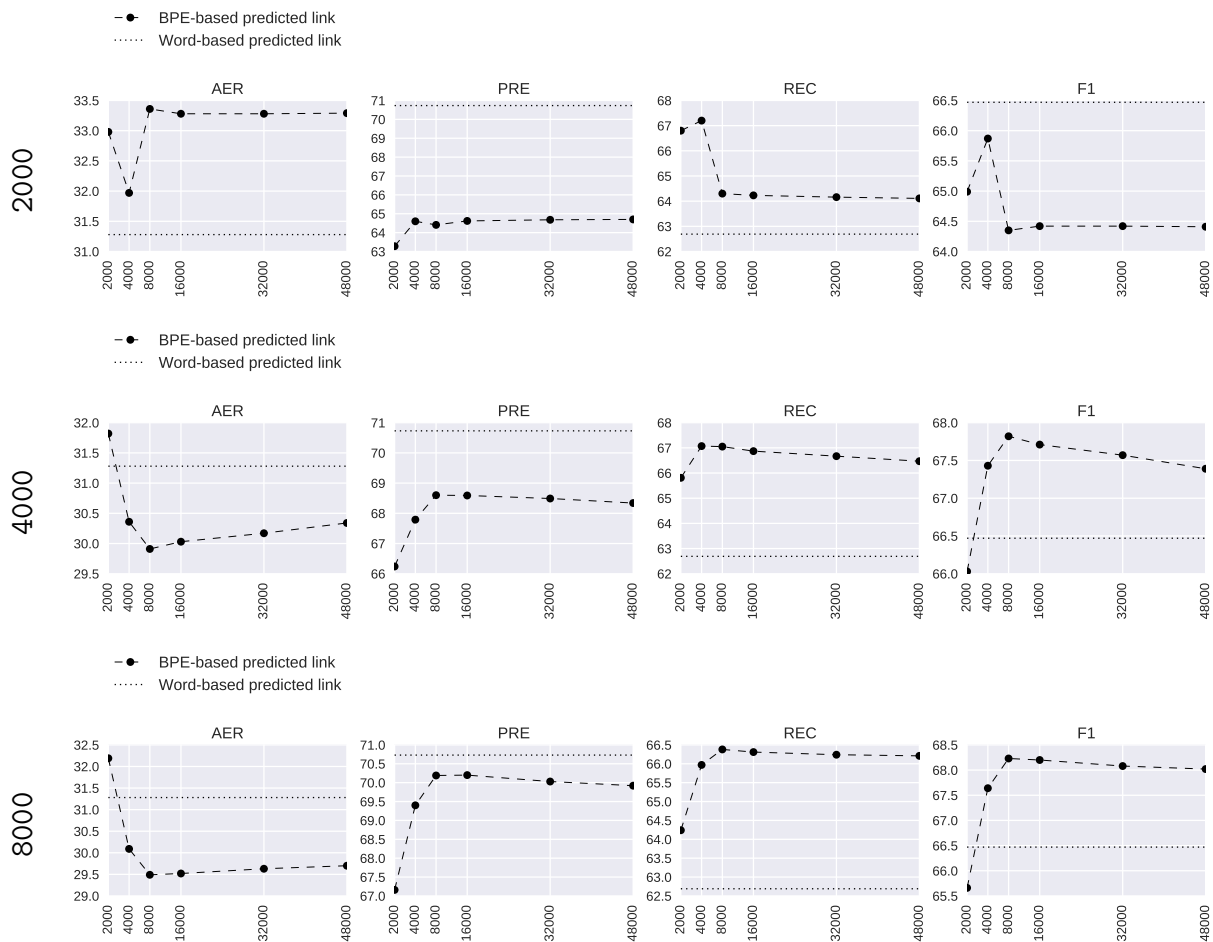


Figure E.9: BPE-based Fastalign for the direction English-German: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.



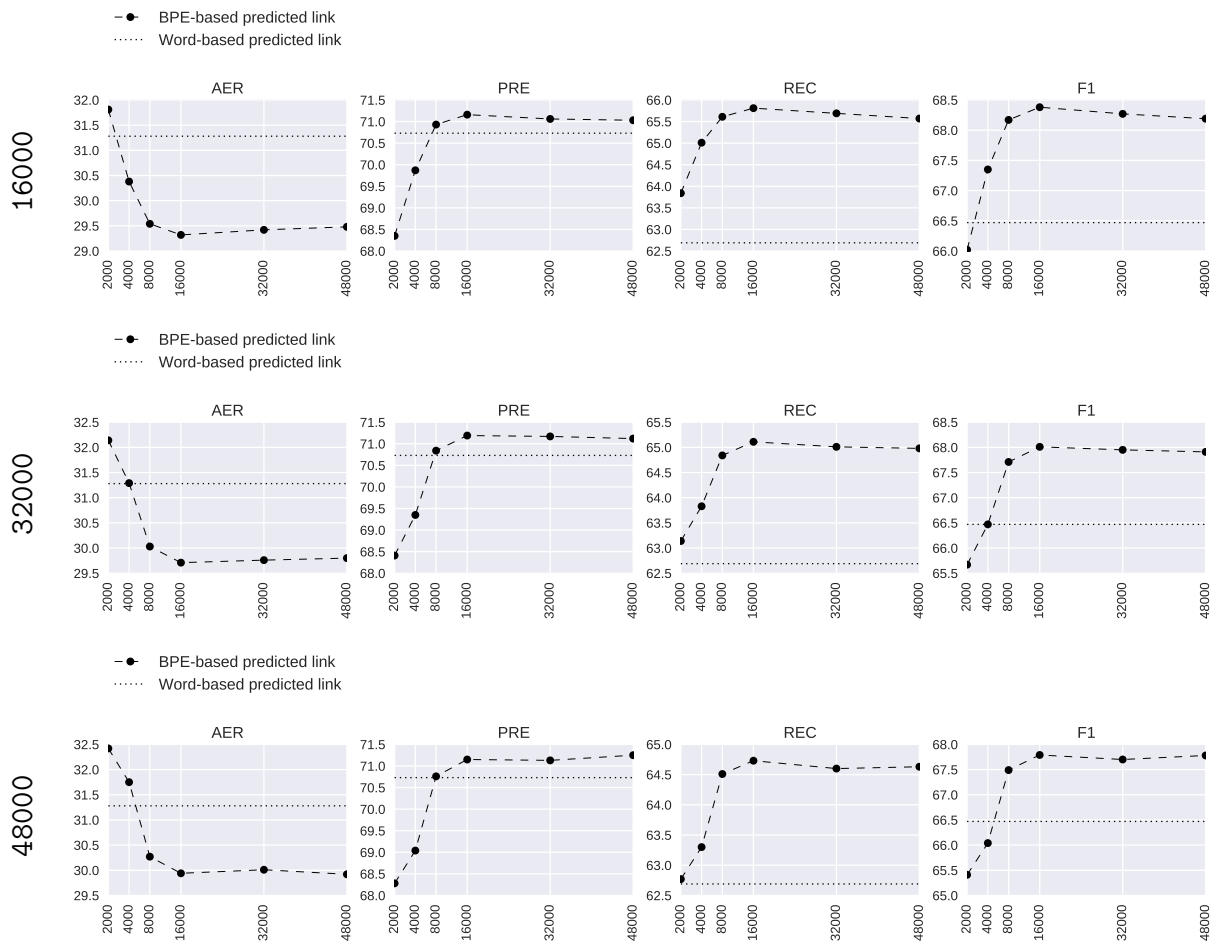
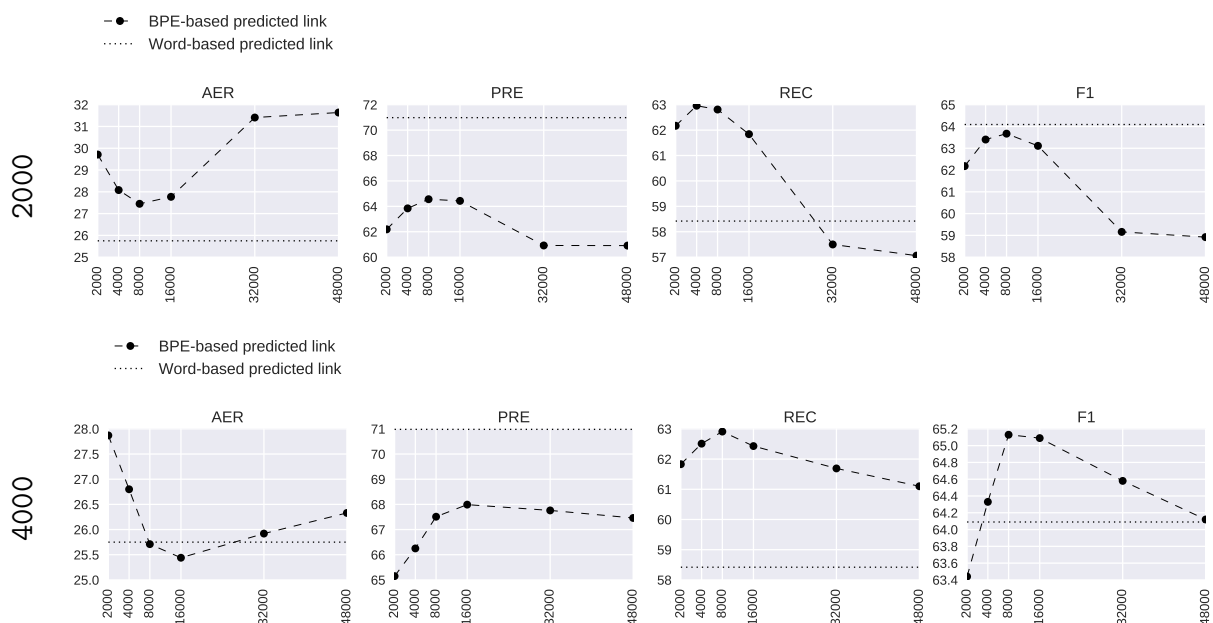


Figure E.10: BPE-based *Fastalign* for the direction German-English: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.



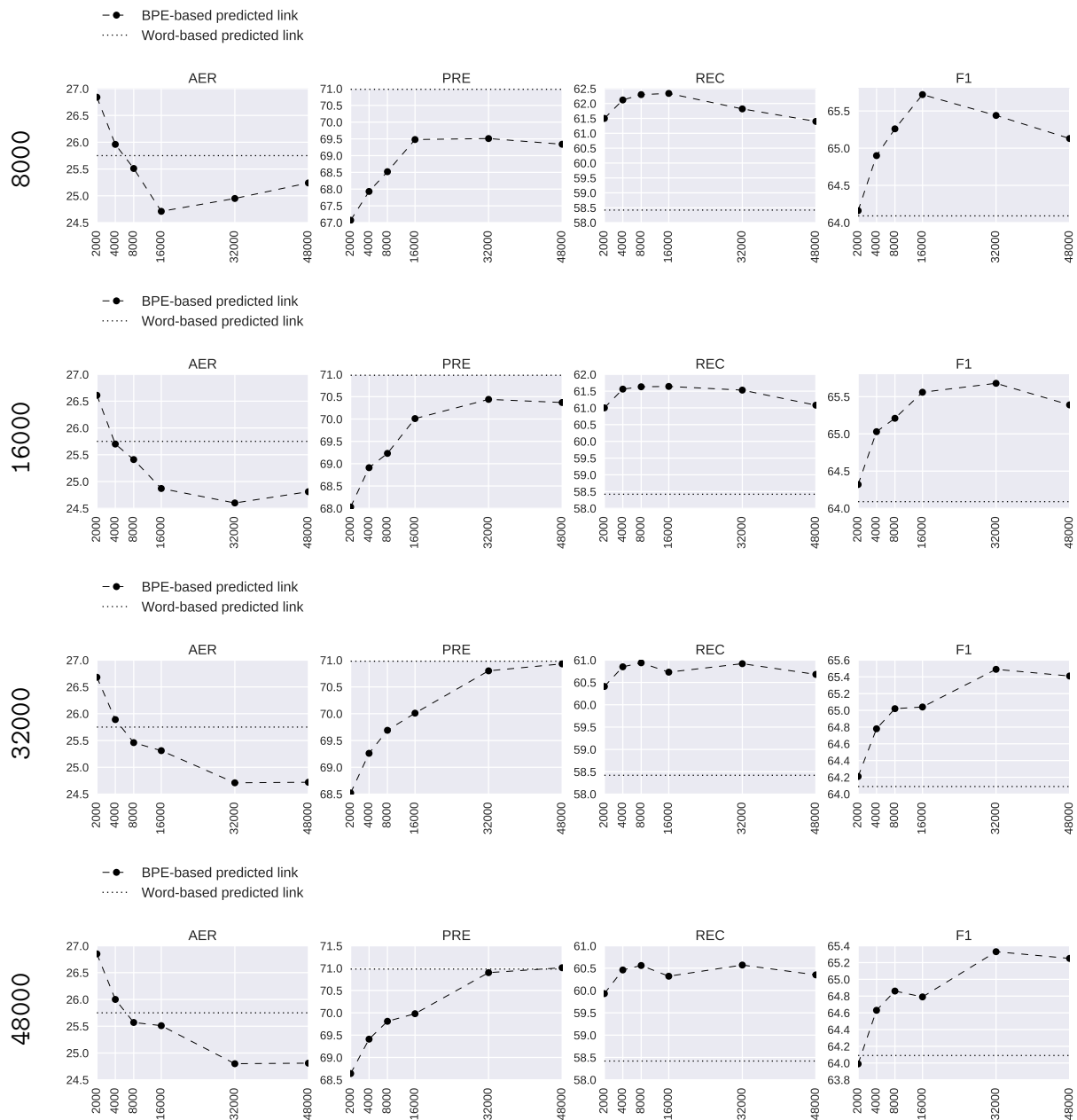
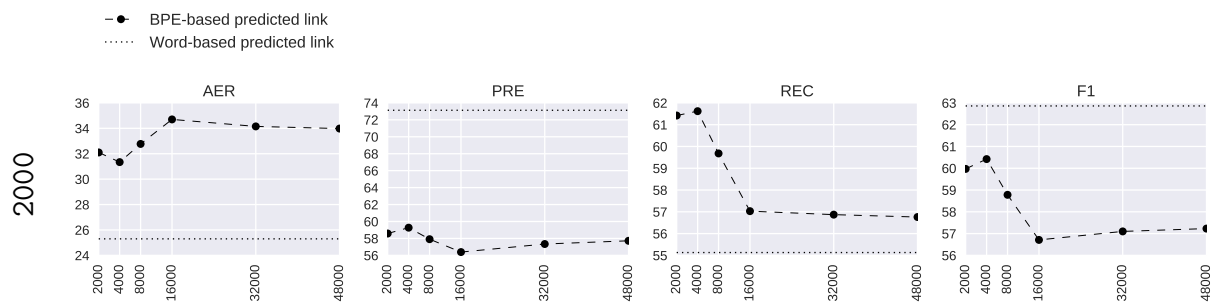


Figure E.11: BPE-based Fastalign for the direction English-Czech: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.



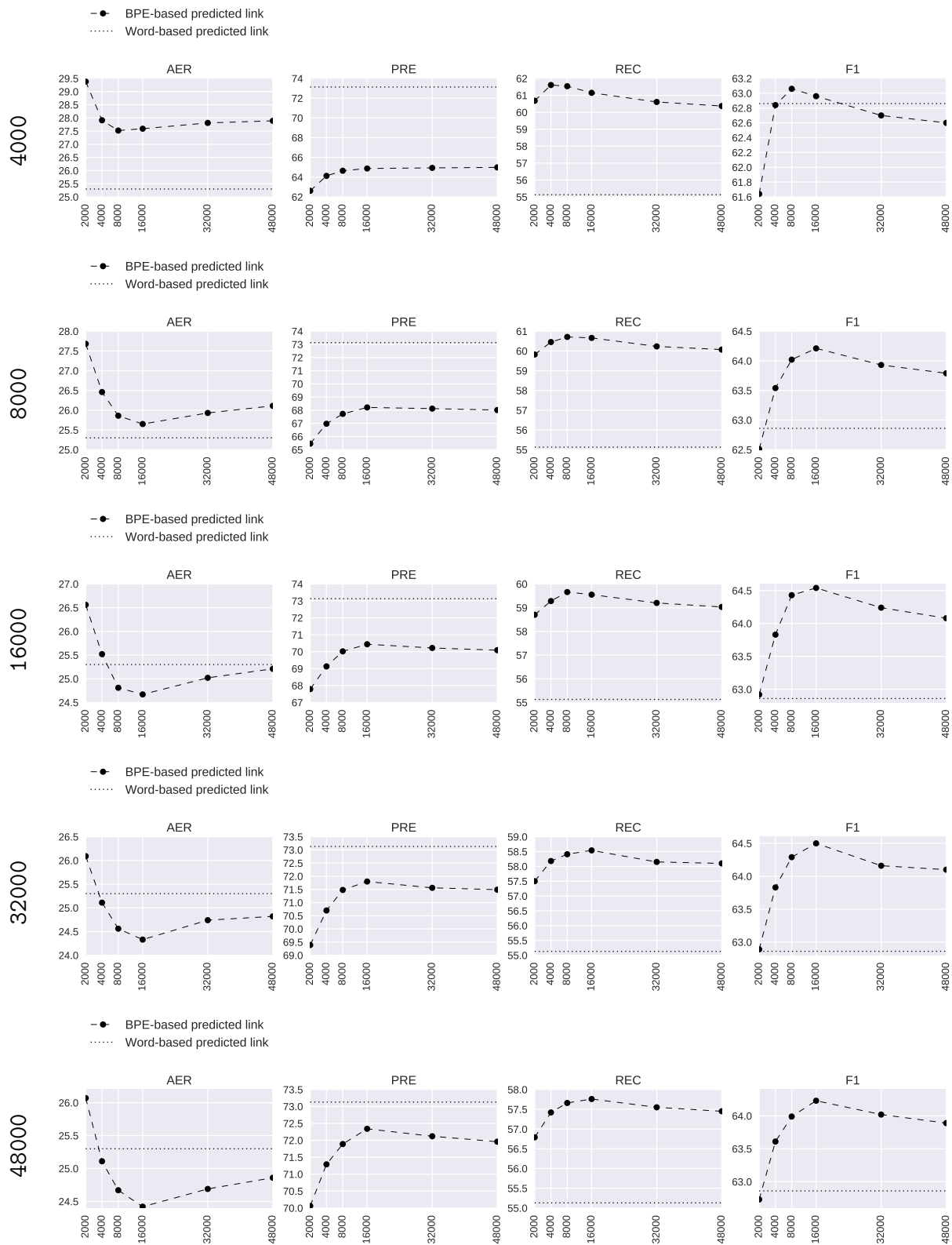


Figure E.12: BPE-based *Fastalign* for the direction Czech-English: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.

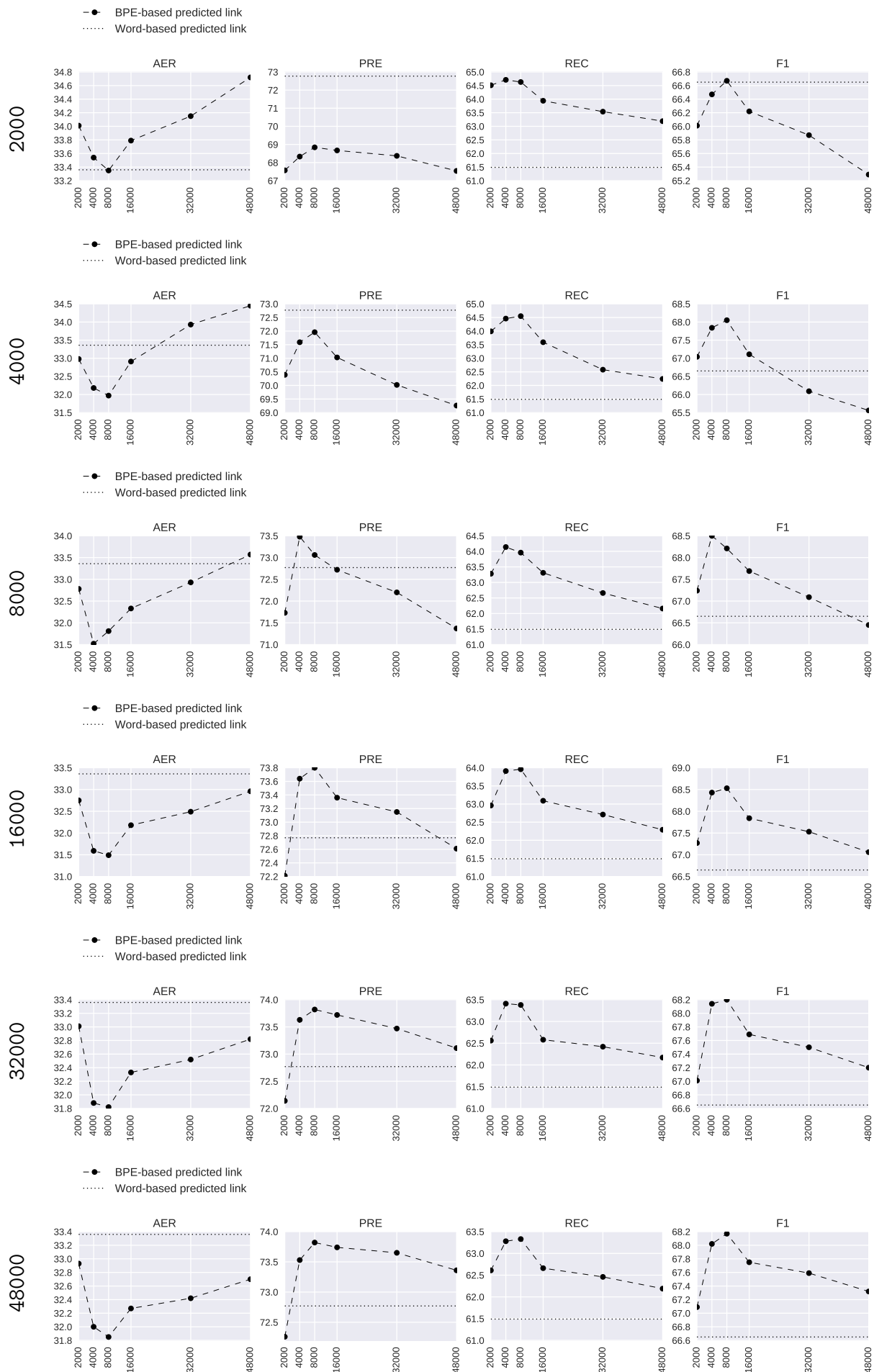
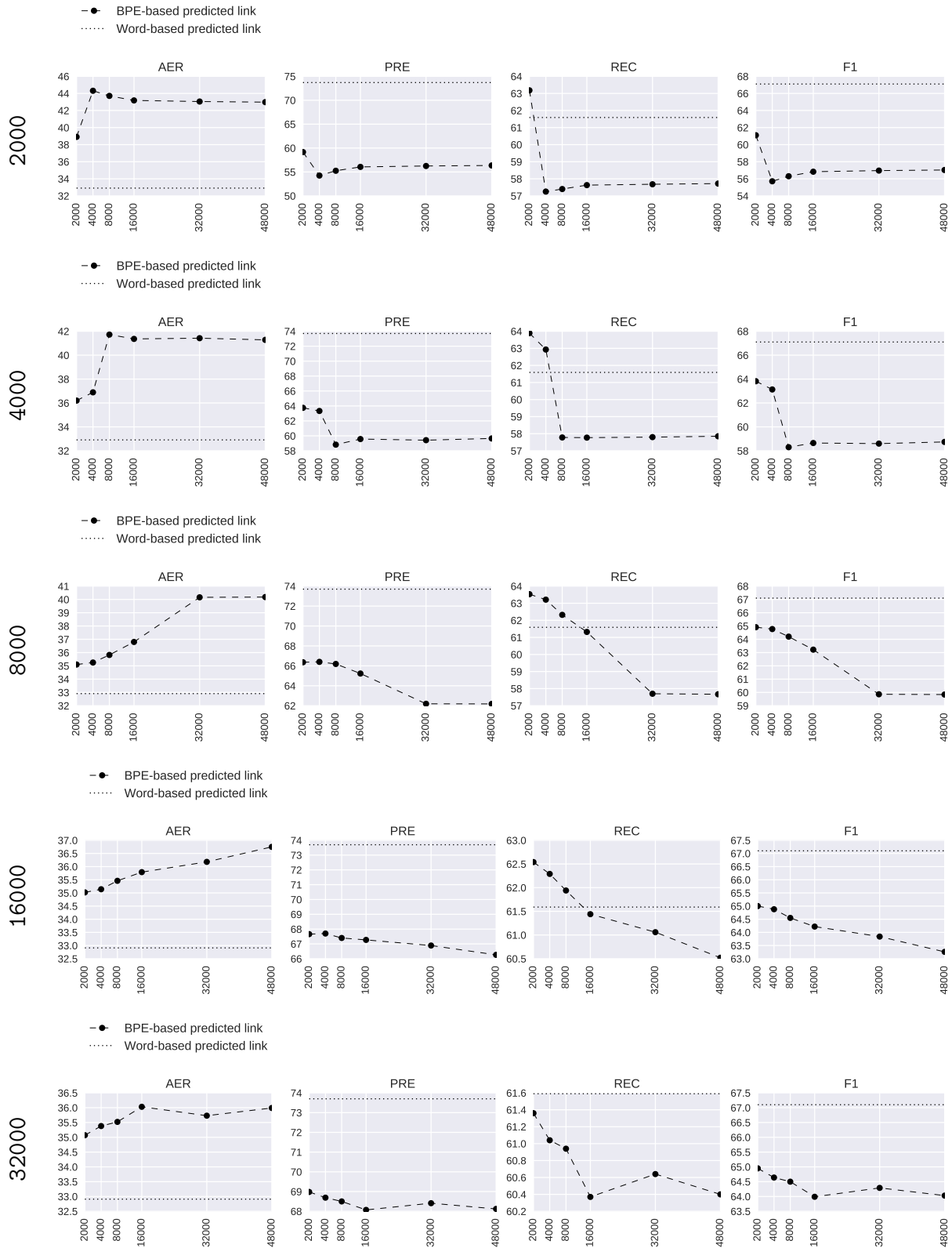


Figure E.13: BPE-based Fastalign for the direction English-Romanian: For each source vo-

cabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.



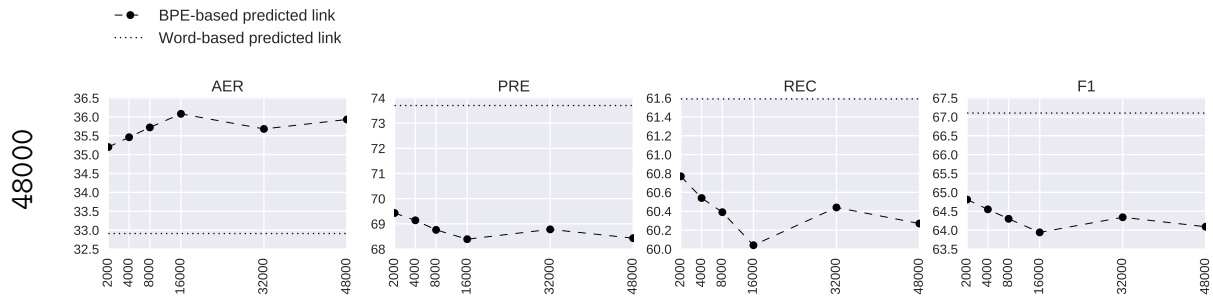
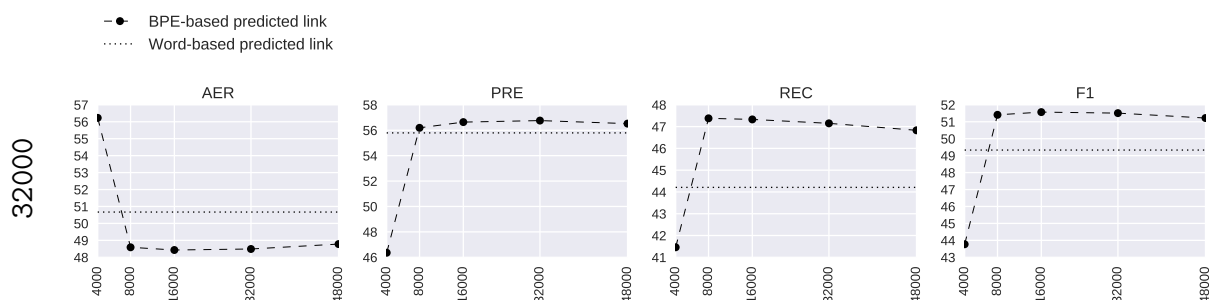
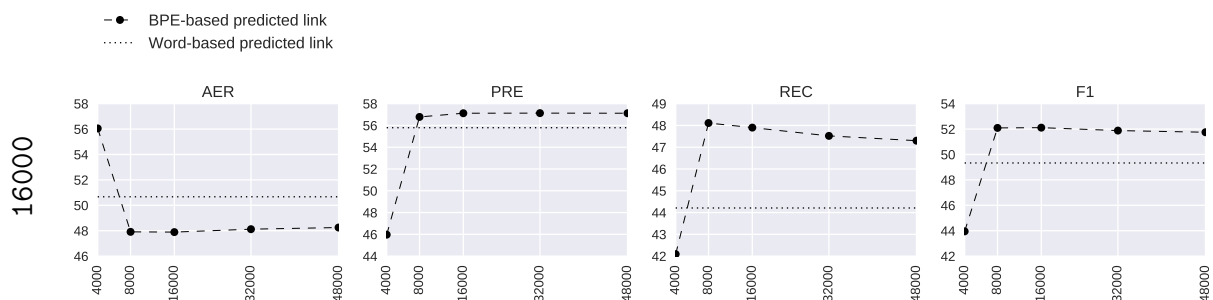
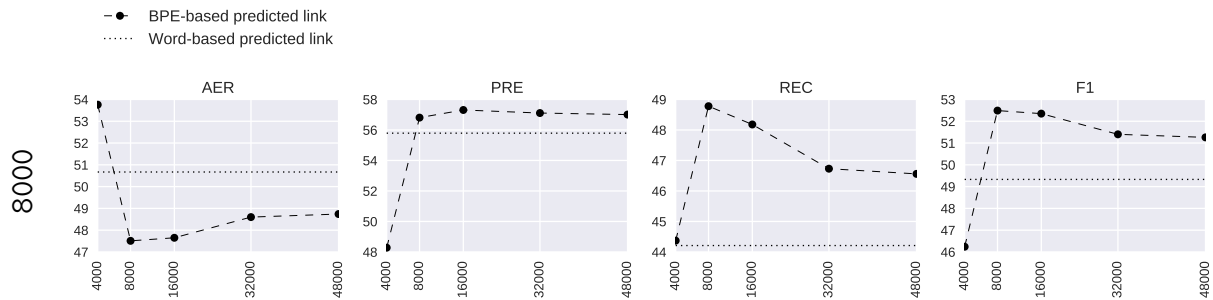
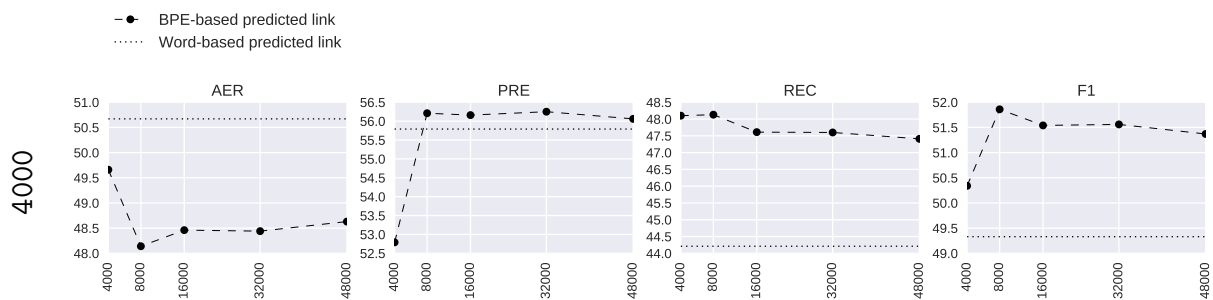


Figure E.14: BPE-based Fastalign for the direction Romanian-English: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.



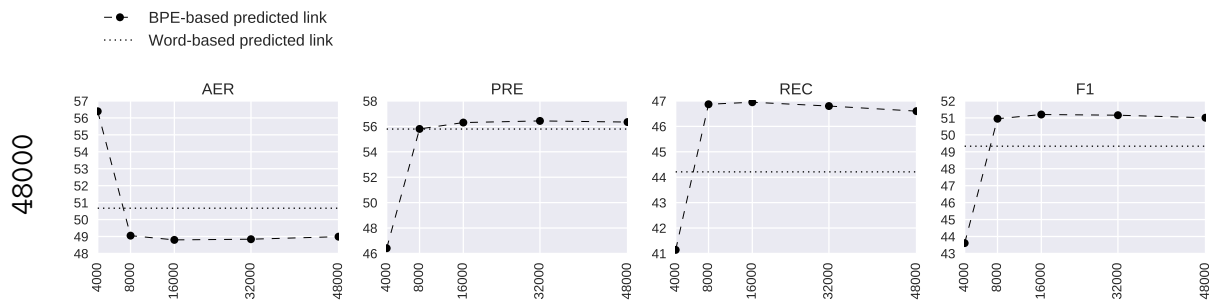
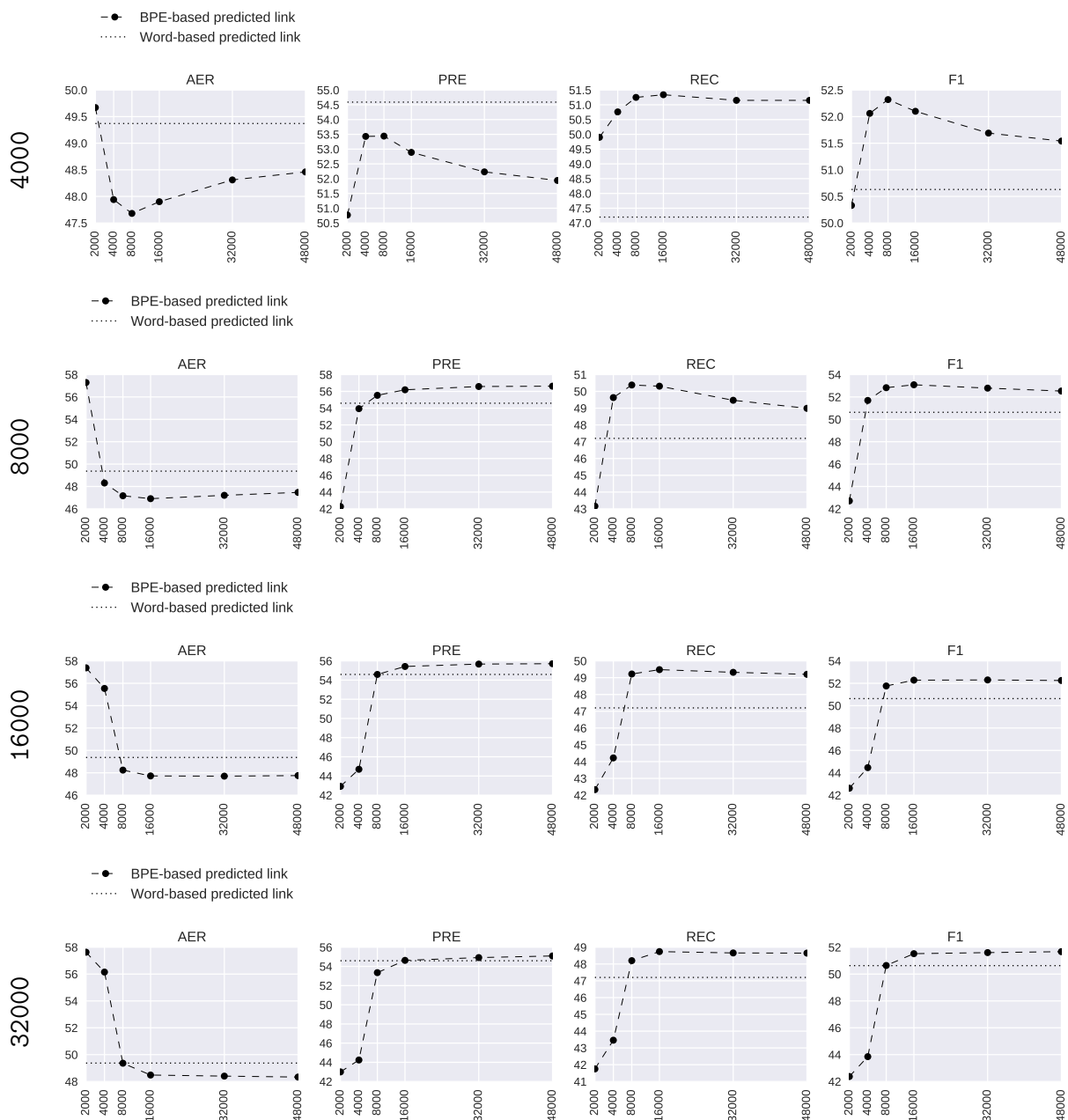


Figure E.15: BPE-based *Fastalign* for the direction English-Japanese: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.



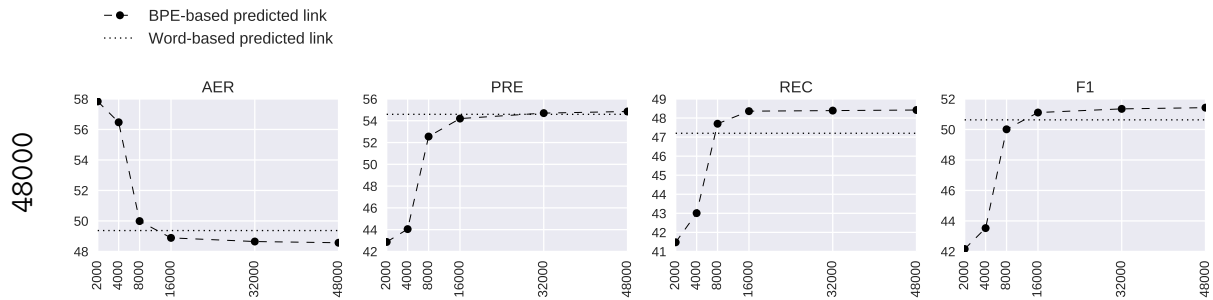
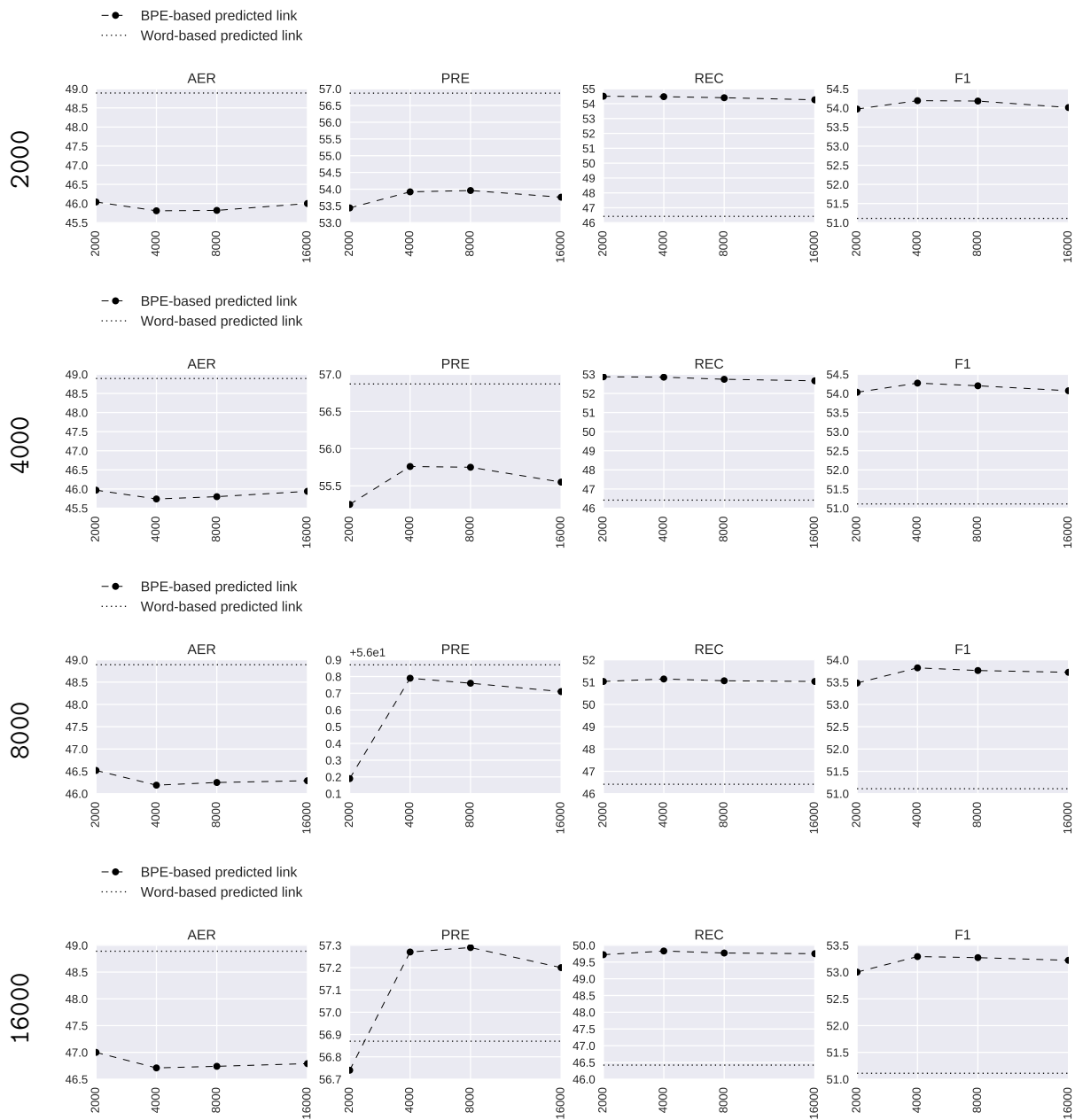


Figure E.16: BPE-based **Fastalign** for the direction Japanese-English: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.



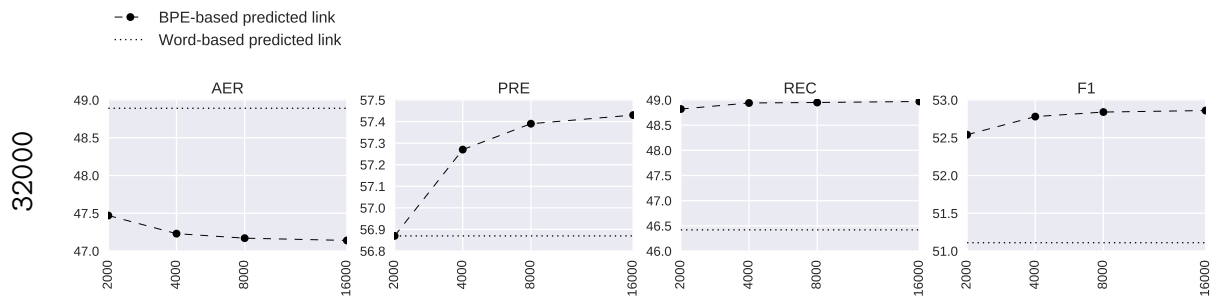


Figure E.17: BPE-based Fastalign for the direction English-Vietnamese: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.

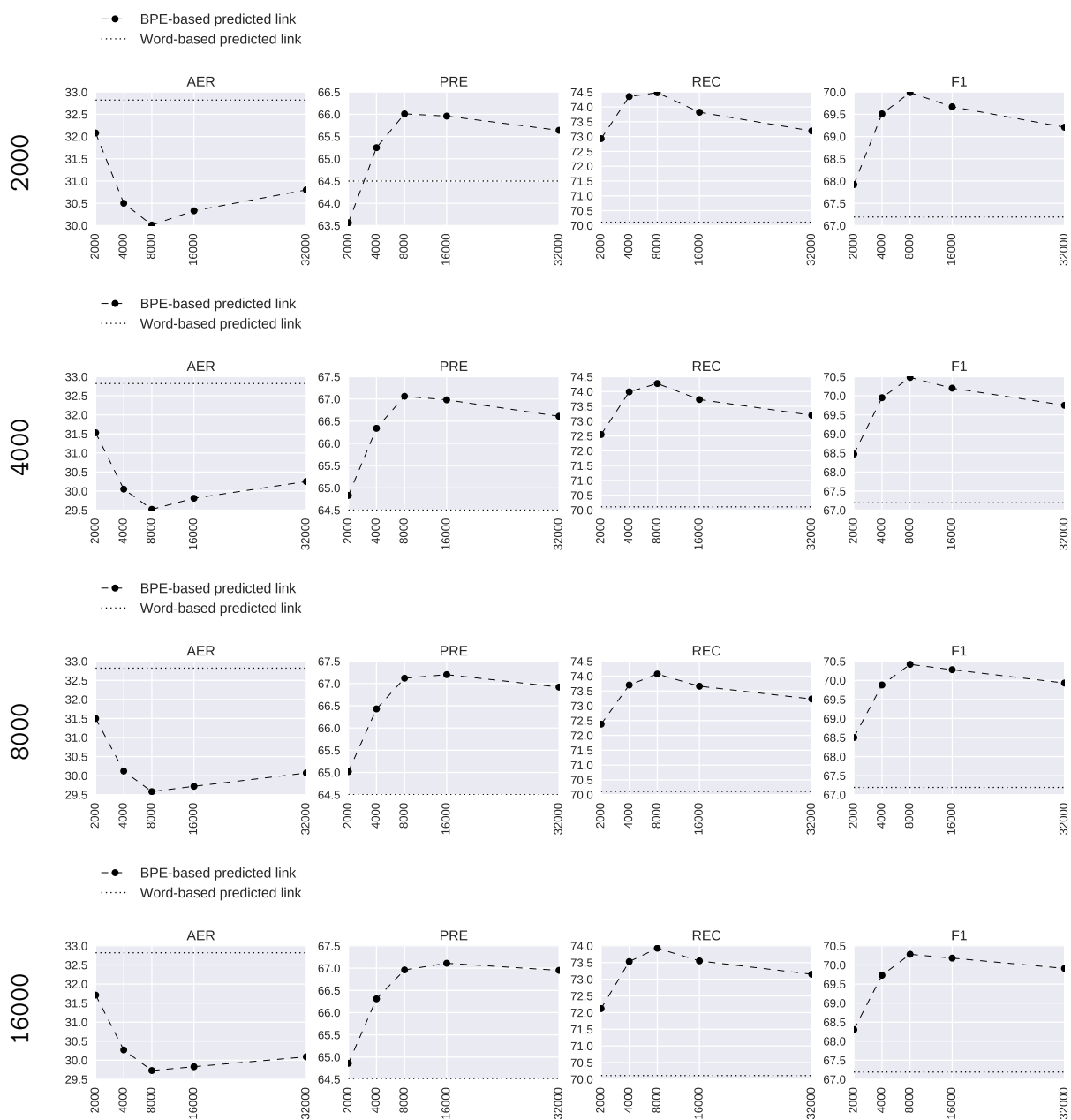


Figure E.18: BPE-based Fastalign for the direction Vietnamese-English: For each source vocabulary size, we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.

recall (REC) as a function of the target vocabulary size.

E.1.3 Fastalign with BPE-dropout: Matrix of scores

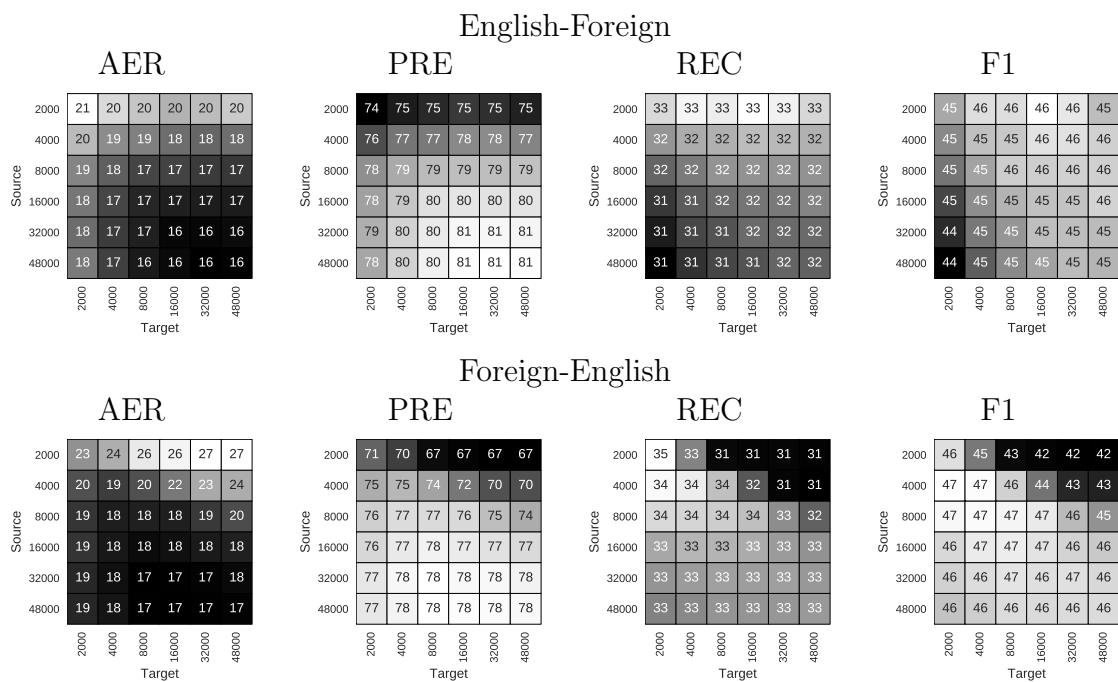


Figure E.19: BPE-based Fastalign with BPE-dropout for English-French: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).

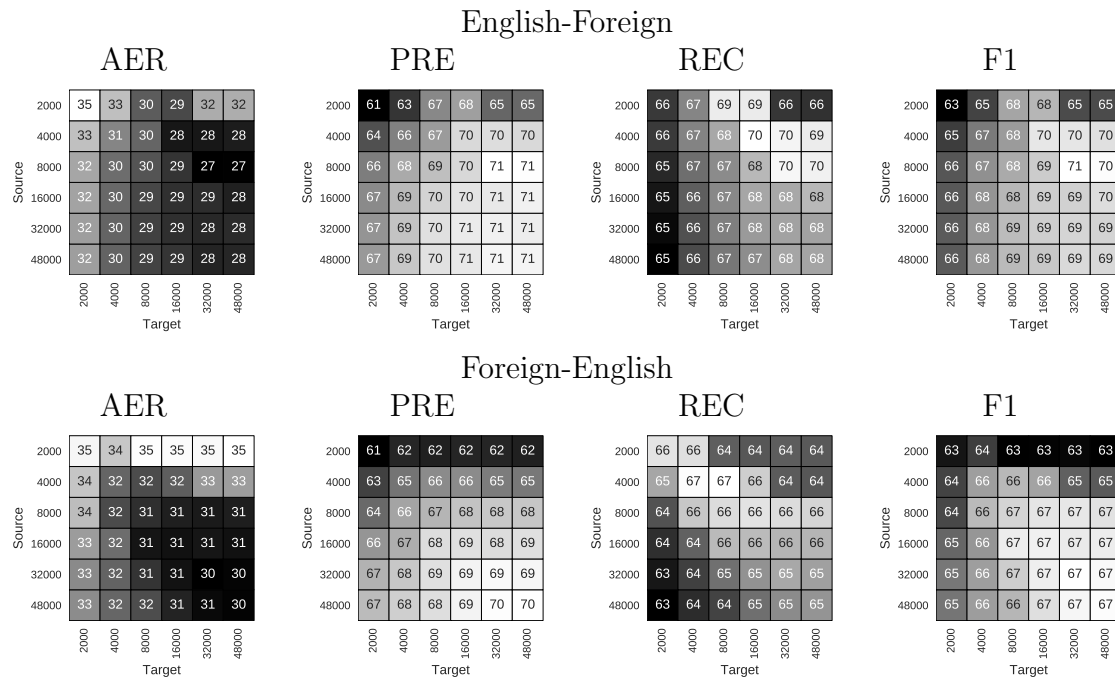


Figure E.20: BPE-based Fastalign with BPE-dropout for English-German: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).

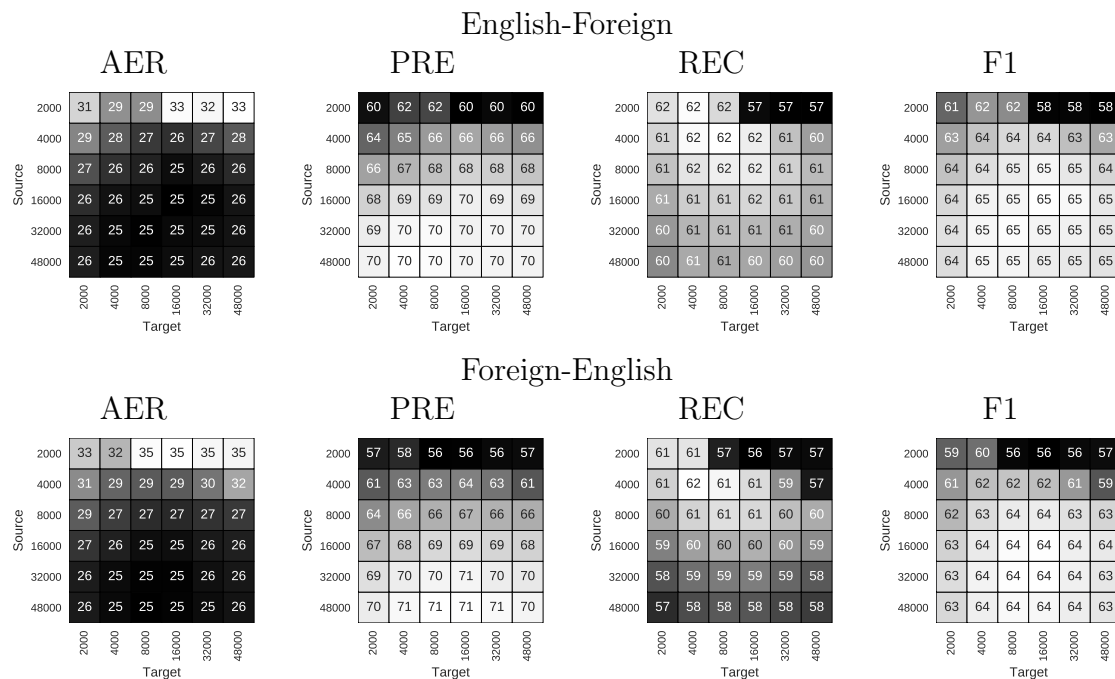


Figure E.21: BPE-based Fastalign with BPE-dropout for English-Czech: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).

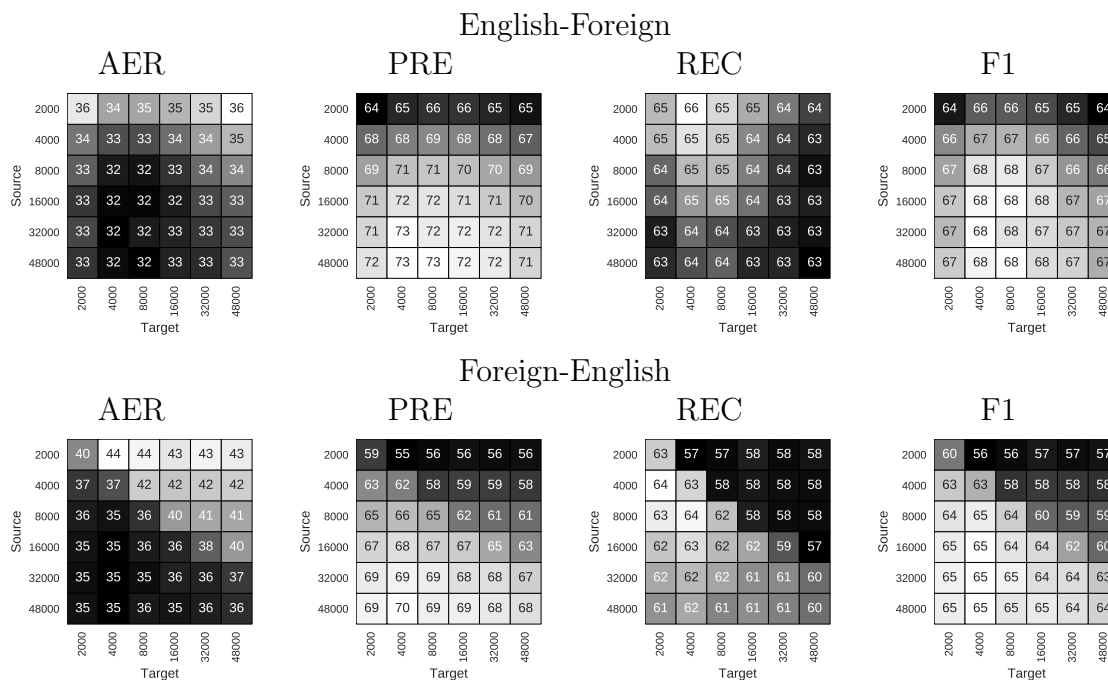


Figure E.22: BPE-based `Fastalign` with BPE-dropout for English-Romanian: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).

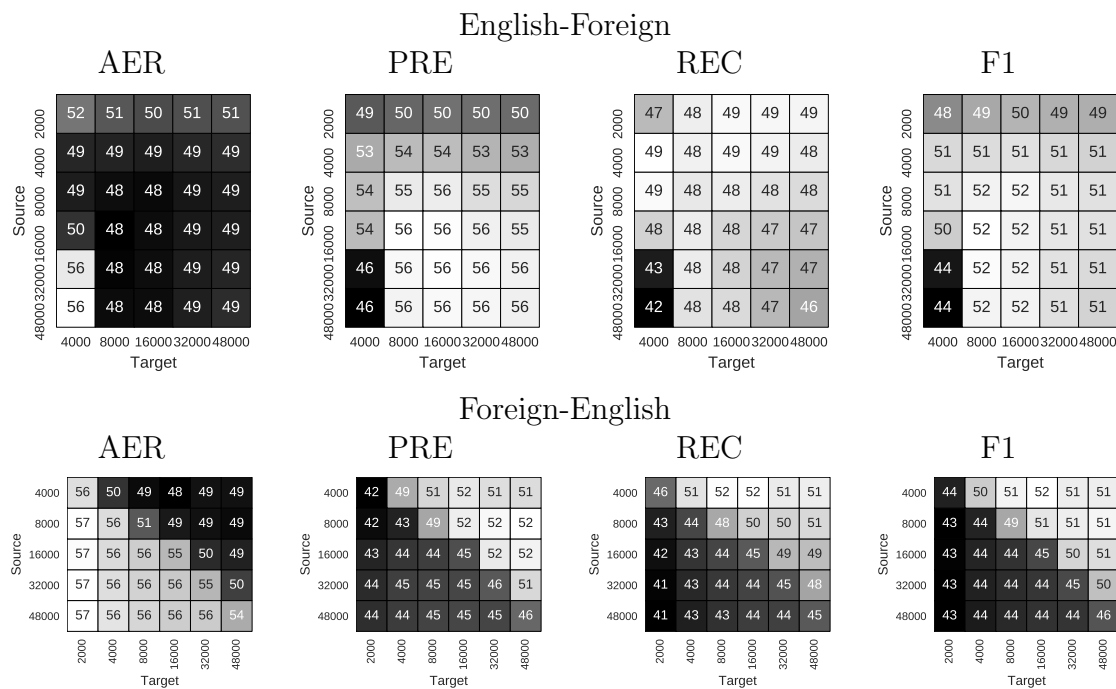


Figure E.23: BPE-based `Fastalign` with BPE-dropout for English-Japanese: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).

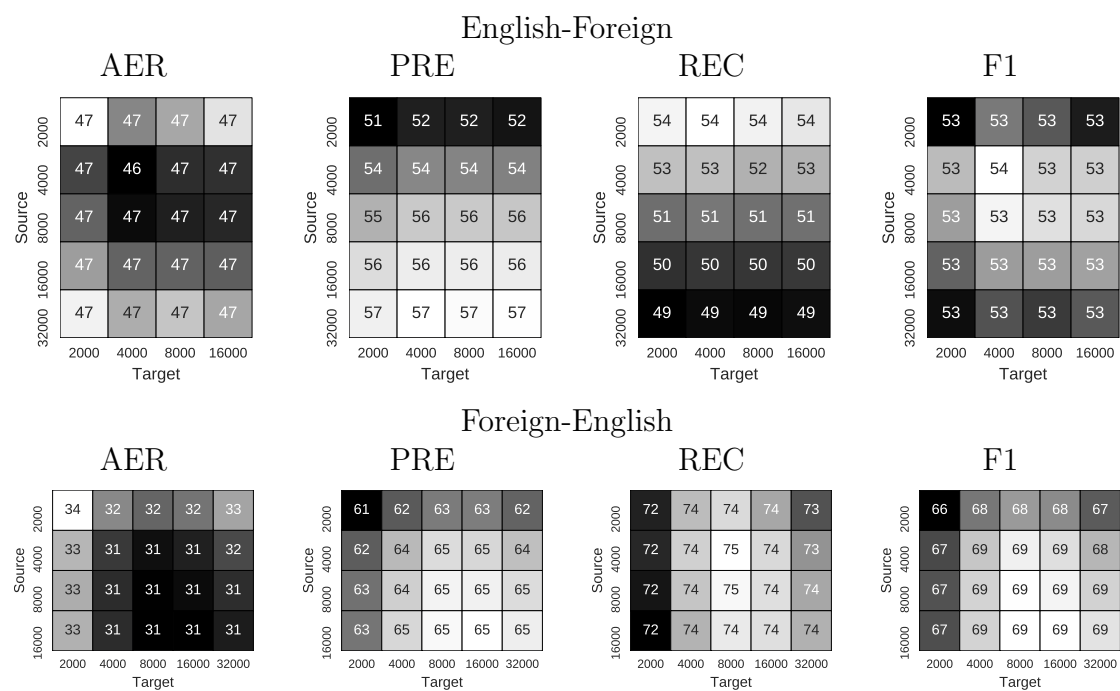


Figure E.24: BPE-based `Fastalign` with BPE-dropout for English-Vietnamese: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).

E.1.4 Eflomal: Matrix of scores

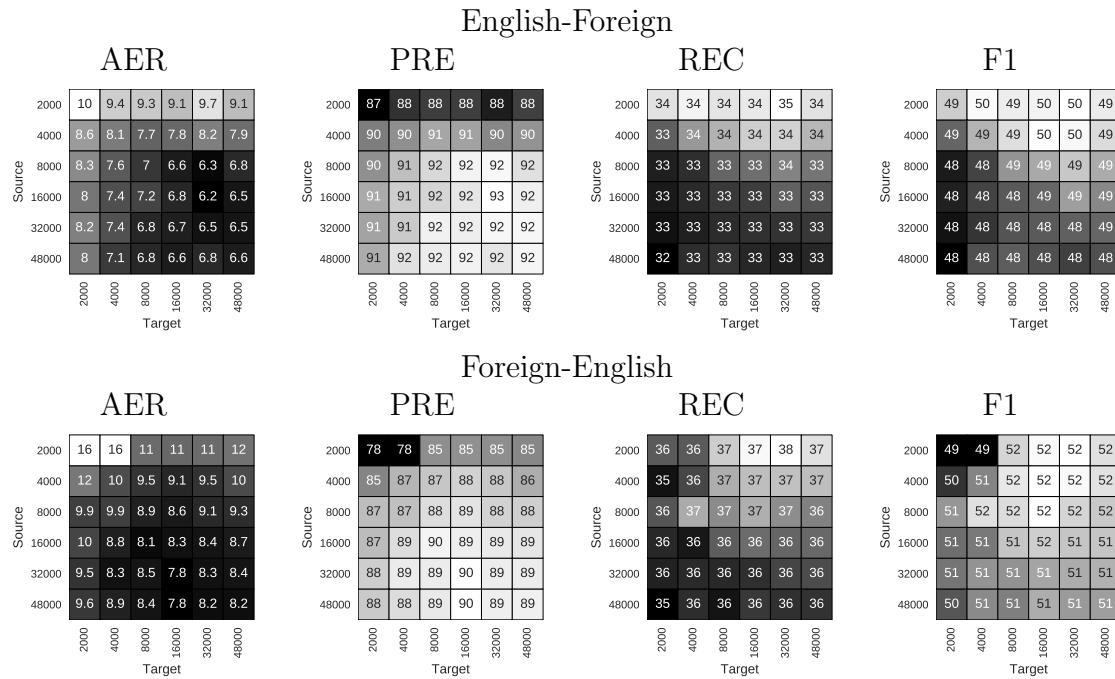


Figure E.25: BPE-based Eflomal for English-French: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).

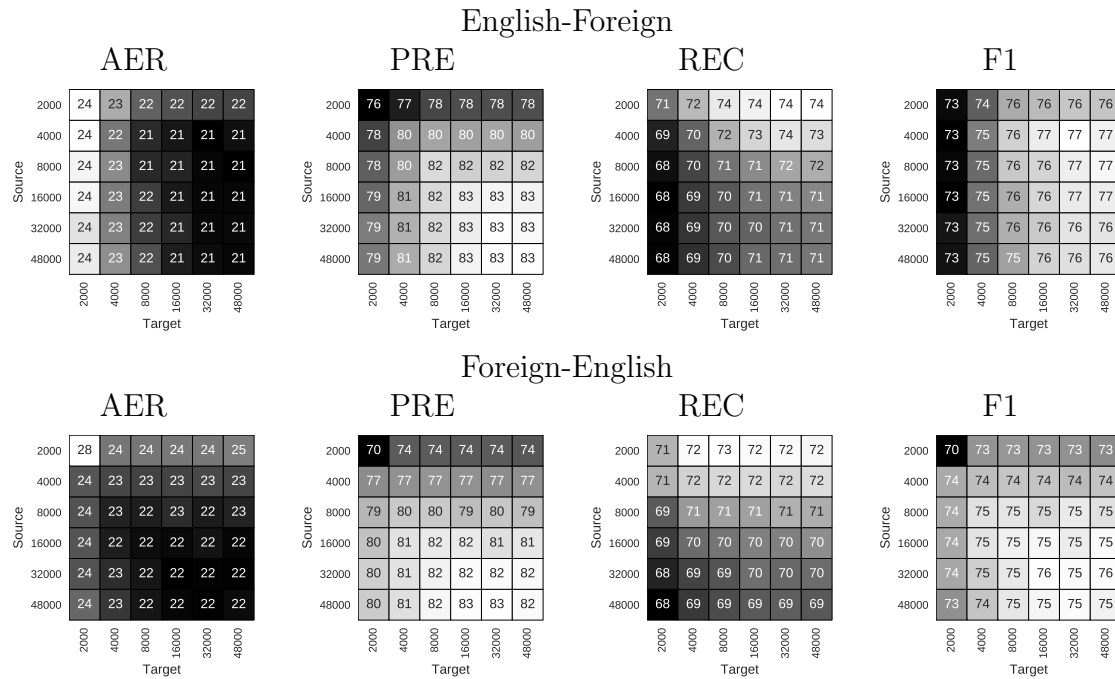


Figure E.26: BPE-based E_{f1oma1} for English-German: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).

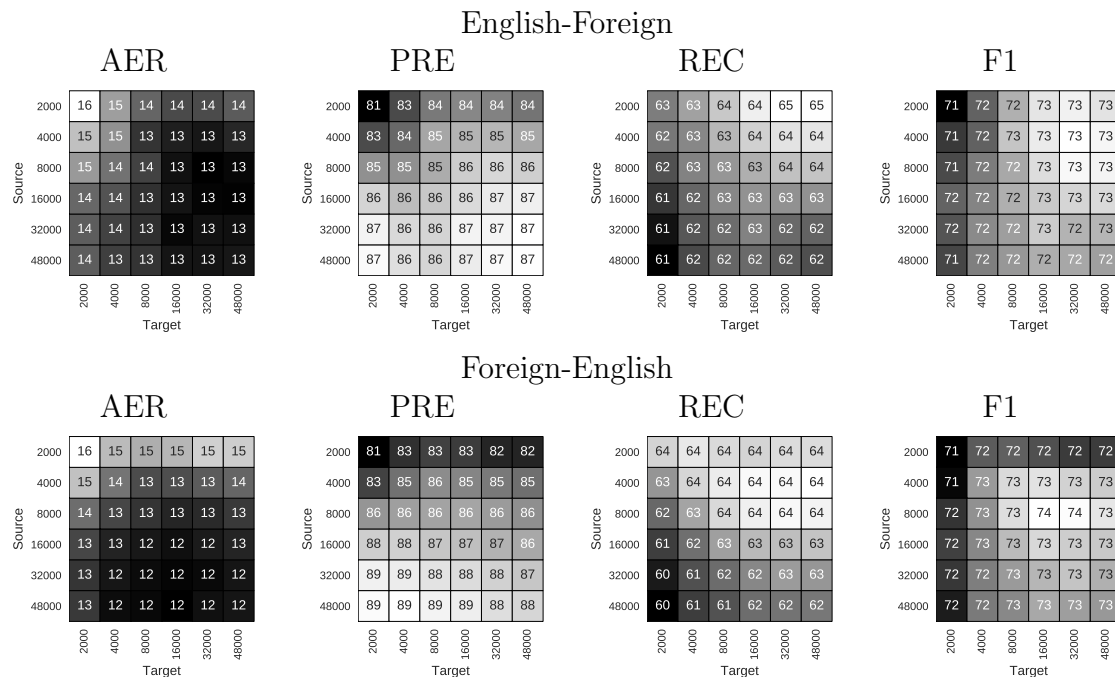


Figure E.27: BPE-based E_{f1oma1} for English-Czech: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).

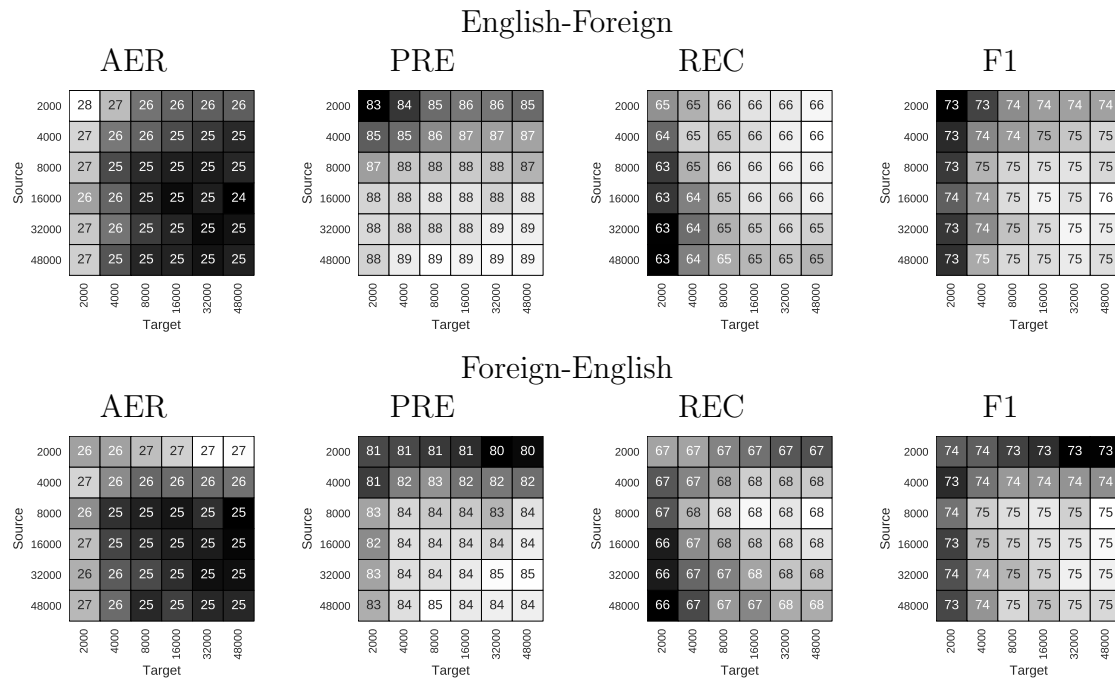


Figure E.28: BPE-based E_{flomal} for English-Romanian: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).

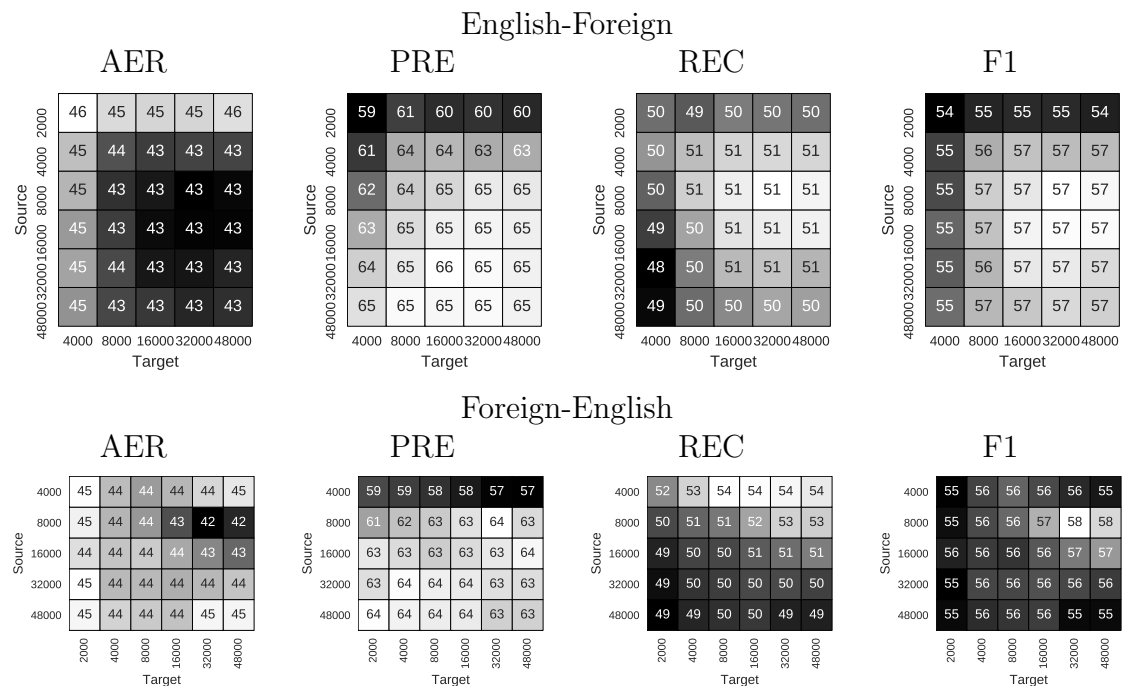


Figure E.29: BPE-based E_{flomal} for English-Japanese: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).

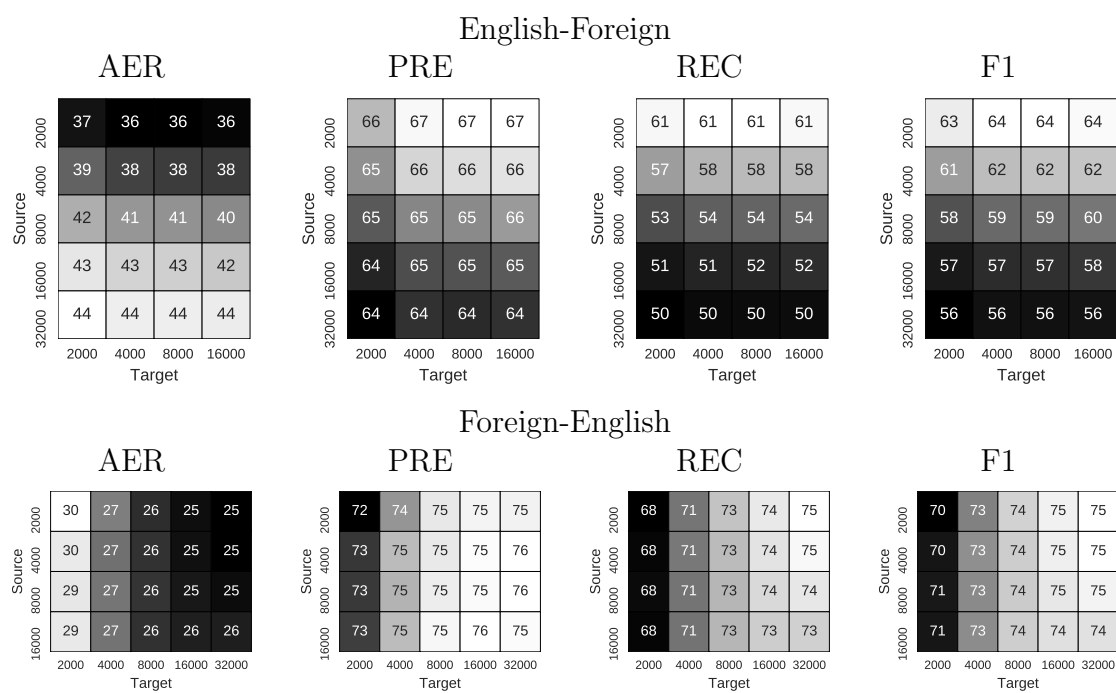


Figure E.30: BPE-based E_{flomal} for English-Vietnamese: For each pair (vocabulary size of source and target), we show Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC).

E.1.5 Neural models with recommended BPE configurations

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
IBM-1 Giza++	40.1	26.7	71.55	16.41	89.01	33.9	36.49	59.24	26.37	88.81
IBM-1+NN	27.96	36.42	69.66	24.65	89.5	27.21	38.08	68.12	26.43	89.52
IBM-1+NNChar	28.76	37.5	67.13	26.01	89.42	31.4	37.21	62.64	26.47	89.11
IBM-1+BPE+B	25.71	37.27	70.64	25.31	89.61	24.05	40.42	70.84	28.28	89.83
IBM-1+BPE+B+C	23.15	39.2	73.93	26.67	89.91	23.01	41.32	72.2	28.94	89.98
Fastalign	15.19	44.98	82.5	30.92	90.78	16.23	46.32	80.08	32.58	90.79
HMM Giza++	11.99	45.18	86.12	30.62	90.94	11.97	45.98	85.2	31.49	90.98
HMM+NN	11.84	45.57	86.68	30.91	91	11.15	46.86	86.17	32.18	91.1
HMM+NNCharTgt	9.17	47.22	89.53	32.07	91.26	9.56	47.87	88.15	32.86	91.27
HMM+NNCharJB	8.47	44.38	91.8	29.26	91.06	7.74	46.26	91.42	30.96	91.23
HMM+BPE+B	9.84	41.74	93.17	26.9	90.84	10.48	43.2	91.57	28.27	90.94
HMM+BPE+B+C	8.93	43.26	94.07	28.09	91.01	9.86	44.18	92.16	29.06	91.05
IBM-4 Giza++	10	44.43	90.61	29.43	91.02	9.64	45.43	89.58	30.43	91.08

Table E.13: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-French

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
IBM-1 Giza++	39.03	58.76	59.1	58.43	96.4	42.66	55.39	57.02	53.84	96.2
IBM-1+NN	37.64	60.07	62.98	57.41	96.65	39.22	58.53	62.22	55.25	96.57
IBM-1+NNChar	36.22	61.55	62.76	60.39	96.69	40.88	56.99	59.75	54.48	96.4
IBM-1+BPE+B	31.36	66.52	73.38	60.83	97.32	34.46	63.34	64.35	62.36	96.84
IBM-1+BPE+B+C	31.02	67.29	72.93	62.45	97.34	33.93	63.88	64.99	62.81	96.89
Fastalign	28.98	68.75	71.11	66.54	97.35	31.28	66.47	70.73	62.69	97.23
HMM Giza++	23.92	73.3	79.23	68.2	97.82	26.33	71.04	79.47	64.23	97.7
HMM+NN	26.78	70.95	73.94	68.2	97.55	29.44	68.21	74.69	62.76	97.44
HMM+NNCharTgt	26.04	71.57	75.99	67.64	97.64	28.11	69.48	75.59	64.29	97.52
HMM+NNCharJB	23.69	73.38	82.38	66.15	97.9	24.9	72.16	83.36	63.61	97.85
HMM+BPE+B	19.61	78.25	85.82	71.92	98.25	20.38	77.38	84.28	71.52	98.17
HMM+BPE+B+C	19.17	79.19	86.61	72.94	98.32	20.36	77.41	84.36	71.52	98.17
IBM-4 Giza++	21.46	75.48	85.79	67.39	98.08	23.31	73.63	86.56	64.06	97.99

Table E.14: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-German

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
IBM-1 Giza++	45.09	46.75	50.4	43.59	95.97	48.47	42.88	49.17	38.02	95.89
IBM-1+NN	42.29	48.64	54.32	44.04	96.22	40.97	49.08	56.81	43.2	96.36
IBM-1+NNChar	40.85	50.35	54.2	47.01	96.24	42.35	47.99	55.68	42.17	96.29
IBM-1+BPE+B	30.76	59.86	72.25	51.1	97.22	32.71	56.45	71.4	46.68	97.08
IBM-1+BPE+B+C	30.58	60.04	72.48	51.24	97.23	31.63	57.67	72.36	47.94	97.14
Fastalign	25.75	64.09	70.98	58.42	97.34	25.3	62.86	73.13	55.13	97.36
HMM Giza++	27.86	61.22	70.81	53.92	97.23	30.38	57.28	69.26	48.83	97.04
HMM+NN	23.5	65.45	74.39	58.43	97.5	24.06	64.03	75.48	55.6	97.46
HMM+NNCharTgt	16.74	69.36	83.82	59.15	97.88	24.61	63.94	73.42	56.63	97.41
HMM+NNCharJB	16.38	69.35	83.84	59.13	97.88	23.88	62.72	78.56	52.19	97.48
HMM+BPE+B	16.24	71.87	85.27	62.1	98.03	18.74	67.93	82.85	57.56	97.79
HMM+BPE+B+C	16.17	71.97	85.36	62.22	98.03	17.81	69.09	83.85	58.75	97.87
IBM-4 Giza++	20.92	65.7	79.48	56	97.63	26.5	59.81	75.58	49.48	97.3

Table E.15: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Czech

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
IBM-1 Giza++	56.02	43.99	58.8	35.14	96.66	53.52	46.49	49.92	43.5	96.26
IBM-1+NN	46.4	53.62	57.71	50.07	96.77	44.9	55.11	60.08	50.9	96.91
IBM-1+NNChar	50.16	49.85	54.28	46.09	96.54	48.28	51.73	56.08	48.01	96.66
IBM-1+BPE+B	43.47	56.54	63.95	50.67	97.09	40.42	59.59	64.19	55.61	97.19
IBM-1+BPE+B+C	41.52	58.5	65.87	52.61	97.22	39.09	60.93	65.67	56.83	97.28
Fastalign	33.36	66.65	72.77	61.49	97.7	32.91	67.1	73.7	61.59	97.75
HMM Giza++	33.36	66.65	75.28	59.8	97.77	36.38	63.64	72.9	56.46	97.59
HMM+NN	30.69	69.33	76.93	63.09	97.92	40.12	59.89	63.85	56.4	97.18
HMM+NNCharTgt	27.54	72.47	78.13	67.59	98.08	28.01	72.01	82.79	63.71	98.15
HMM+NNCharJB	26.85	73.17	83.55	65.08	98.22	29.76	70.26	80.7	62.21	98.04
HMM+BPE+B	34.41	65.61	84.04	53.81	97.9	29.34	70.68	82.91	61.59	98.09
HMM+BPE+B+C	34.87	65.15	81.71	54.18	97.84	30.16	69.86	80.67	61.61	98.02
IBM-4 Giza++	31.04	68.98	79.28	61.04	97.95	32.3	67.72	80.97	58.2	97.93

Table E.16: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Romanian

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
IBM-1 Giza++	63.12	36.88	40.24	34.04	97.22	61.55	38.45	40.13	36.9	97.18
IBM-1+NN	62.64	37.36	48.83	30.25	97.58	56.9	43.1	49.95	37.91	97.61
IBM-1+NNChar	57.96	42.04	57.54	33.12	97.82	54.91	45.09	47.76	42.7	97.52
IBM-1+BPE+B	56.51	43.49	48.67	39.3	97.56	57.27	42.73	44.39	41.18	97.37
IBM-1+BPE+B+C	54.87	45.13	50.38	40.87	97.63	55.05	44.95	46.66	43.36	97.47
Fastalign	50.67	49.33	55.79	44.21	97.83	49.37	50.63	54.59	47.2	97.8
HMM Giza++	57.01	42.99	52.22	36.53	97.69	54.41	45.59	52.45	40.32	97.7
HMM+NN	49.68	50.32	70.46	39.13	98.16	47.09	52.91	56.81	49.51	97.9
HMM+NNCharTgt	43.92	56.08	65.13	49.25	98.16	43.6	56.4	67.58	48.39	98.21
HMM+NNCharJB	40.92	59.08	84.77	45.35	98.5	37.71	62.29	67.77	57.62	98.33
HMM+BPE+B	38.33	61.65	72.6	53.56	98.41	38.93	61.04	62.67	59.5	98.19
HMM+BPE+B+C	38.3	61.7	71.71	54.15	98.4	38.93	61.07	62.42	59.78	98.18
IBM-4 Giza++	52.52	47.48	61.27	38.76	97.95	49.23	50.77	61.46	43.24	98

Table E.17: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Japanese

Models	English-Foreign					Foreign-English				
	AER	F1	PRE	REC	ACC	AER	F1	PRE	REC	ACC
IBM-1 Giza++	69.43	30.58	43.42	23.59	96.51	46.45	53.56	50.79	56.64	96.8
IBM-1+NN	58.87	41.13	57.25	32.1	97.01	42.25	57.75	57.33	58.19	97.23
IBM-1+NNChar	55.06	44.94	62.54	35.07	97.2	41.15	58.86	63.9	54.56	97.51
IBM-1+BPE+B	55.85	44.15	49.18	40.06	96.7	37.72	62.26	64.34	60.31	97.62
IBM-1+BPE+B+C	54.33	45.67	51.19	41.22	96.8	34.08	65.93	68.05	63.93	97.85
Fastalign	48.89	51.11	56.87	46.42	97.11	32.82	67.19	64.5	70.11	97.77
HMM Giza++	57.86	42.14	49.18	36.86	96.7	37.57	62.43	60.28	64.74	97.46
HMM+NN	49.27	50.73	63.52	42.24	97.33	31.45	68.56	67.75	69.39	97.93
HMM+NNCharTgt	47.52	52.49	67.71	42.86	97.47	30.94	69.06	75.69	63.51	98.15
HMM+NNCharJB	43.28	56.73	84.49	42.7	97.88	27.59	72.42	72.7	72.14	98.21
HMM+BPE+B	47.03	52.97	62.75	45.83	97.35	27.76	72.24	74.13	70.45	98.24
HMM+BPE+B+C	45.85	54.15	64.41	46.71	97.42	26.05	73.95	75.61	72.37	98.34
IBM-4 Giza++	51.91	48.1	58.38	40.89	97.12	33.19	66.82	67.17	66.47	97.85

Table E.18: Alignment error rate (AER), accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) for English-Vietnamese

E.2 Correct and incorrect alignment links

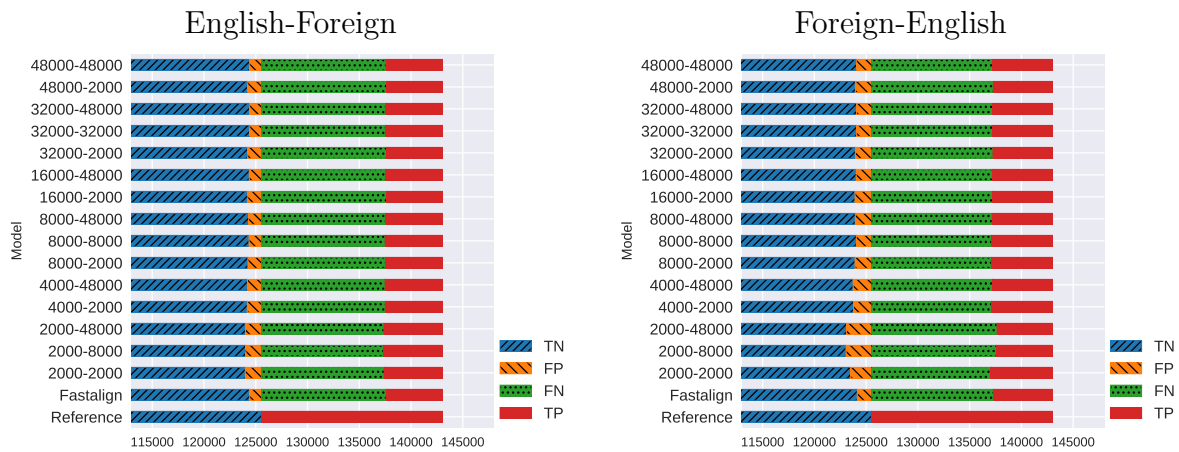


Figure E.31: BPE-based Fastalign for English-French: We observe correct and incorrect alignment links.

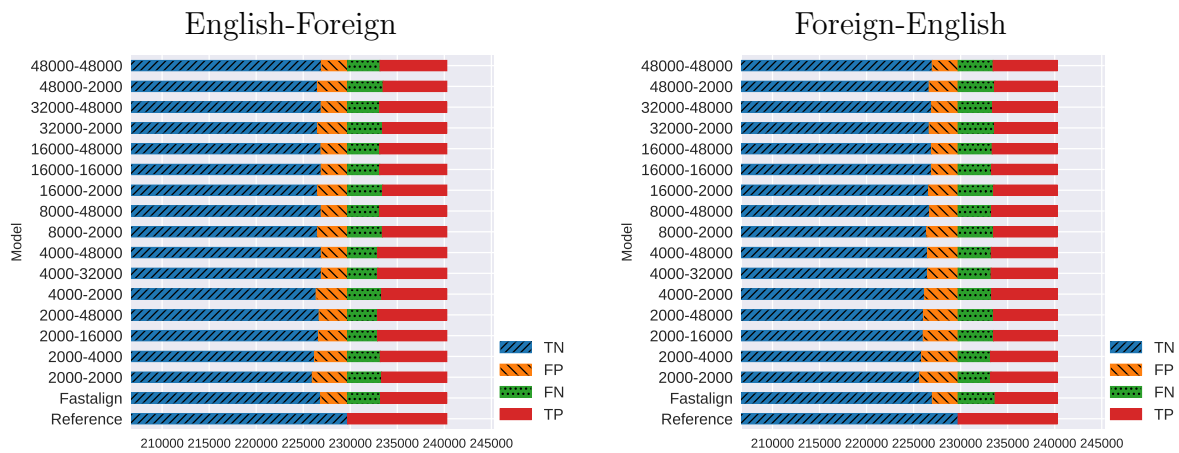


Figure E.32: BPE-based Fastalign for English-German: We observe correct and incorrect alignment links.

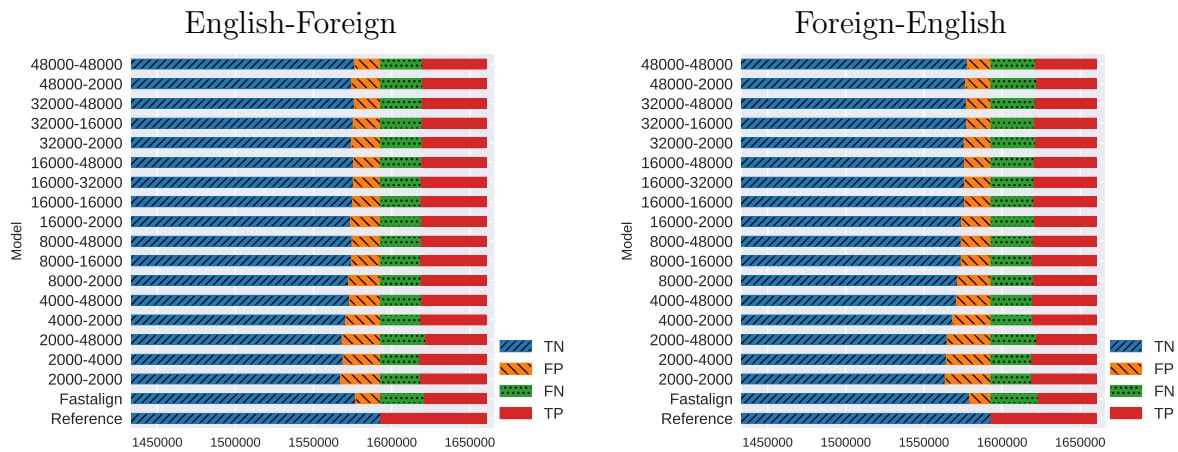


Figure E.33: BPE-based Fastalign for English-Czech: We observe correct and incorrect alignment links.

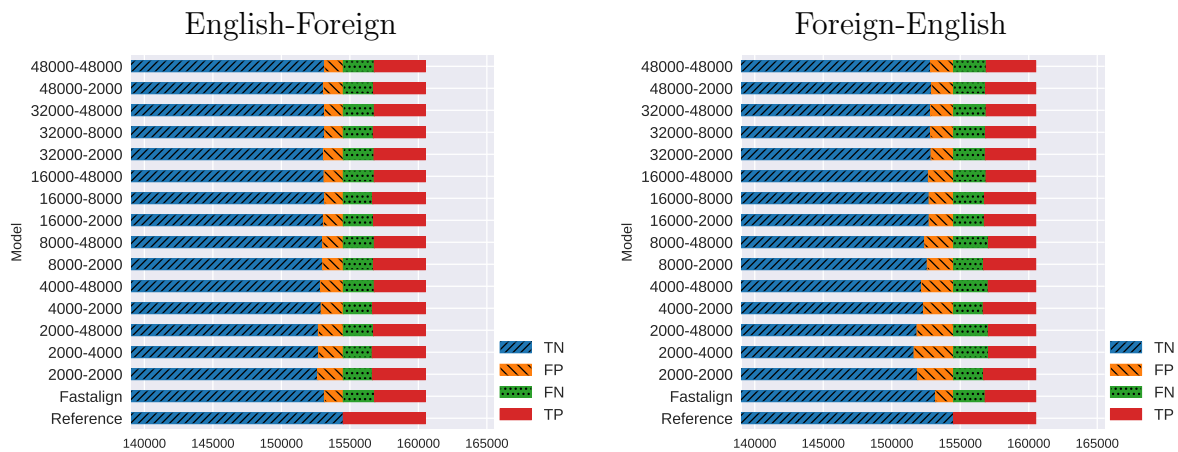


Figure E.34: BPE-based Fastalign for English-Romanian: We observe correct and incorrect alignment links.

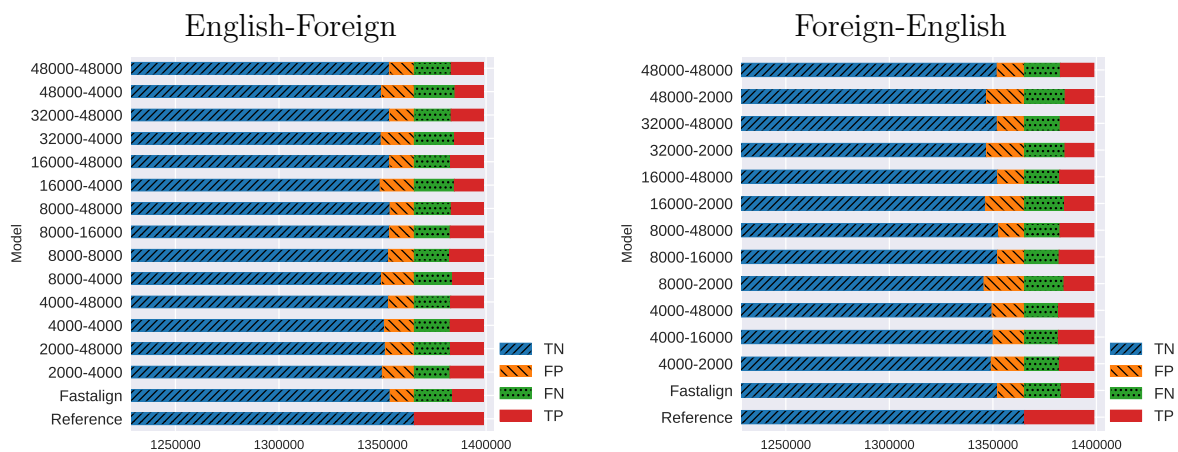


Figure E.35: BPE-based Fastalign for English-Japanese: We observe correct and incorrect alignment links.

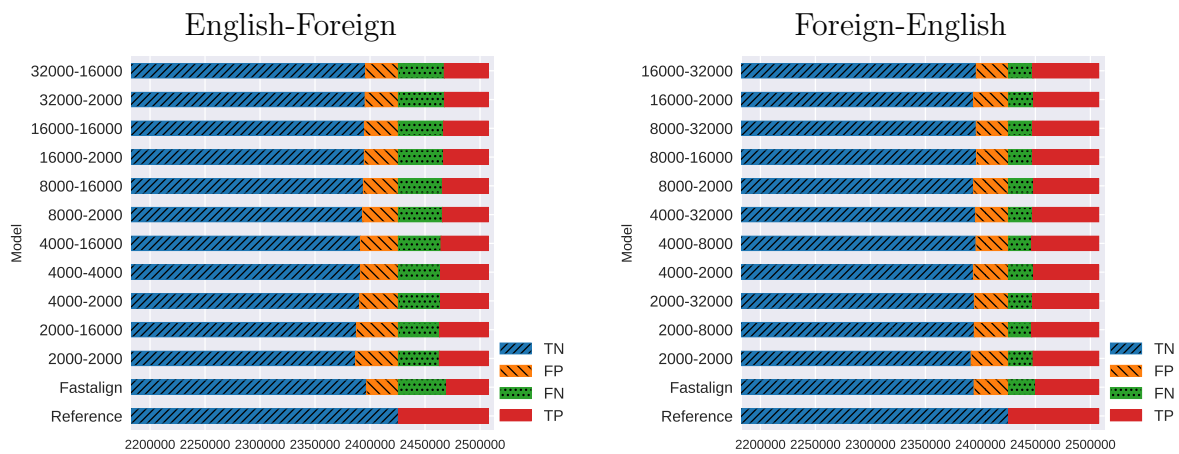


Figure E.36: BPE-based Fastalign for English-Vietnamese: We observe correct and incorrect alignment links.

E.3 Unaligned words

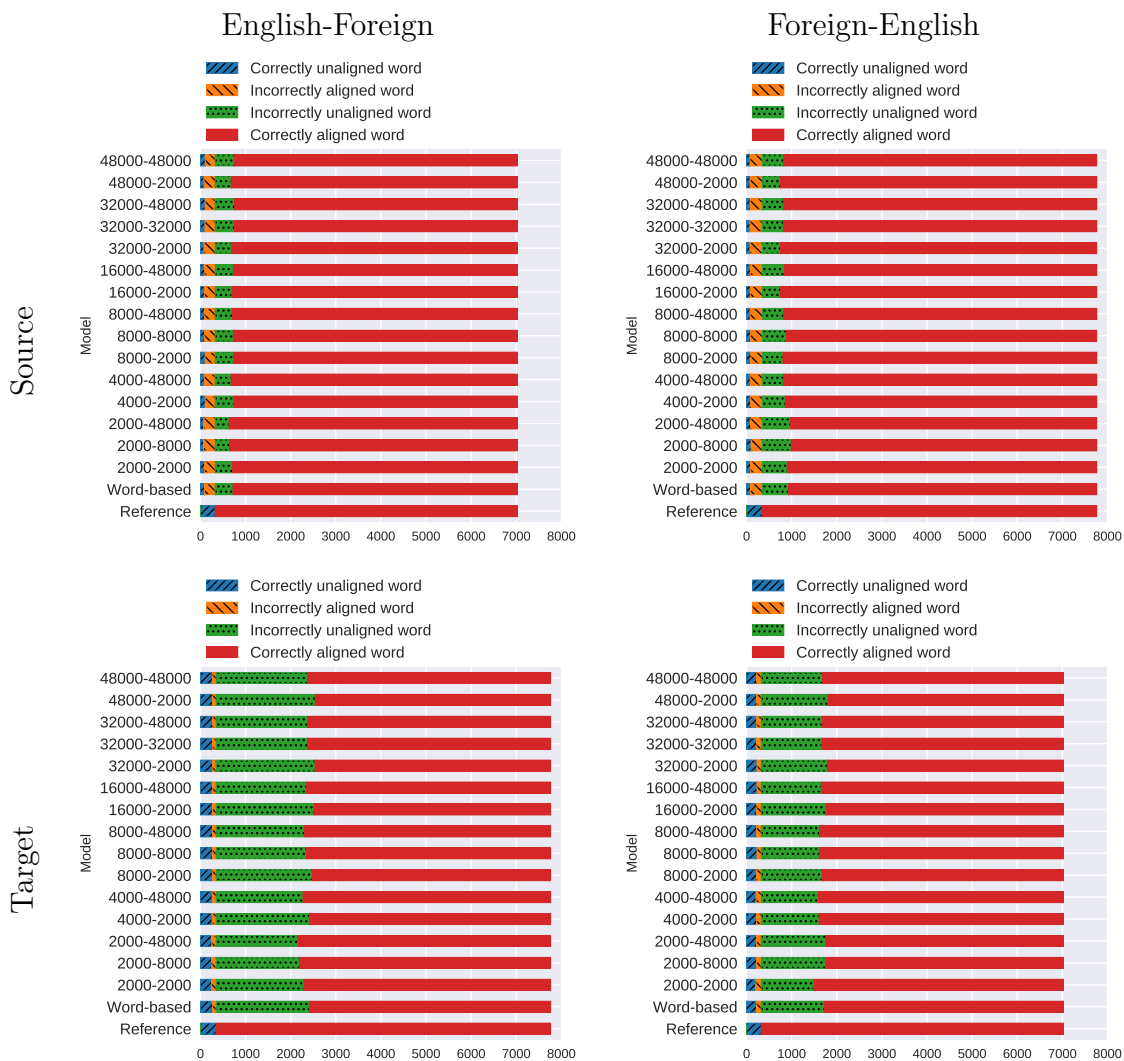


Figure E.37: BPE-based Fastalign: Unaligned words for English-French

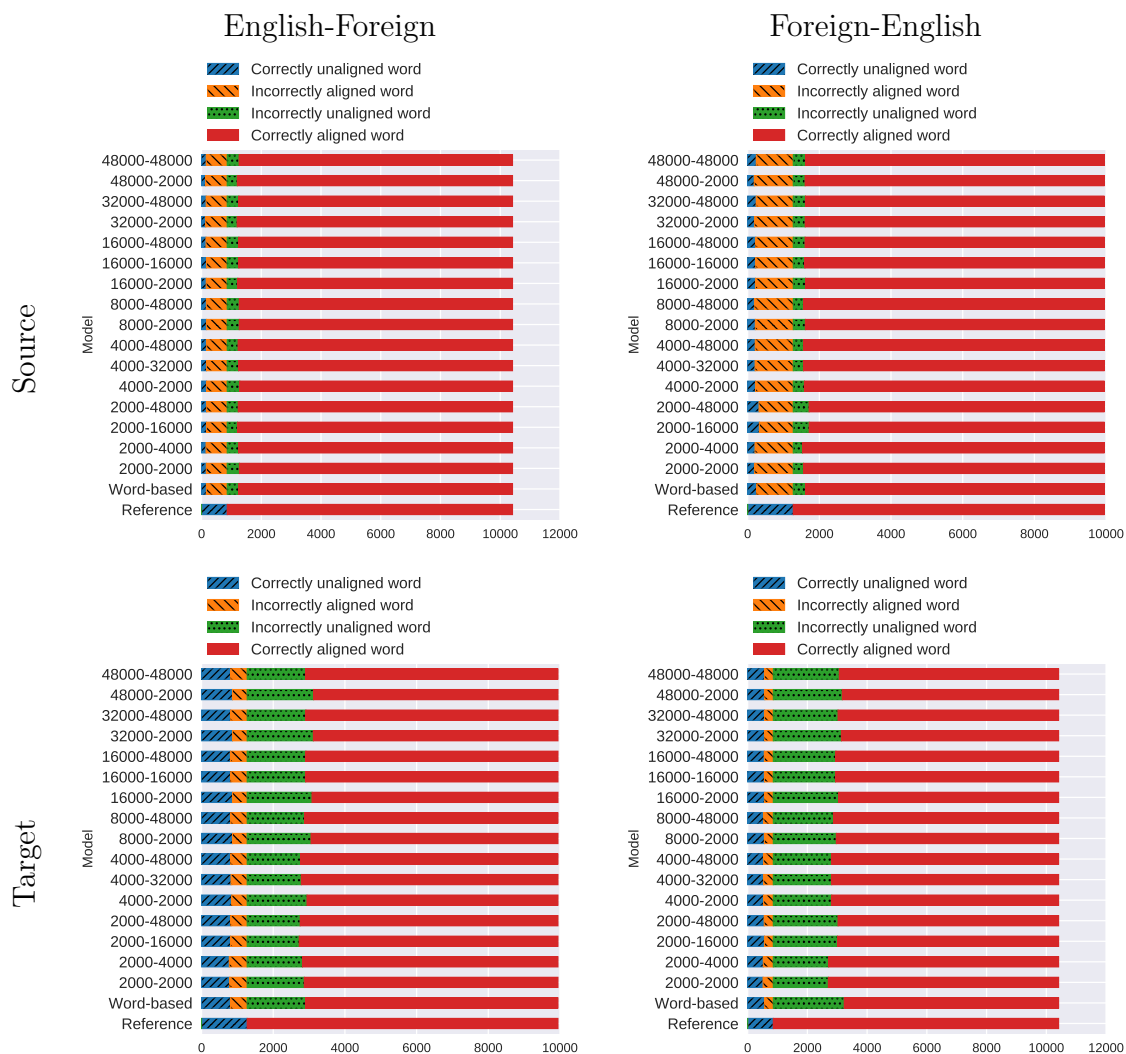


Figure E.38: BPE-based Fastalign: Unaligned words for English-German

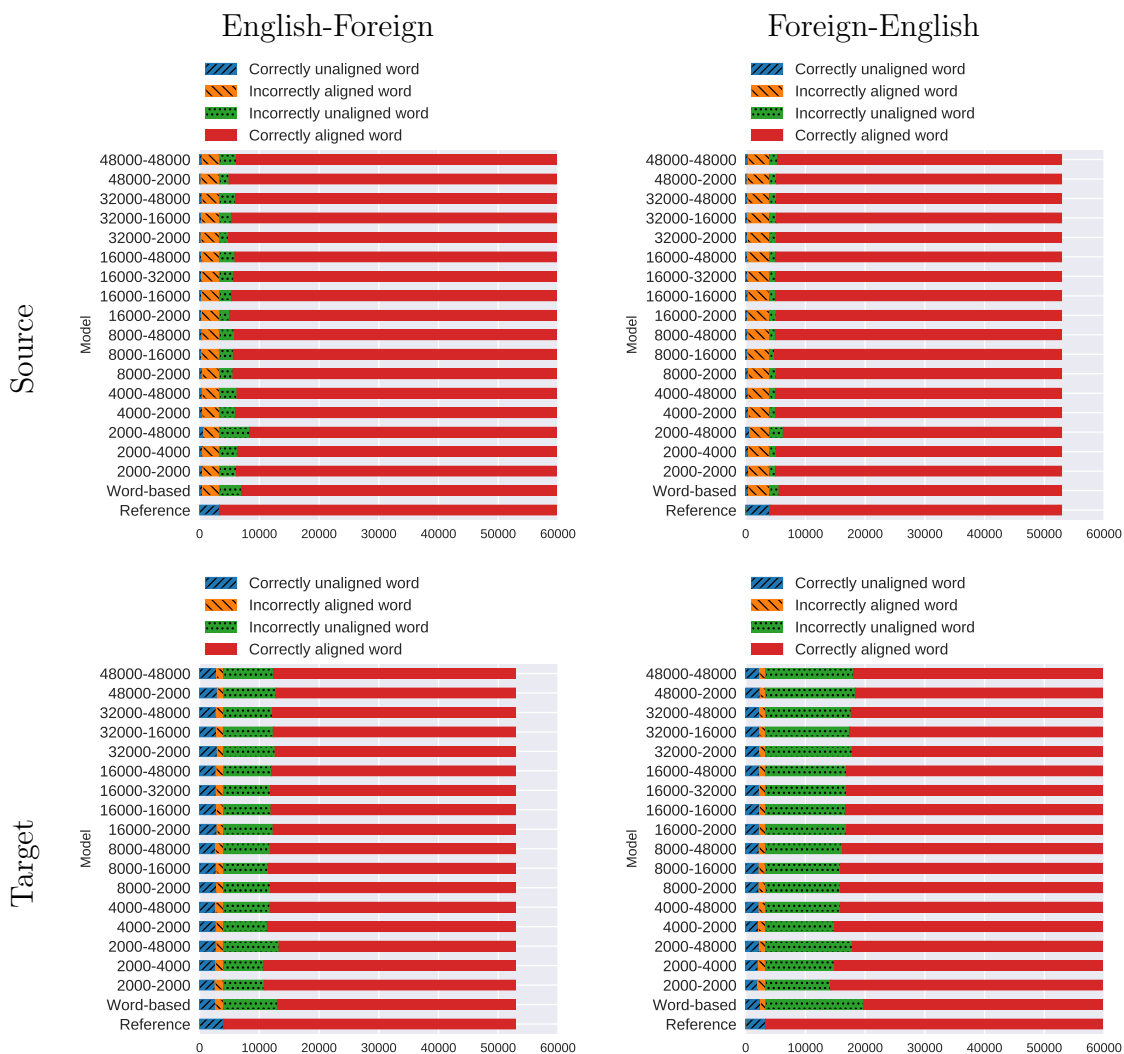


Figure E.39: BPE-based Fastalign: Unaligned words for English-Czech

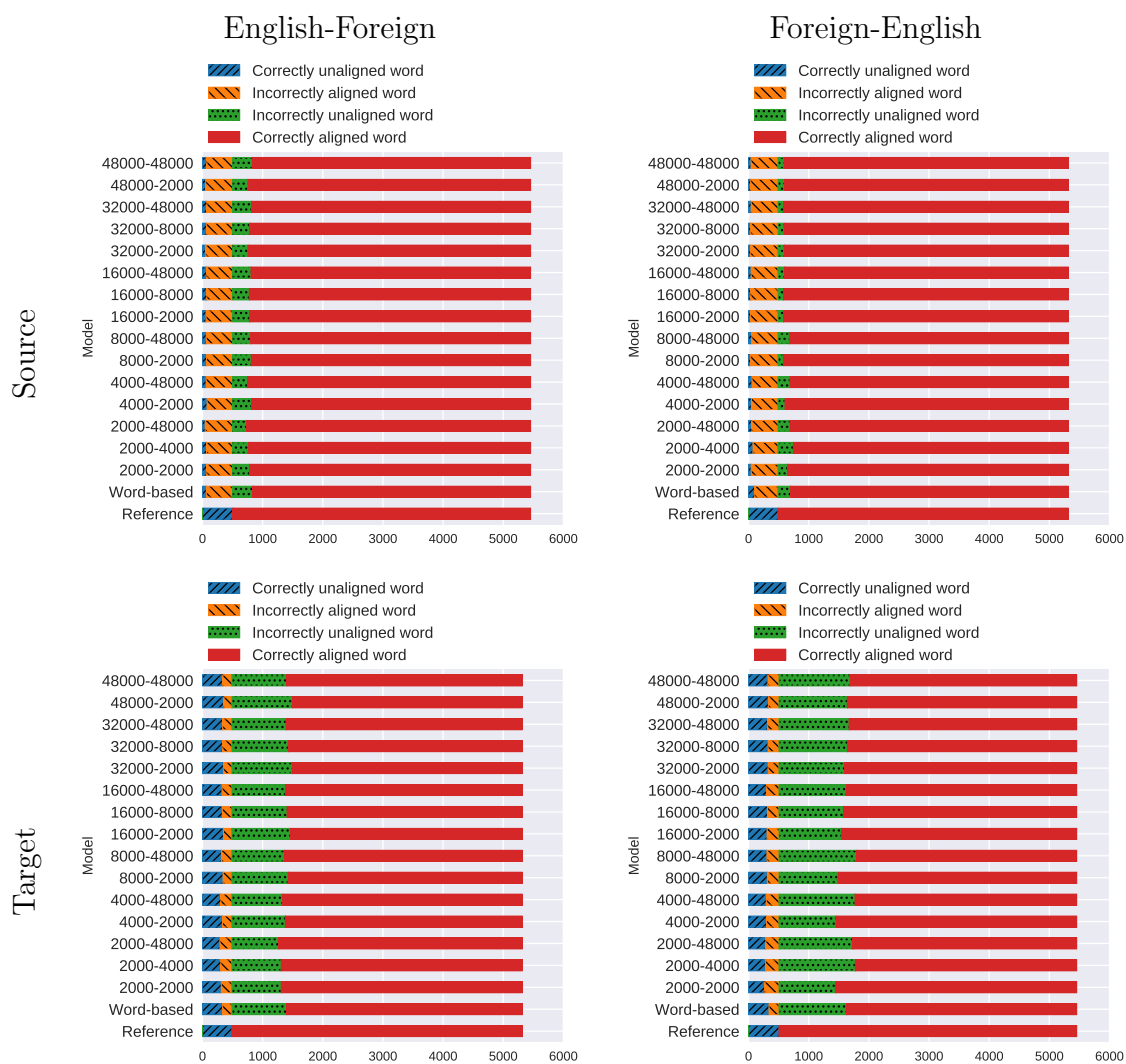


Figure E.40: BPE-based Fastalign: Unaligned words for English-Romanian

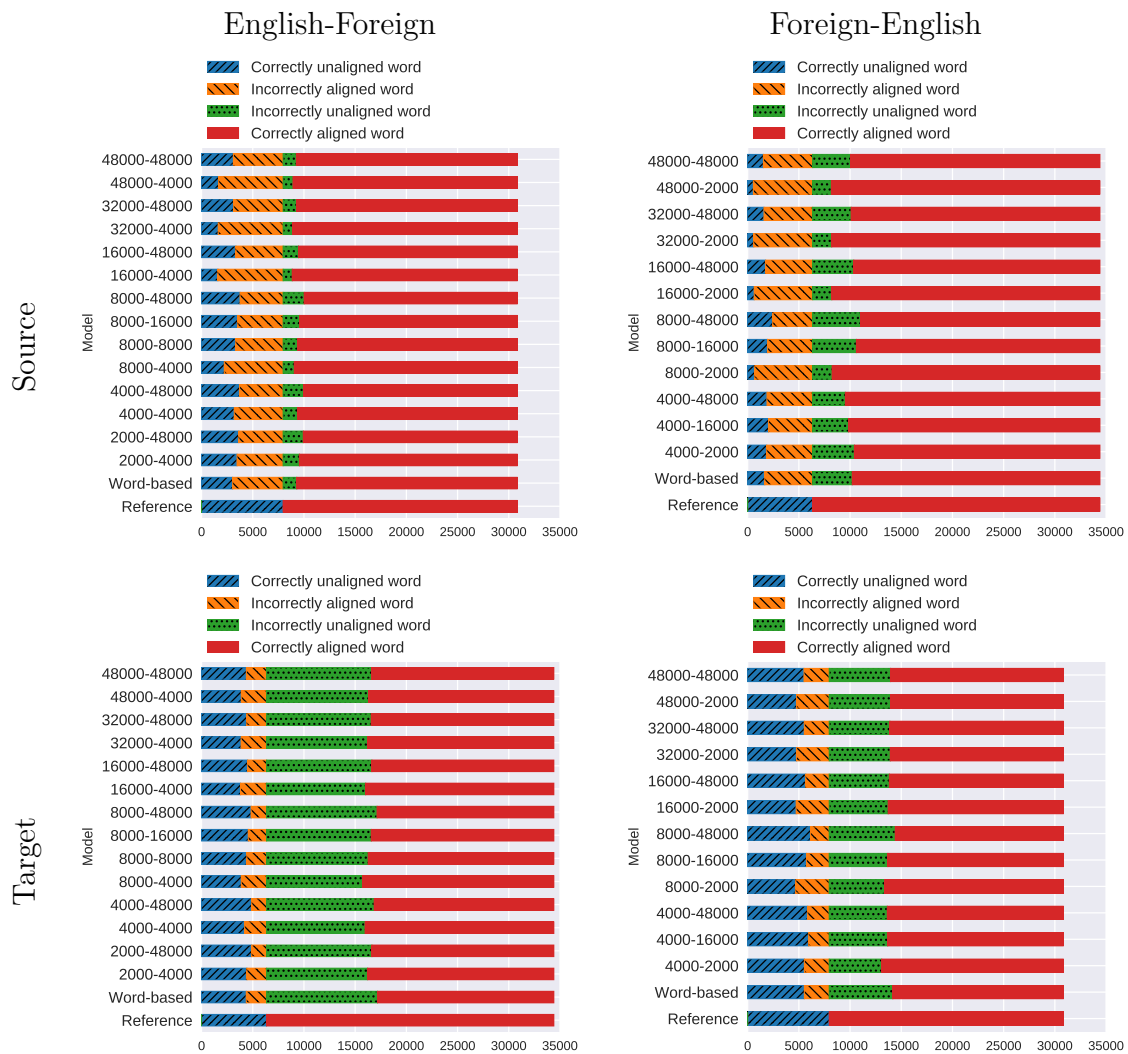


Figure E.41: BPE-based Fastalign: Unaligned words for English-Japanese

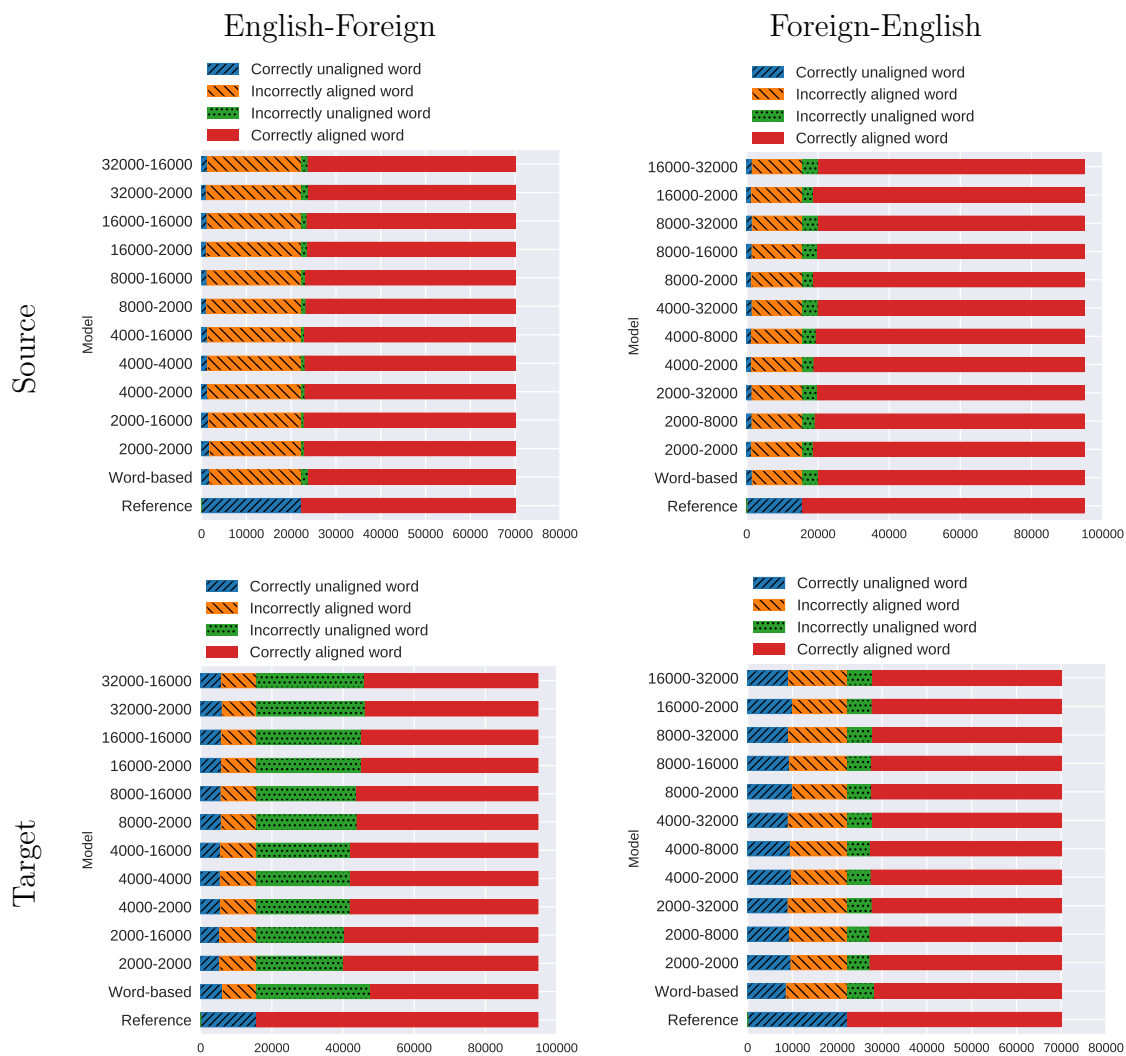
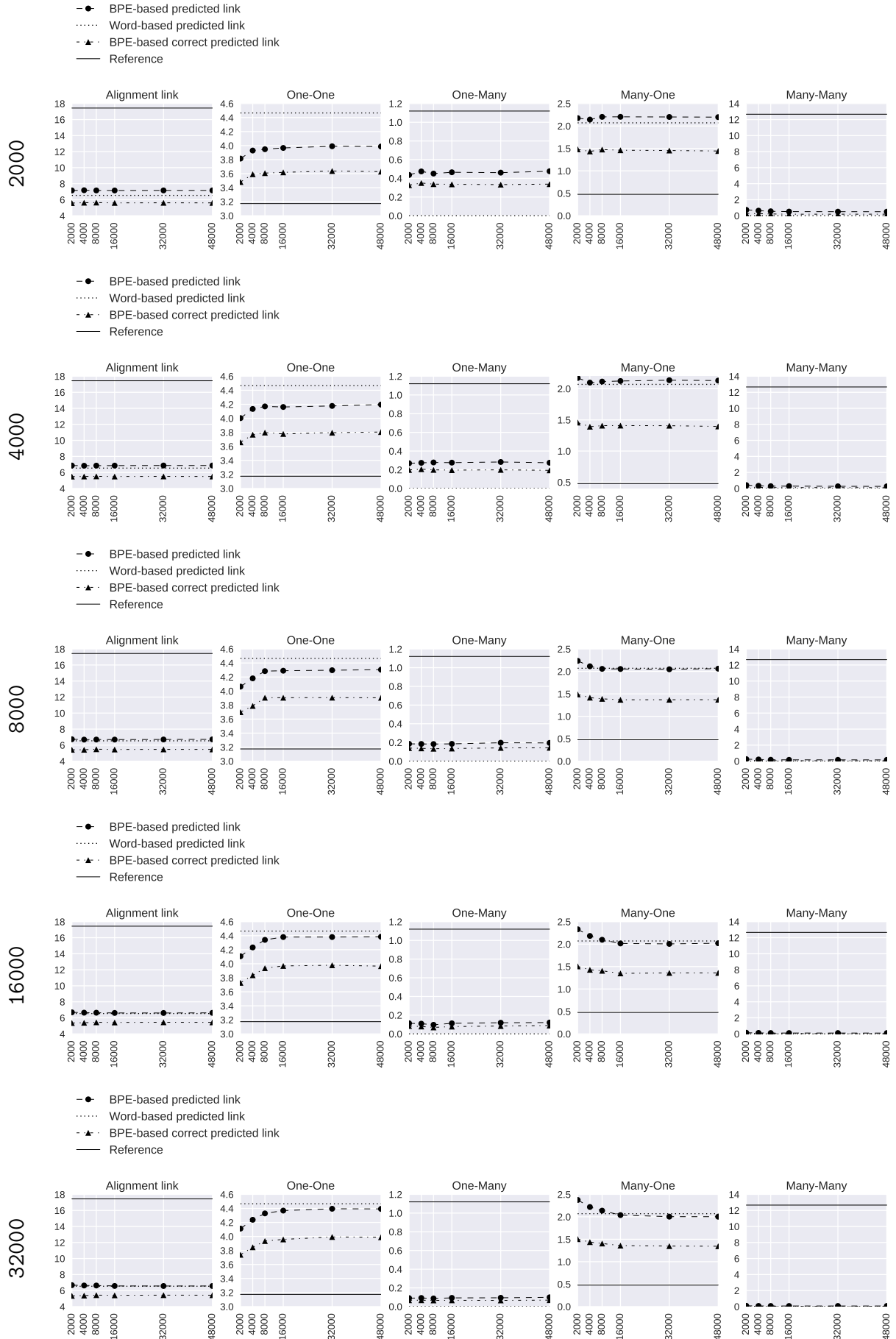


Figure E.42: BPE-based Fastalign: Unaligned words for English-Vietnamese

E.4 Alignment types



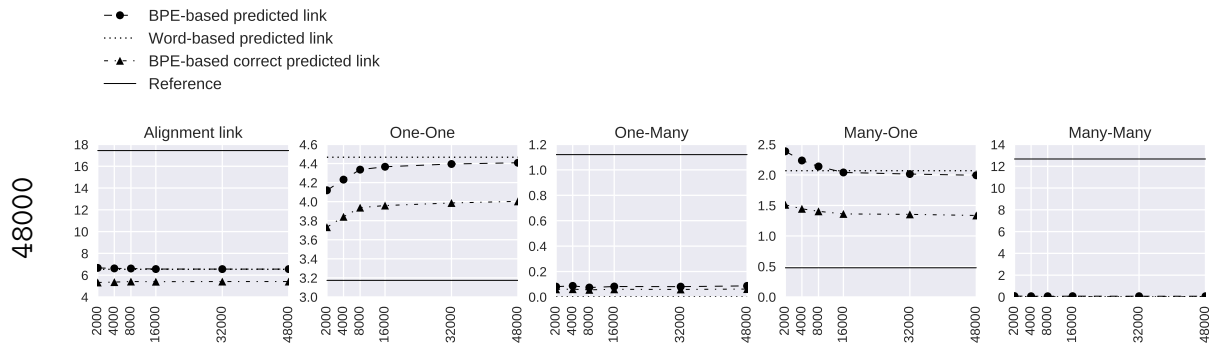
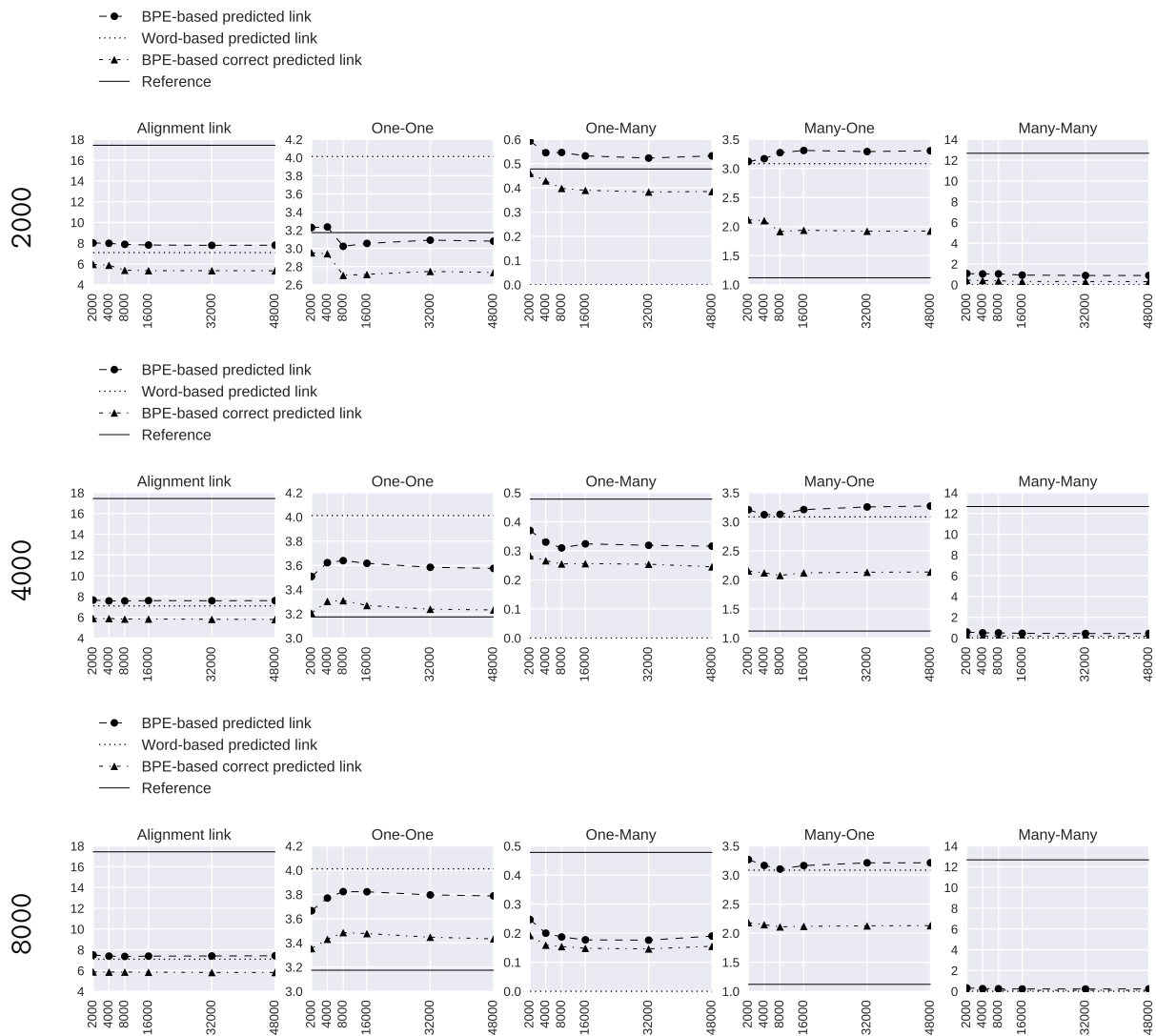


Figure E.43: BPE-based **Fastalign** for the direction English-French: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).



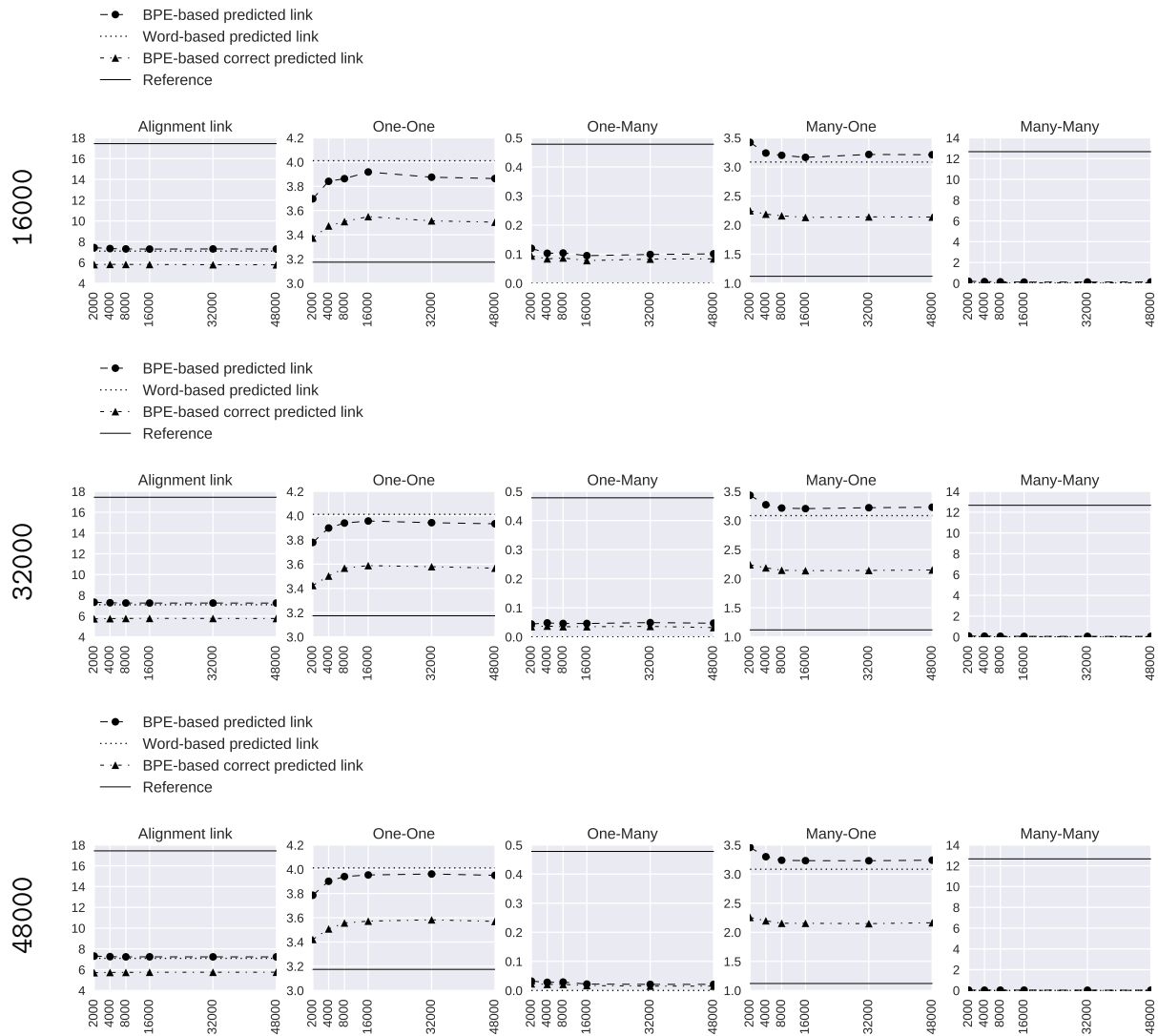
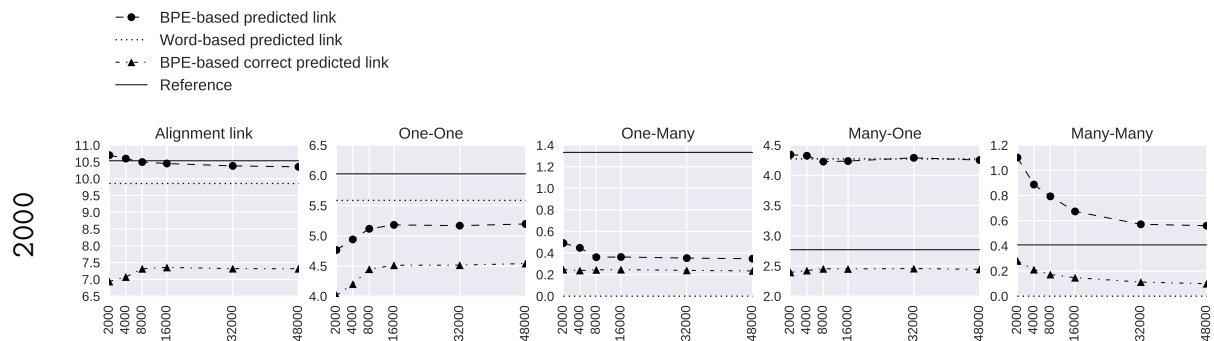


Figure E.44: BPE-based `Fastalign` for the direction French-English: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).



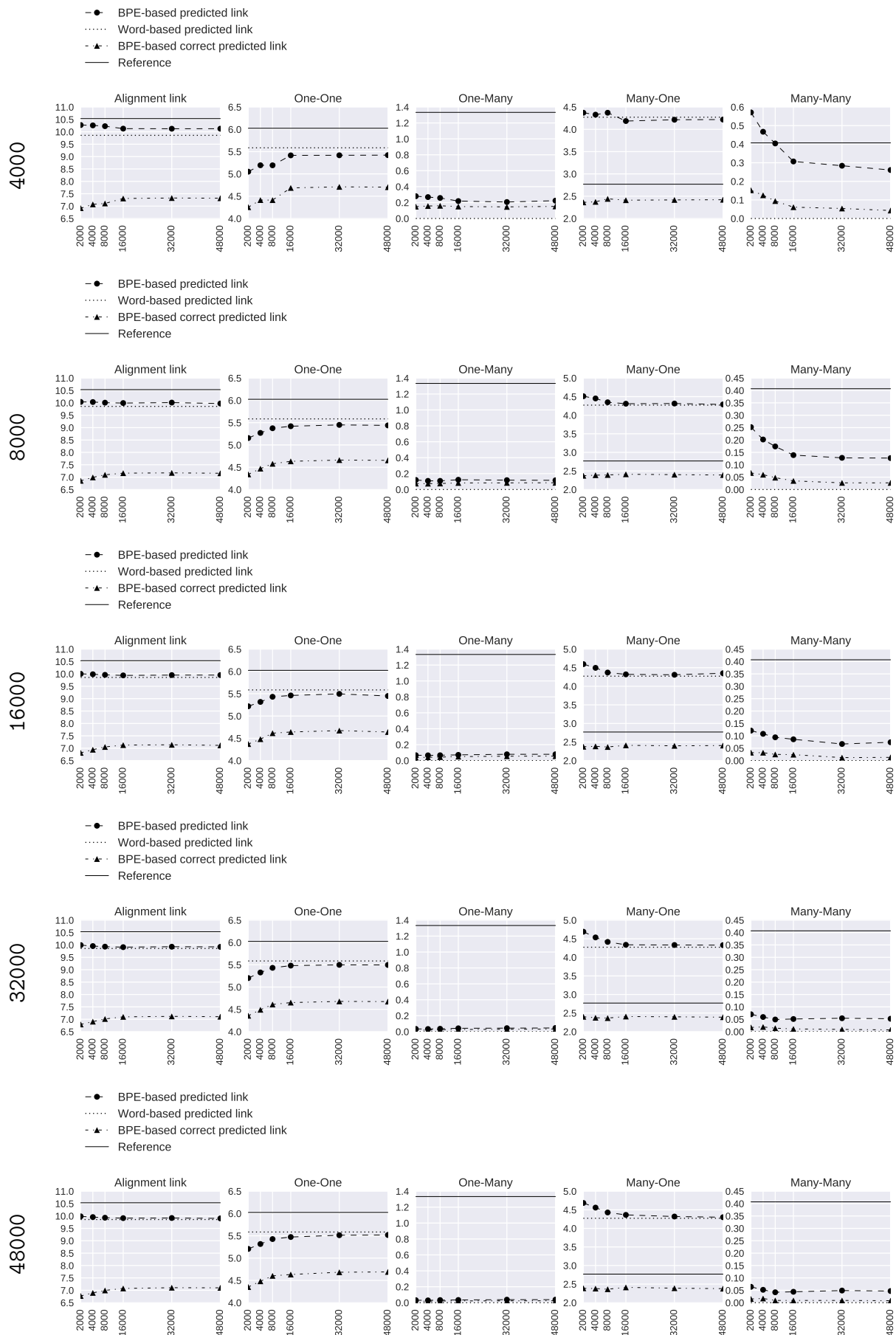
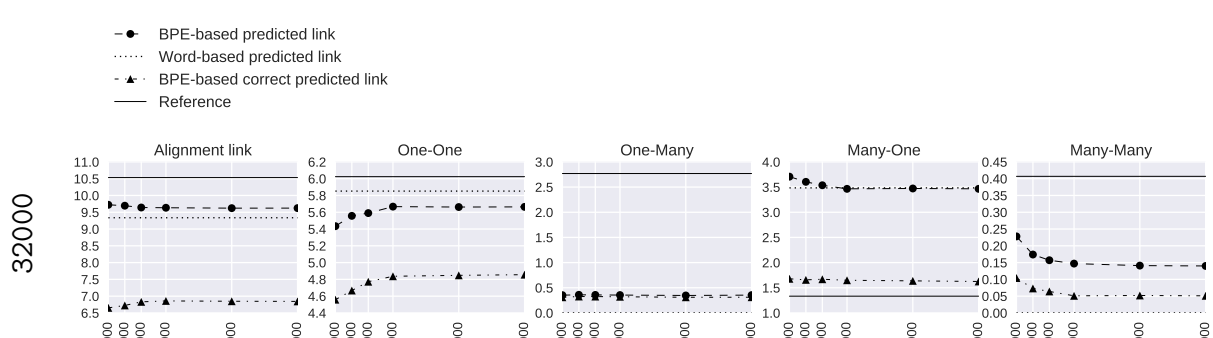
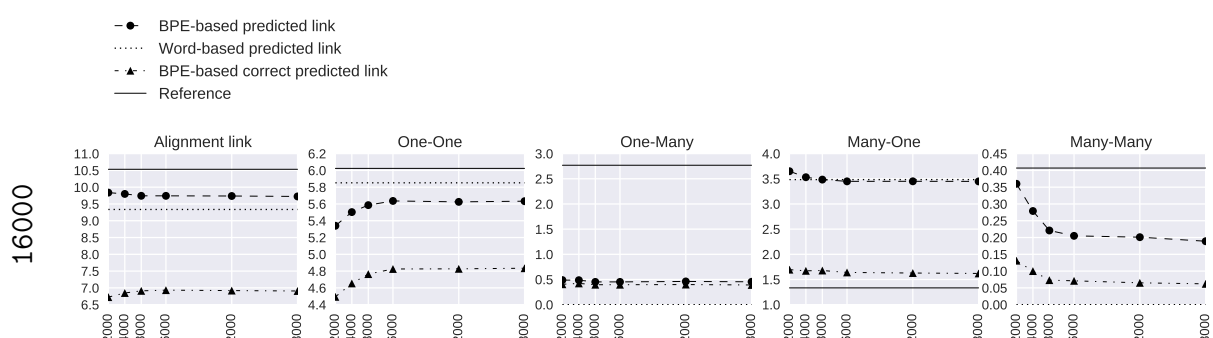
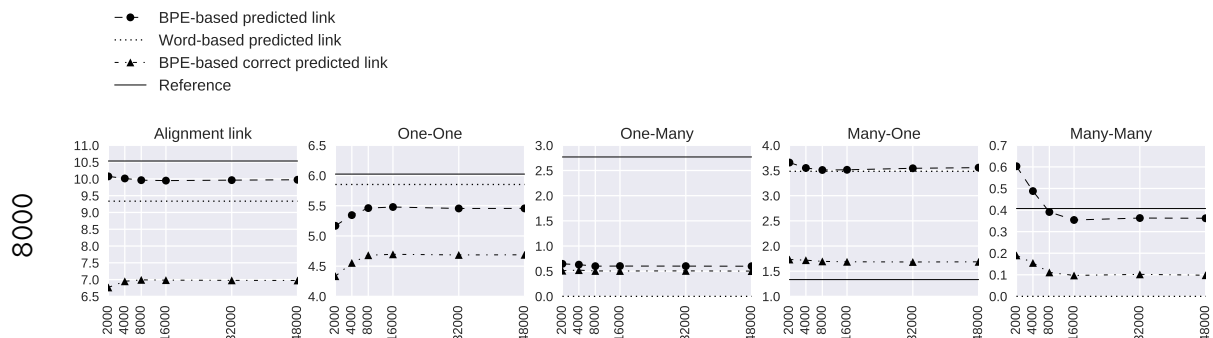
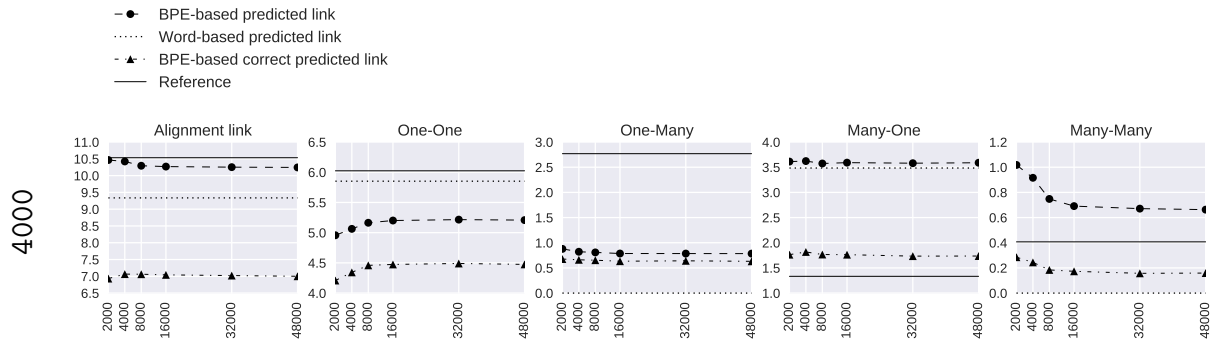


Figure E.45: BPE-based Fastalign for the direction English-German: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).



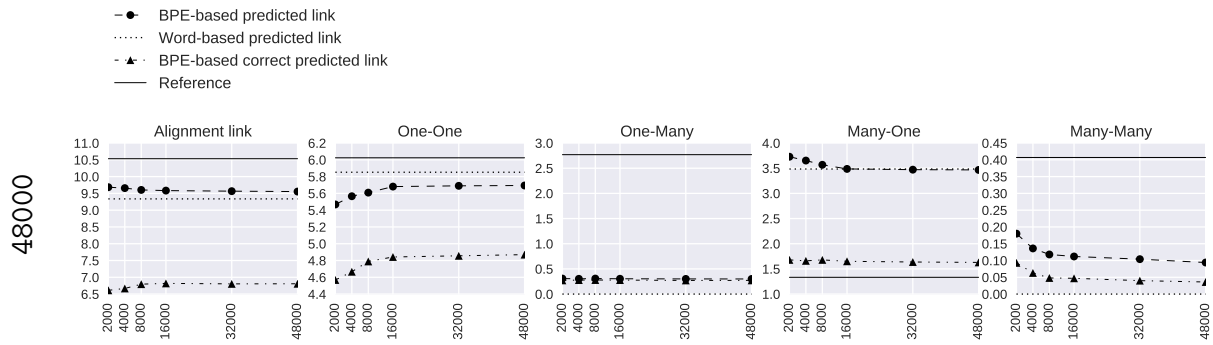
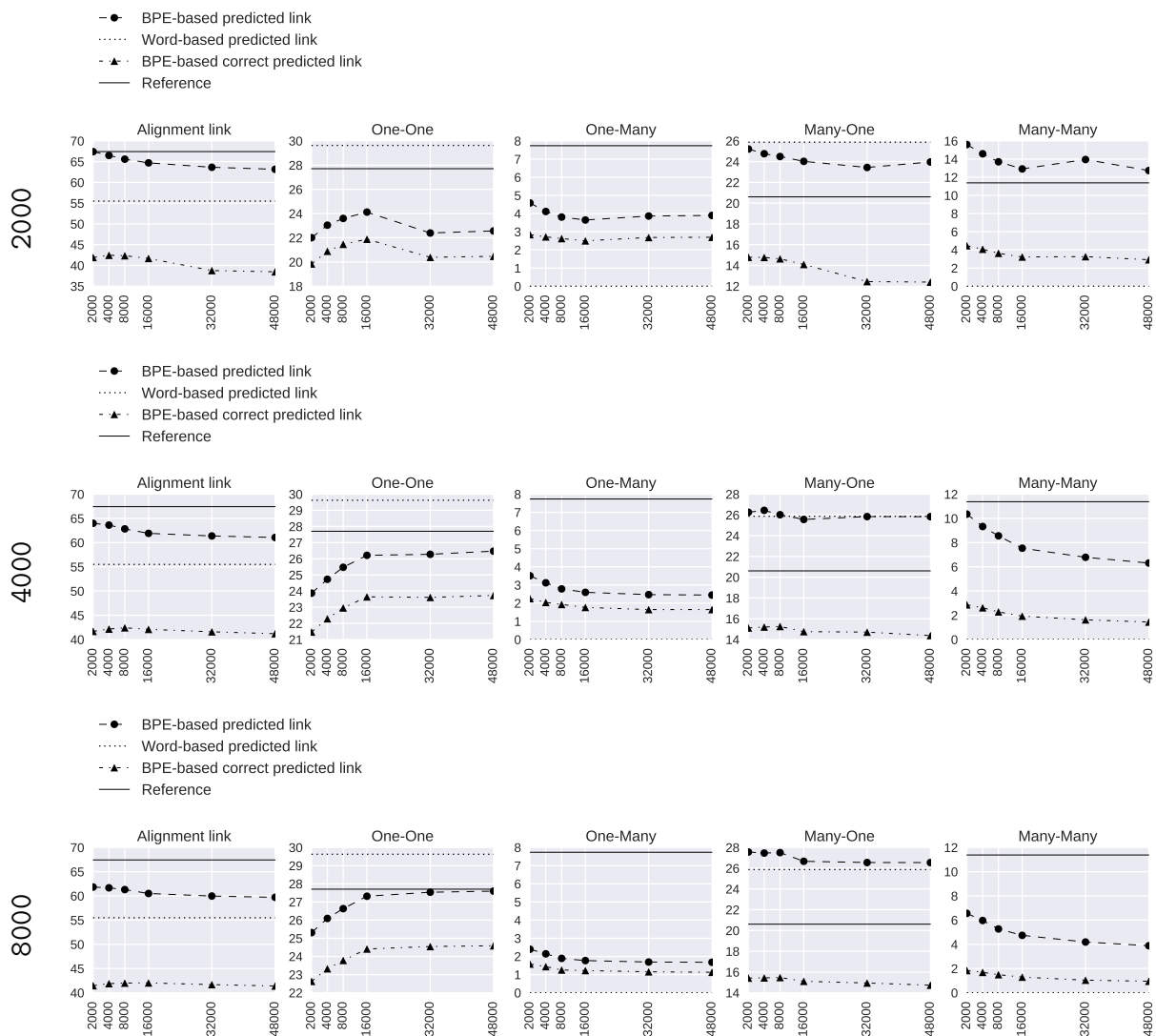


Figure E.46: BPE-based Fastalign for the direction German-English: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).



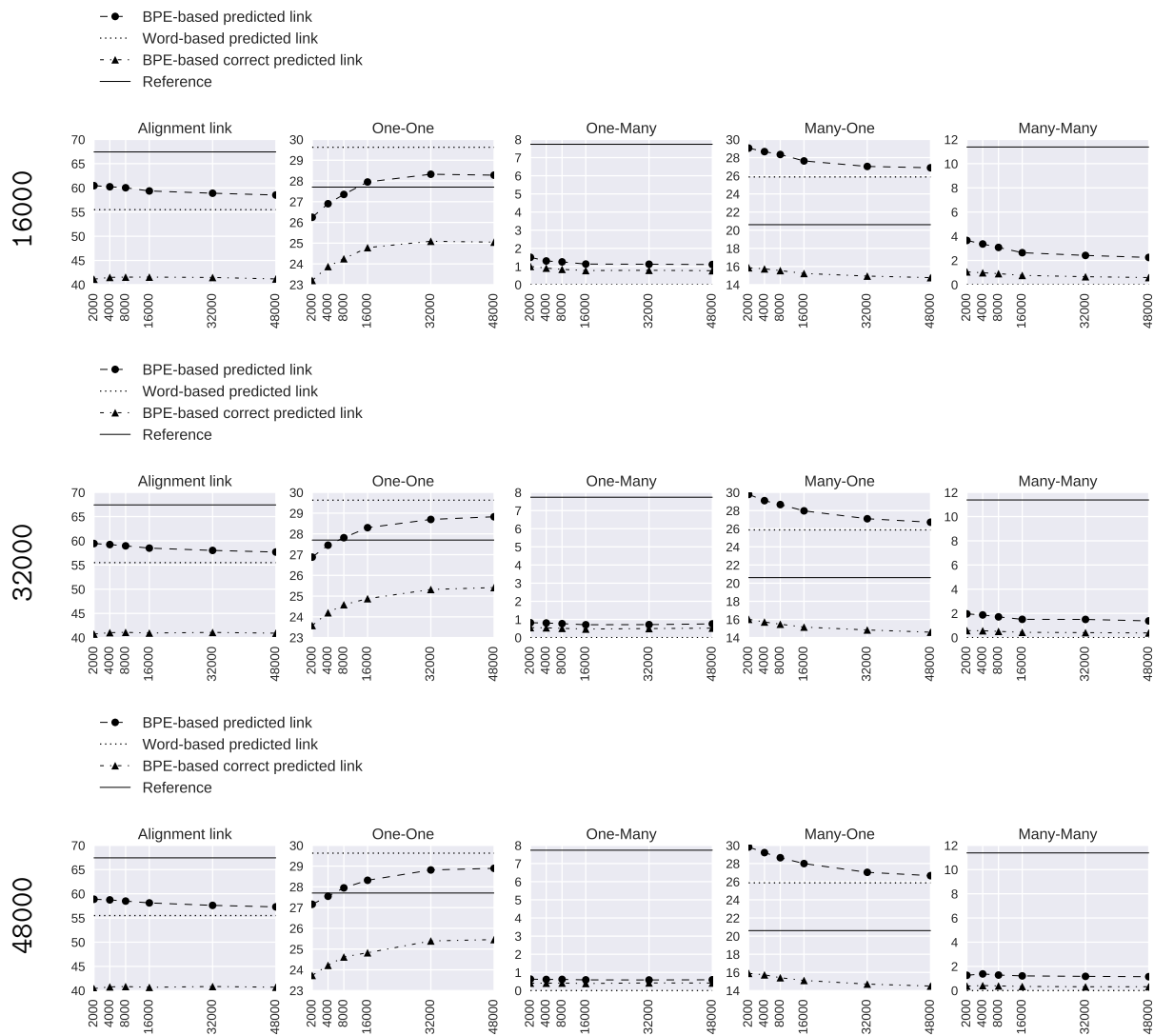
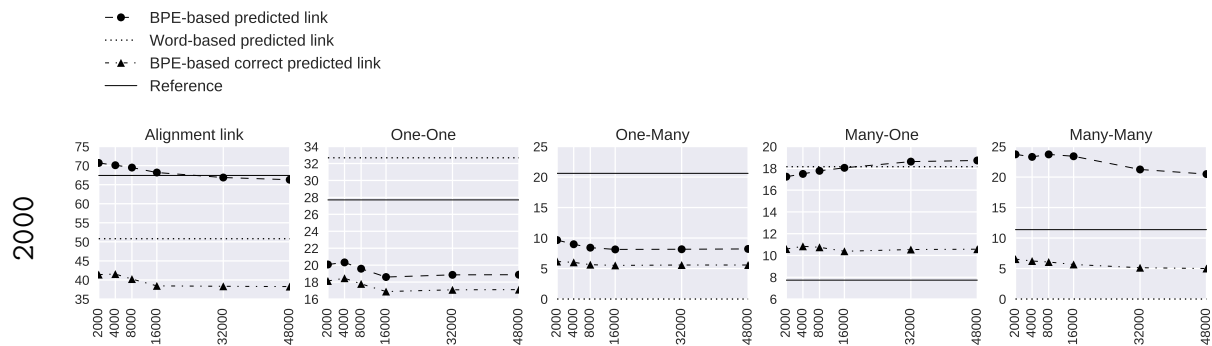


Figure E.47: BPE-based `Fastalign` for the direction English-Czech: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).



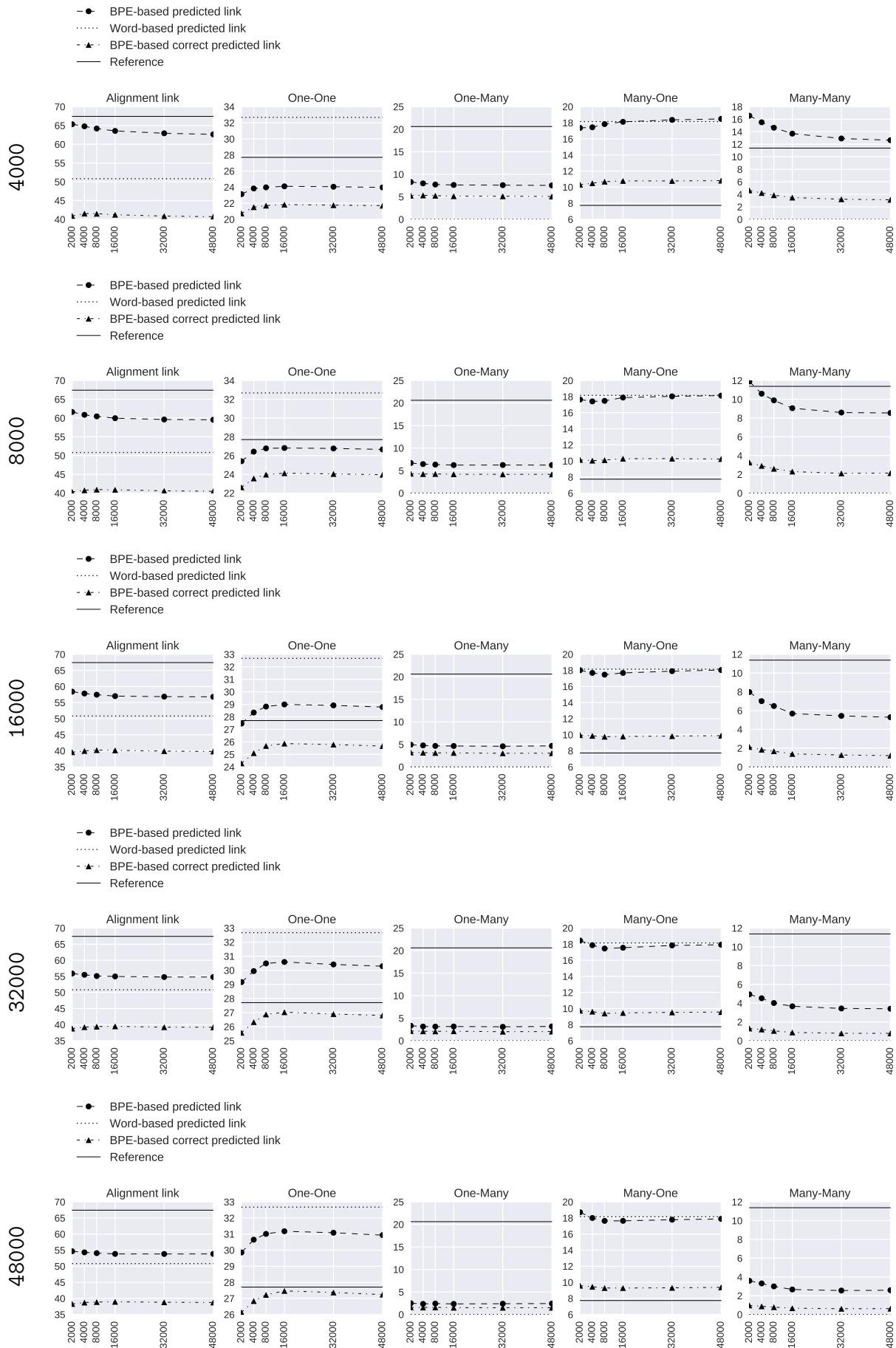
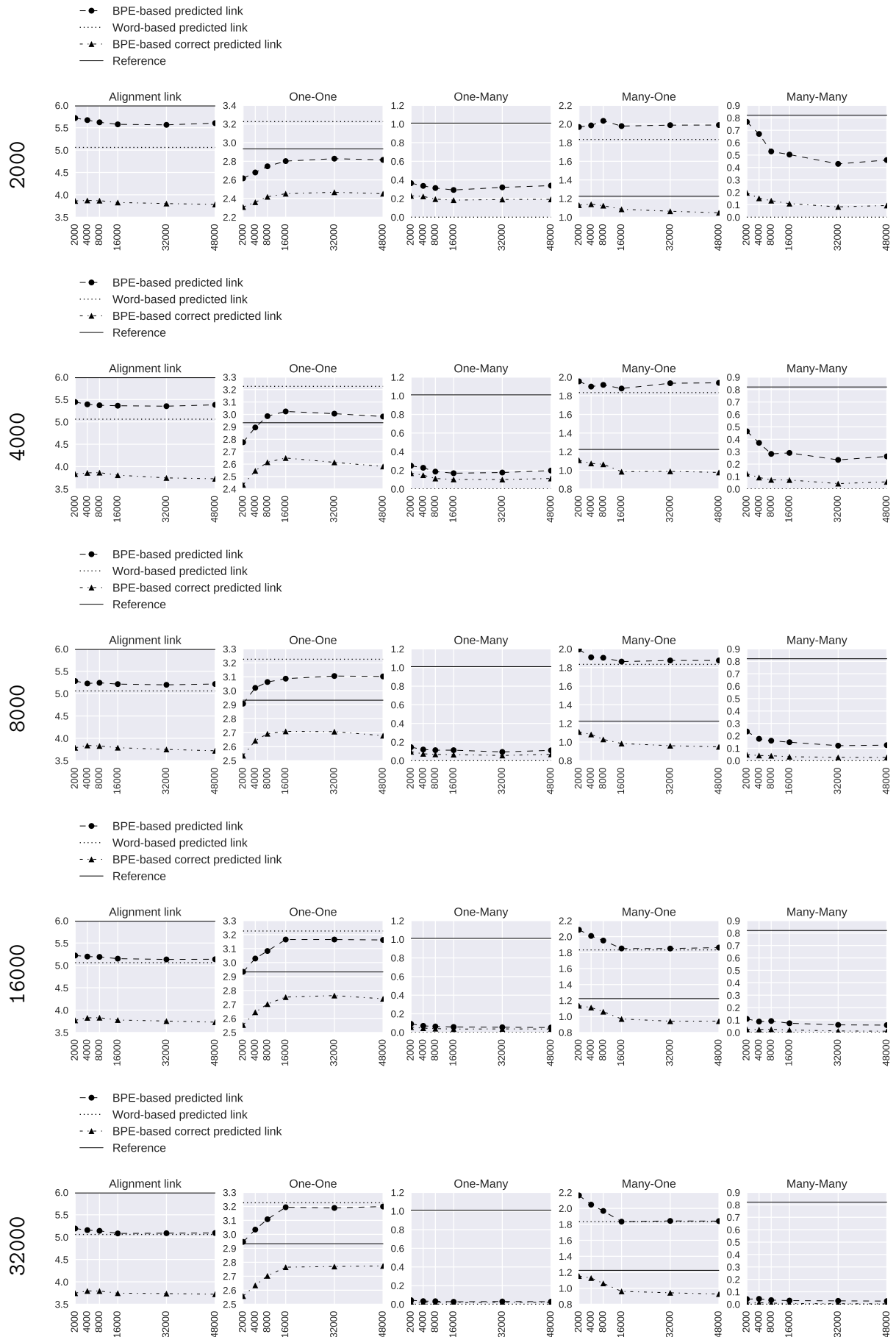


Figure E.48: BPE-based `Fastalign` for the direction Czech-English: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).



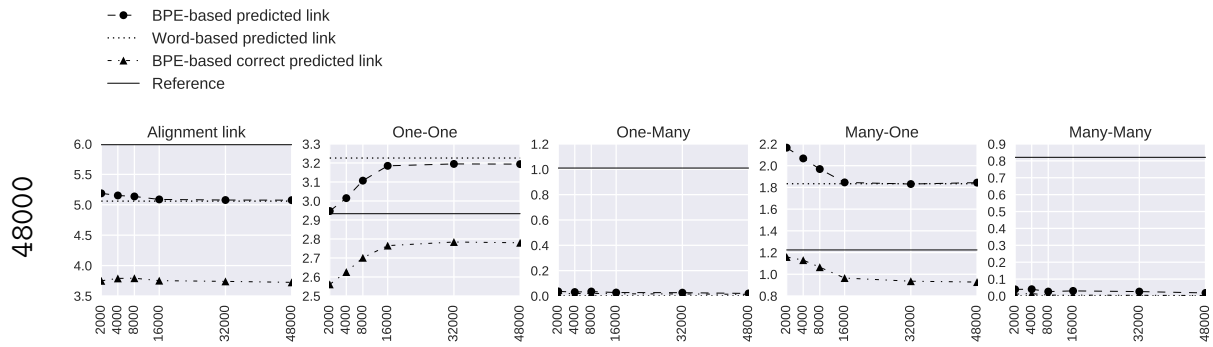
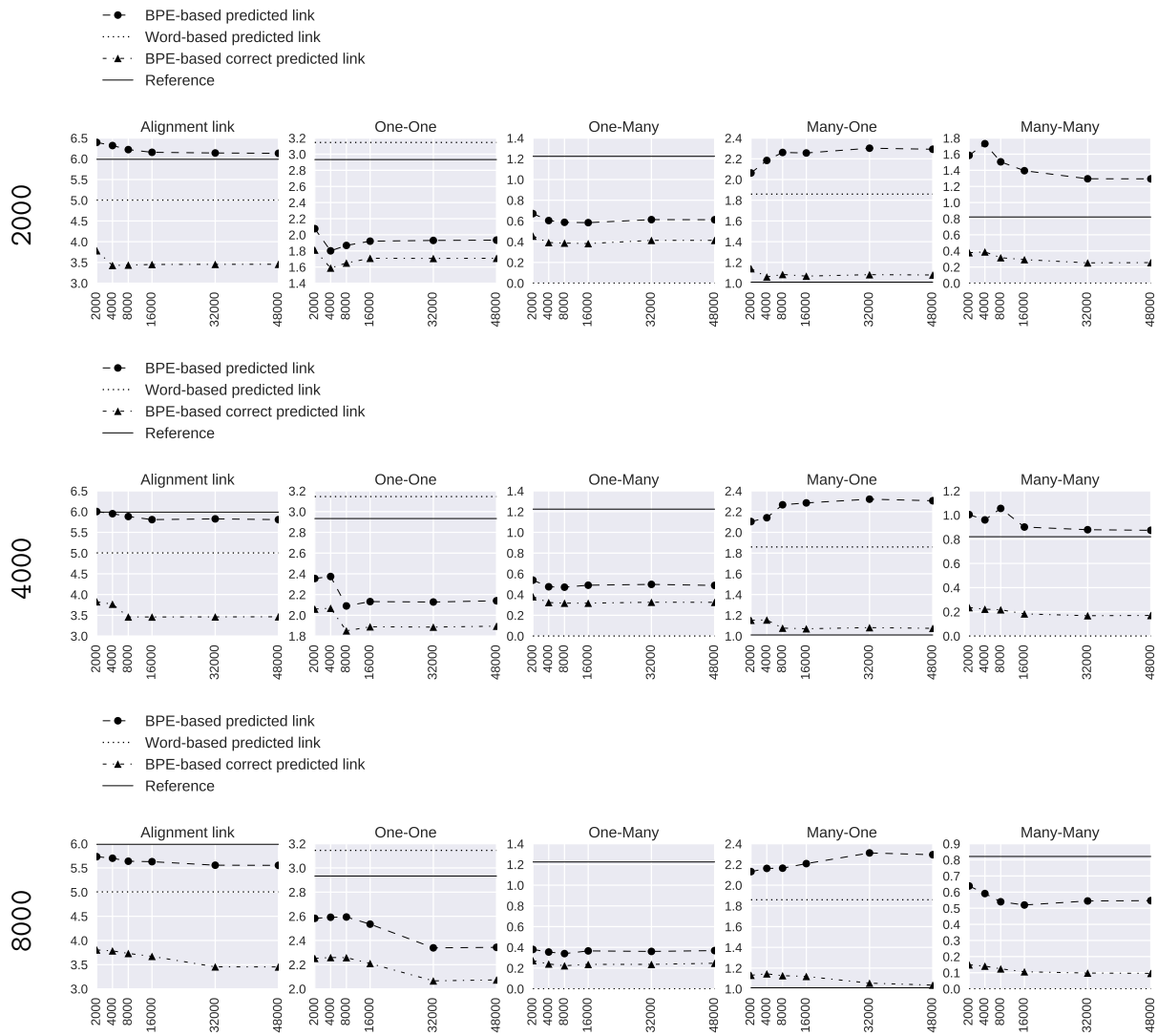


Figure E.49: BPE-based Fastalign for the direction English-Romanian: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).



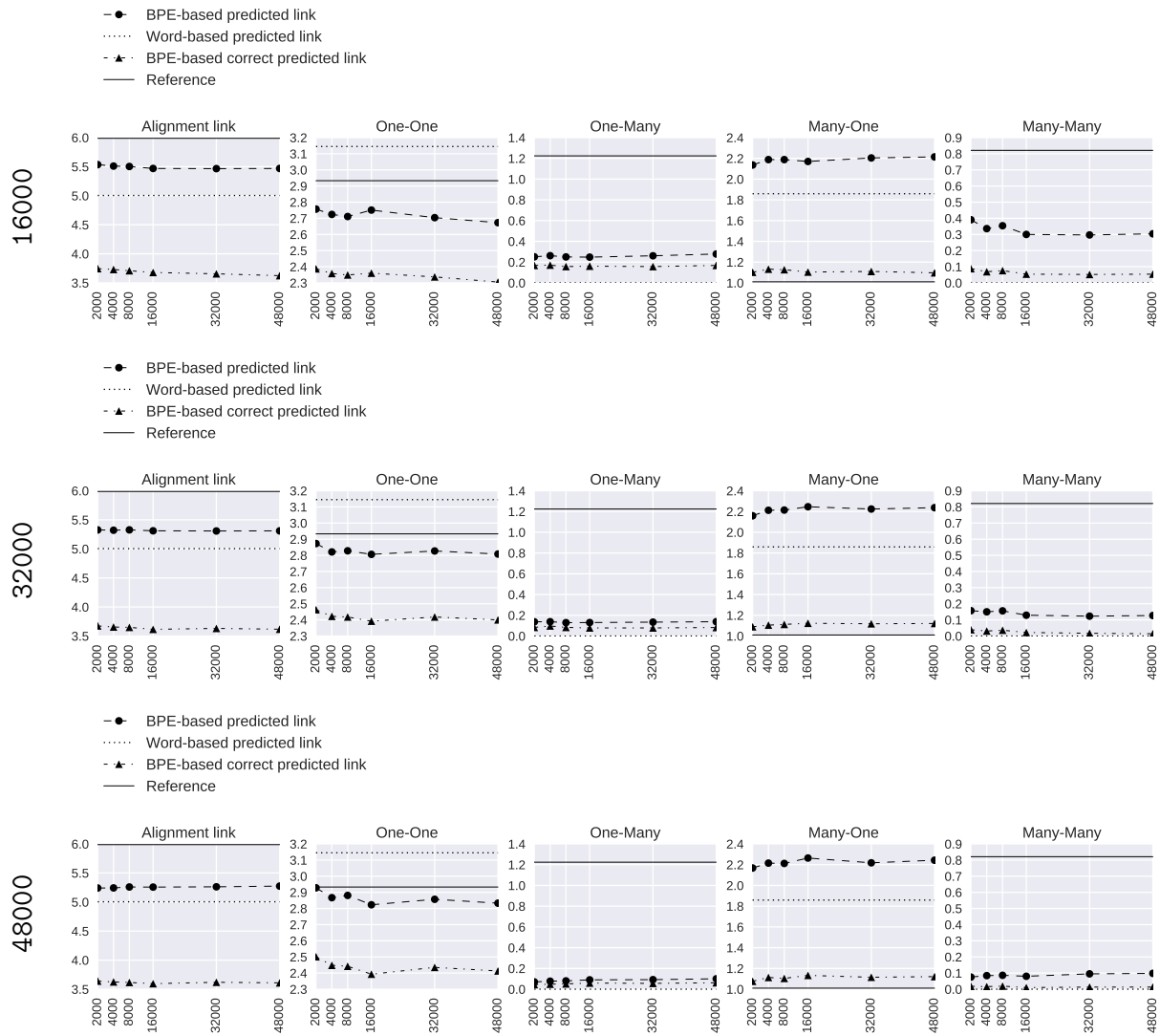
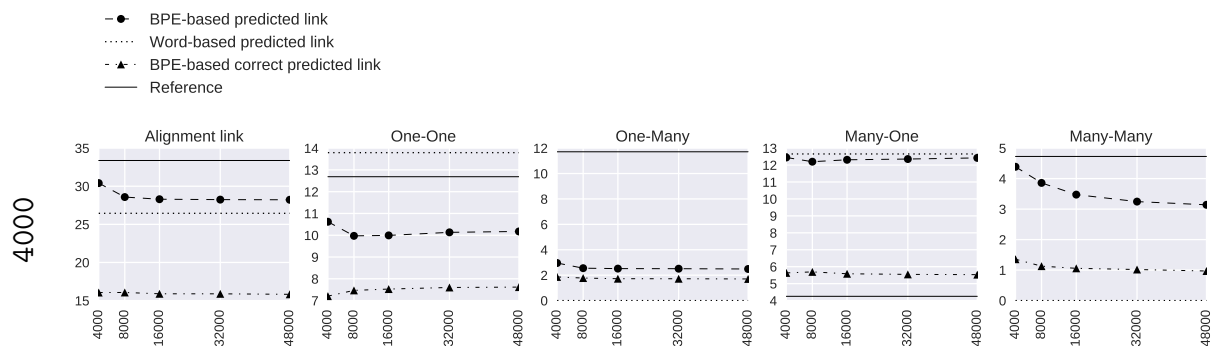


Figure E.50: BPE-based Fastalign for the direction Romanian-English: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).



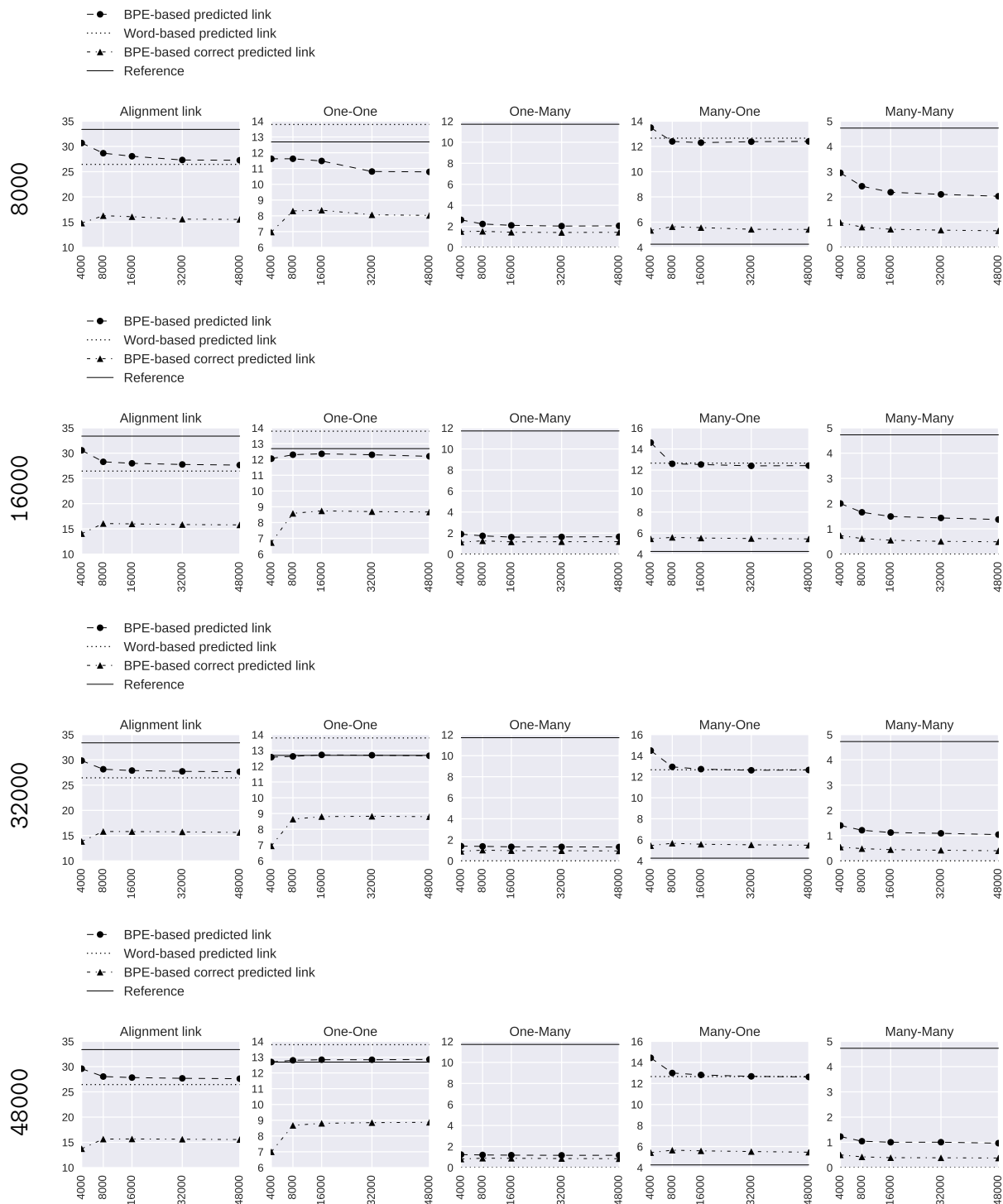


Figure E.51: BPE-based Fastalign for the direction English-Japanese: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).

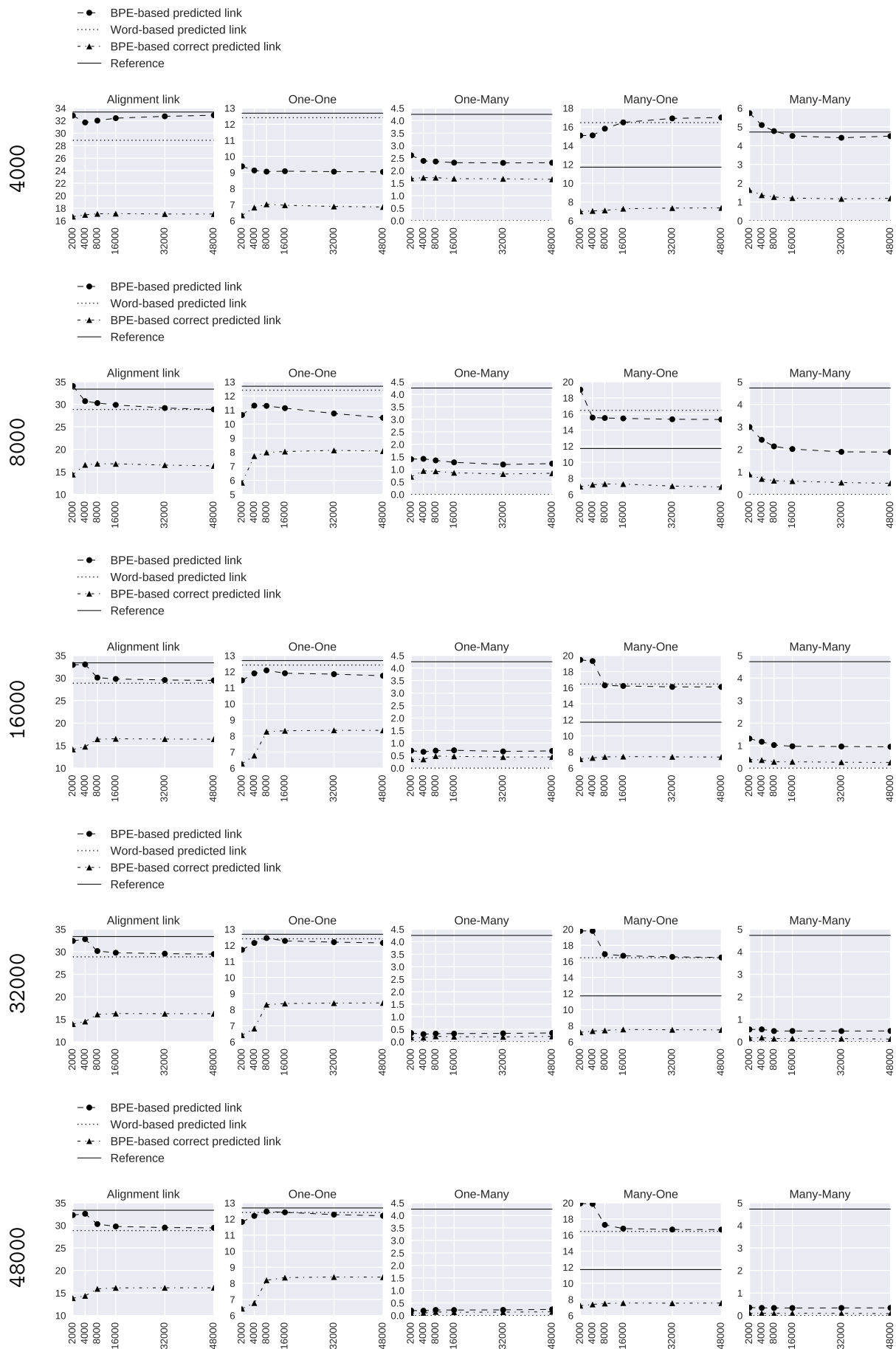


Figure E.52: BPE-based Fastalign for the direction Japanese-English: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).

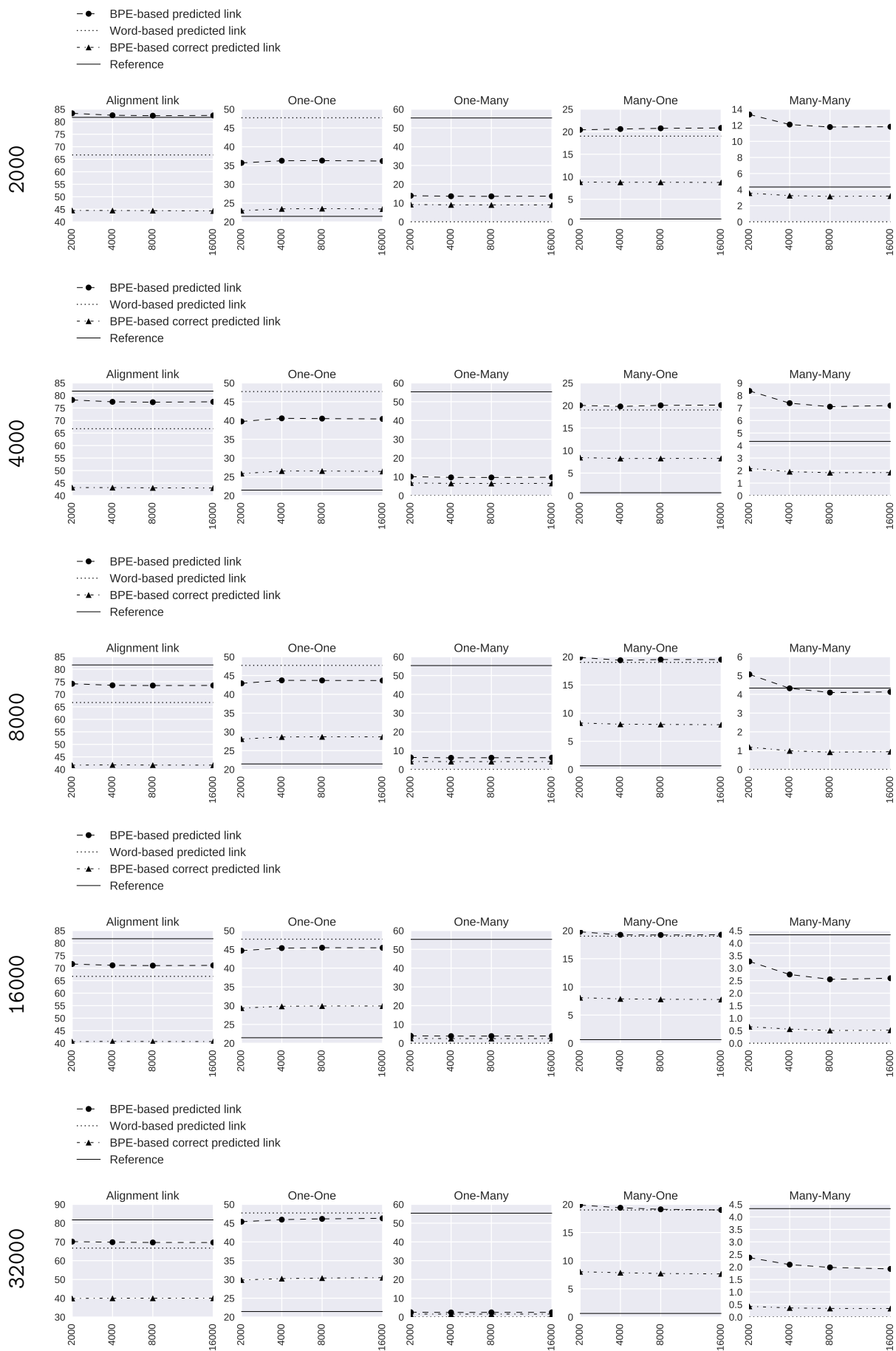


Figure E.53: BPE-based Fastalign for the direction English-Vietnamese: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).

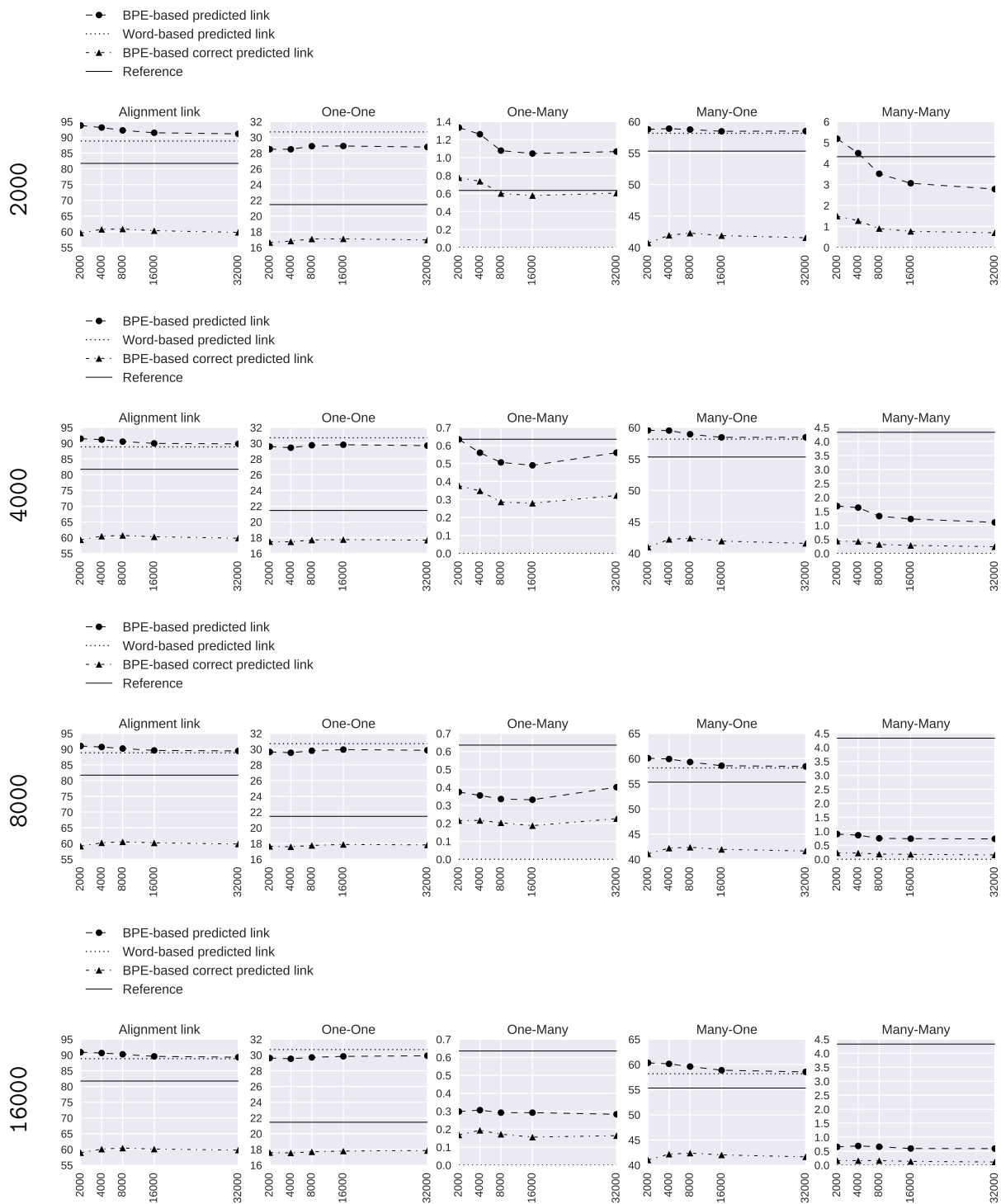


Figure E.54: BPE-based Fastalign for the direction Vietnamese-English: We observe the alignment types. For each source vocabulary size, we show the number of links as a function of the target vocabulary size. The y axis corresponds to the number of links ($\times 1000$).

E.5 Rare words

E.5.1 Matrix of scores

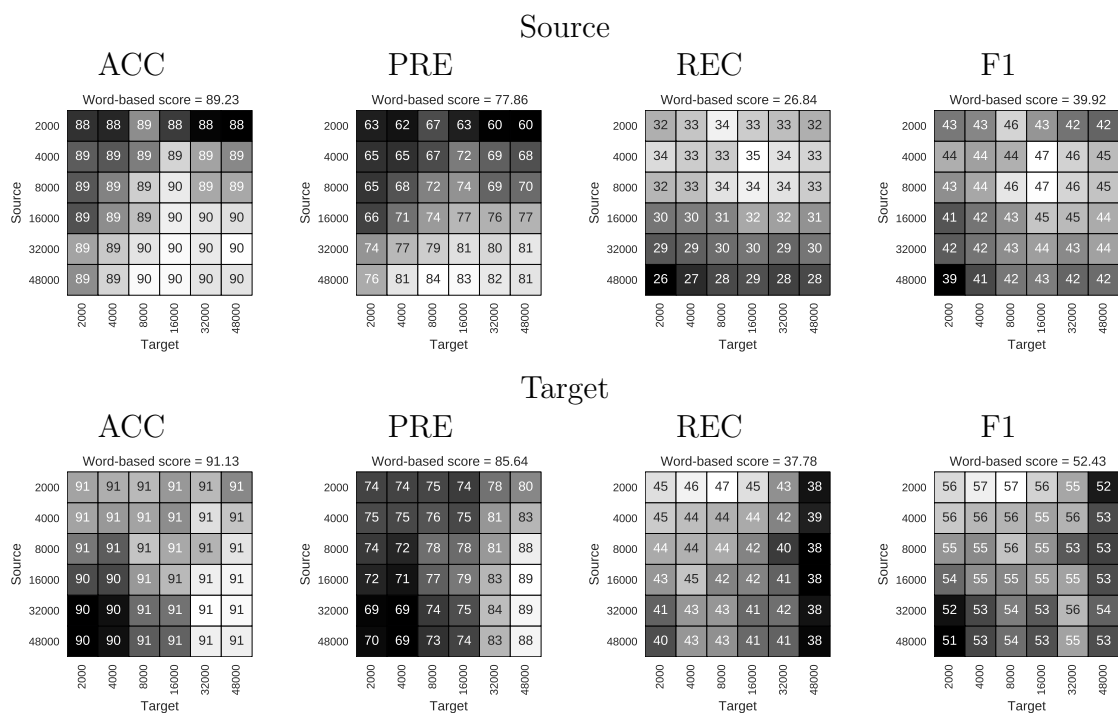


Figure E.55: BPE-based Fastalign for the direction English-French: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.

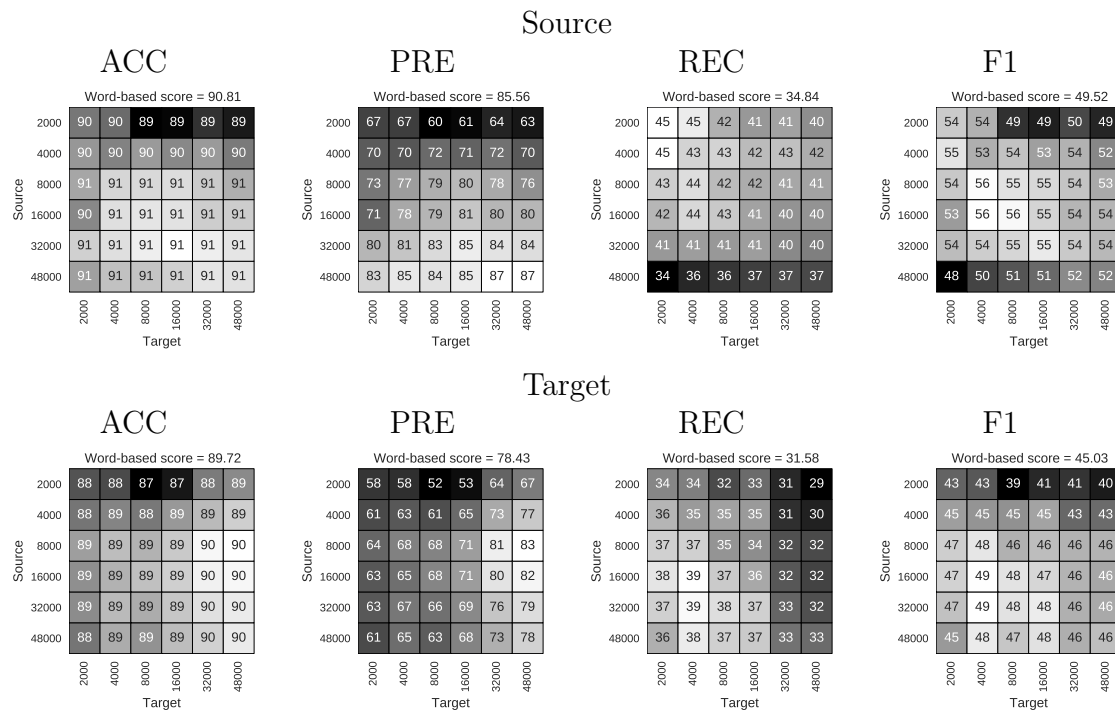


Figure E.56: BPE-based *Fastalign* for the direction French-English: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.

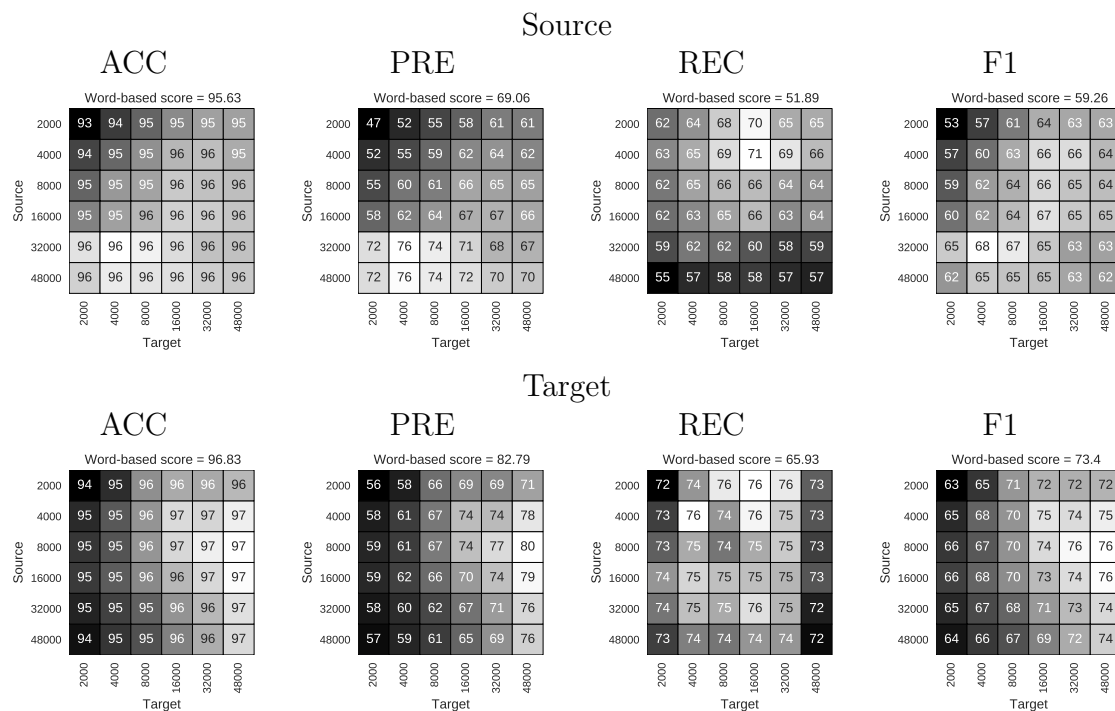


Figure E.57: BPE-based *Fastalign* for the direction English-German: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.

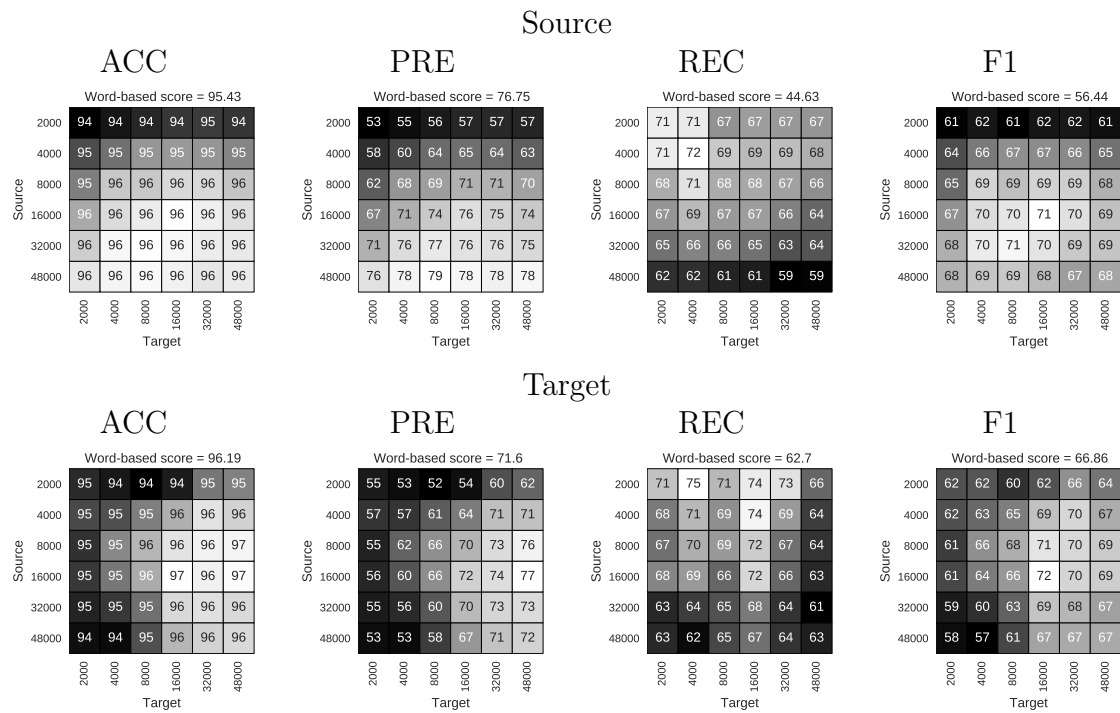


Figure E.58: BPE-based Fastalign for the direction German-English: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.

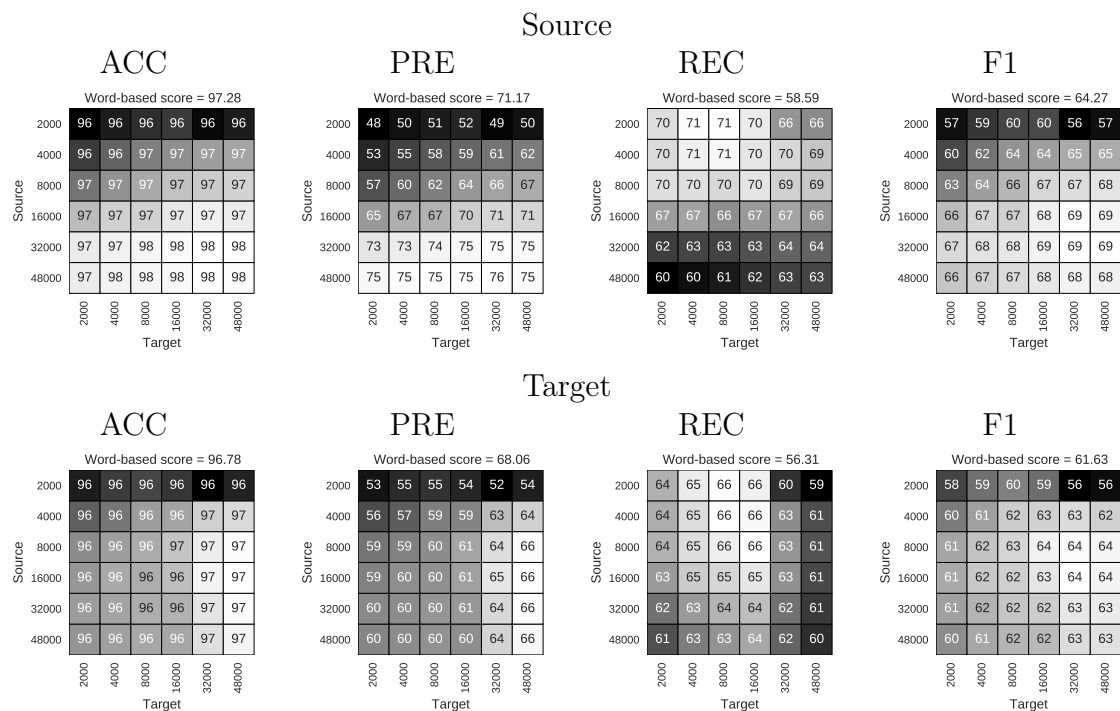


Figure E.59: BPE-based Fastalign for the direction English-Czech: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.

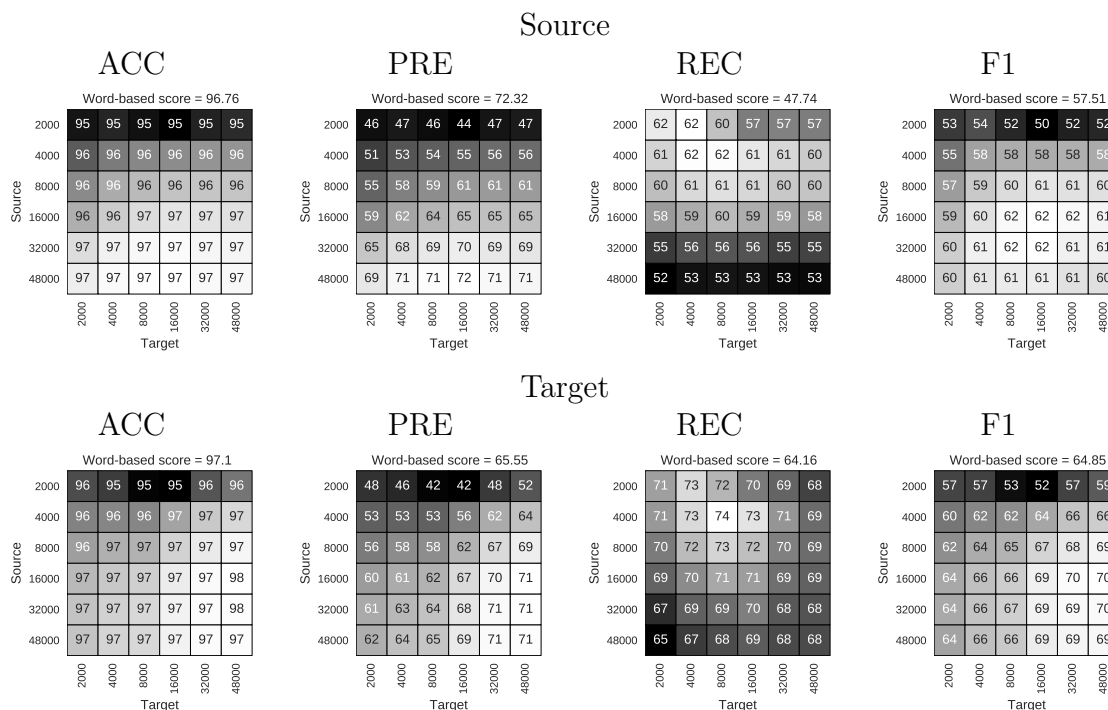


Figure E.60: BPE-based **Fastalign** for the direction Czech-English: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.

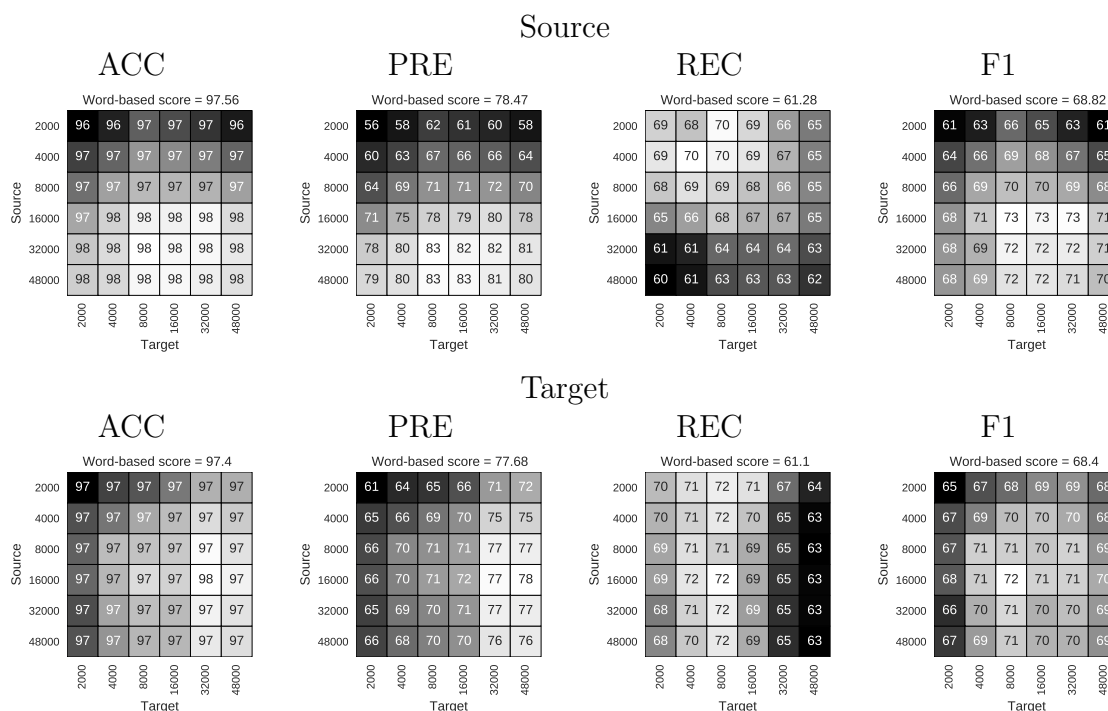


Figure E.61: BPE-based **Fastalign** for the direction English-Romanian: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.

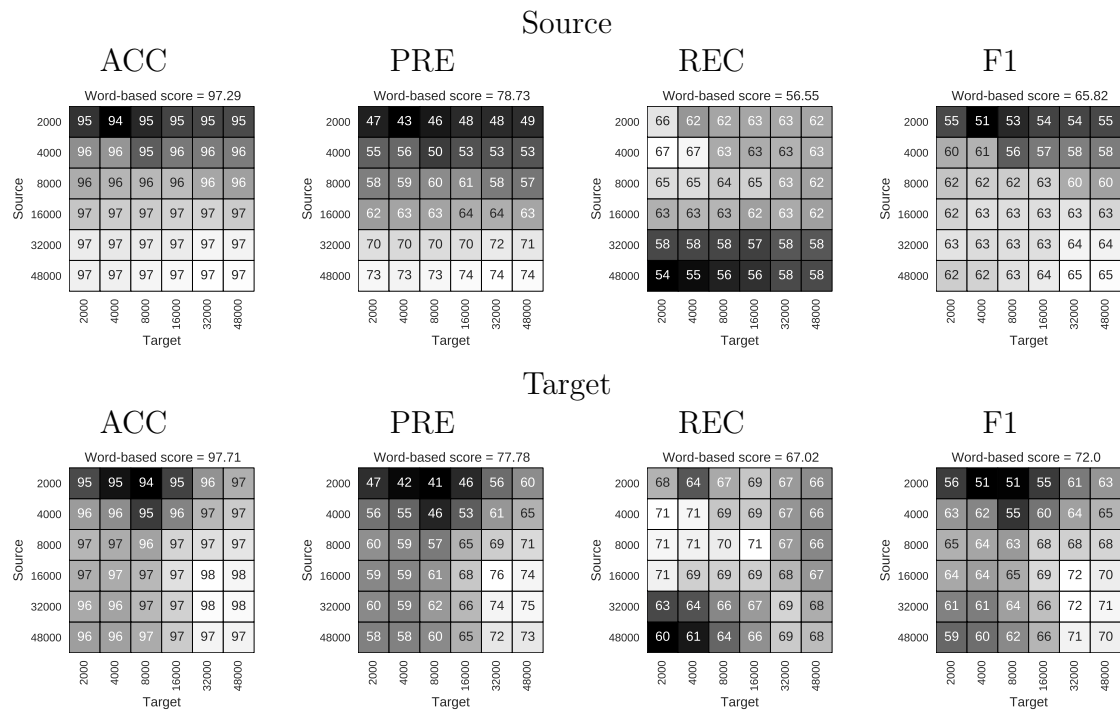


Figure E.62: BPE-based **Fastalign** for the direction Romanian-English: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.

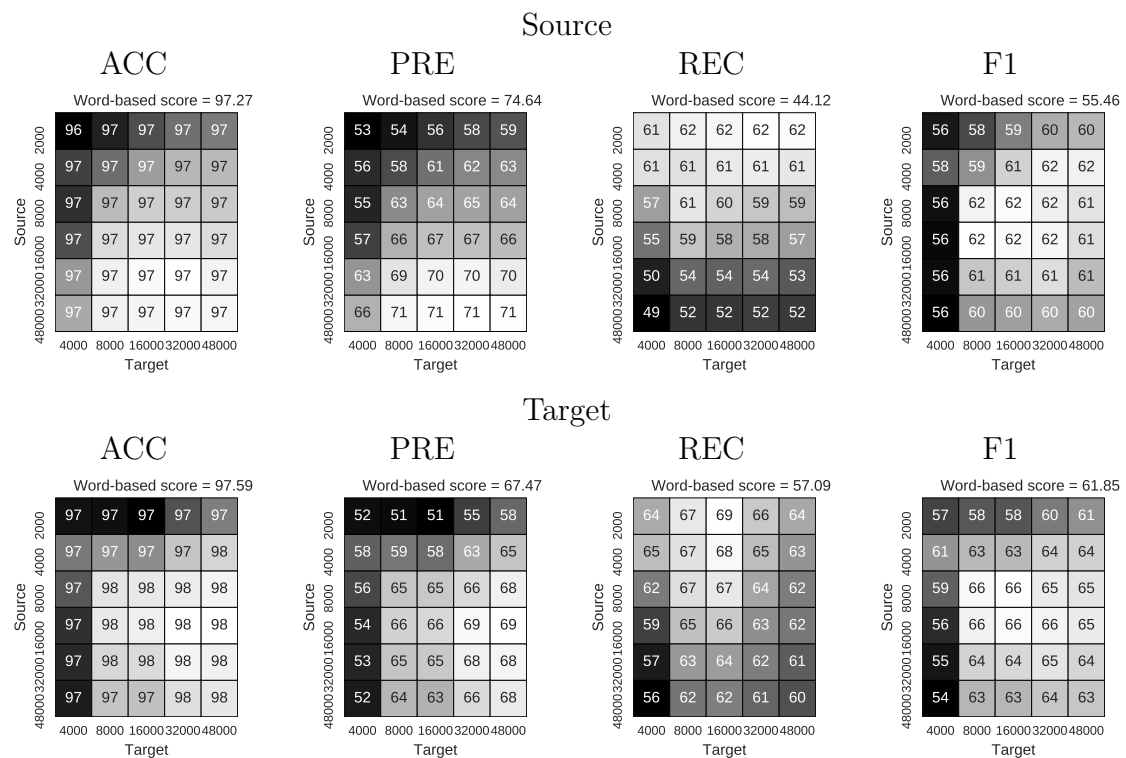


Figure E.63: BPE-based **Fastalign** for the direction English-Japanese: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.

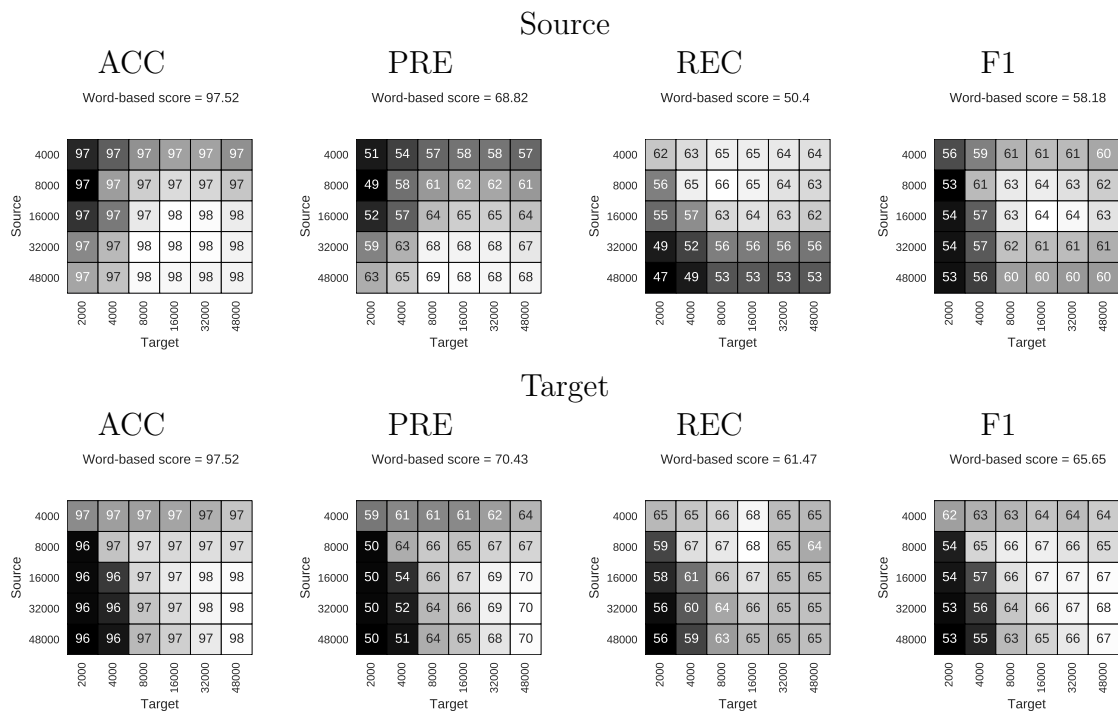


Figure E.64: BPE-based `Fastalign` for the direction Japanese-English: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.

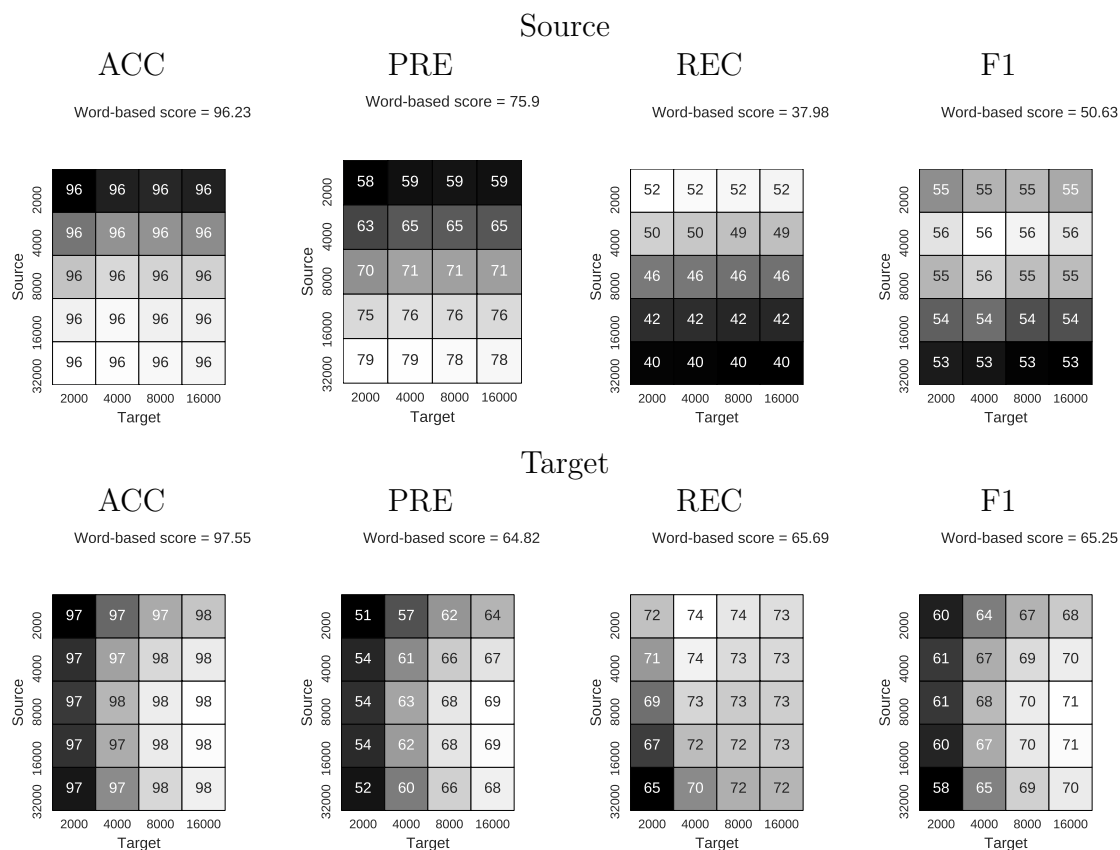


Figure E.65: BPE-based `Fastalign` for the direction English-Vietnamese: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.

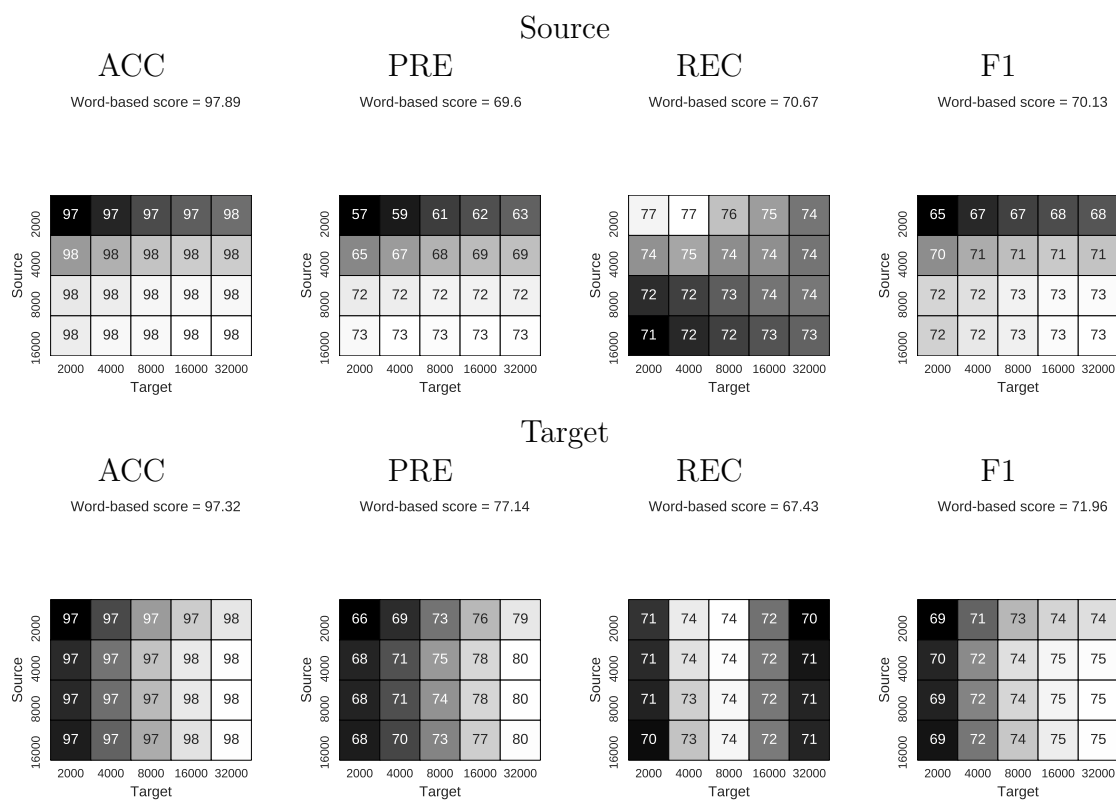
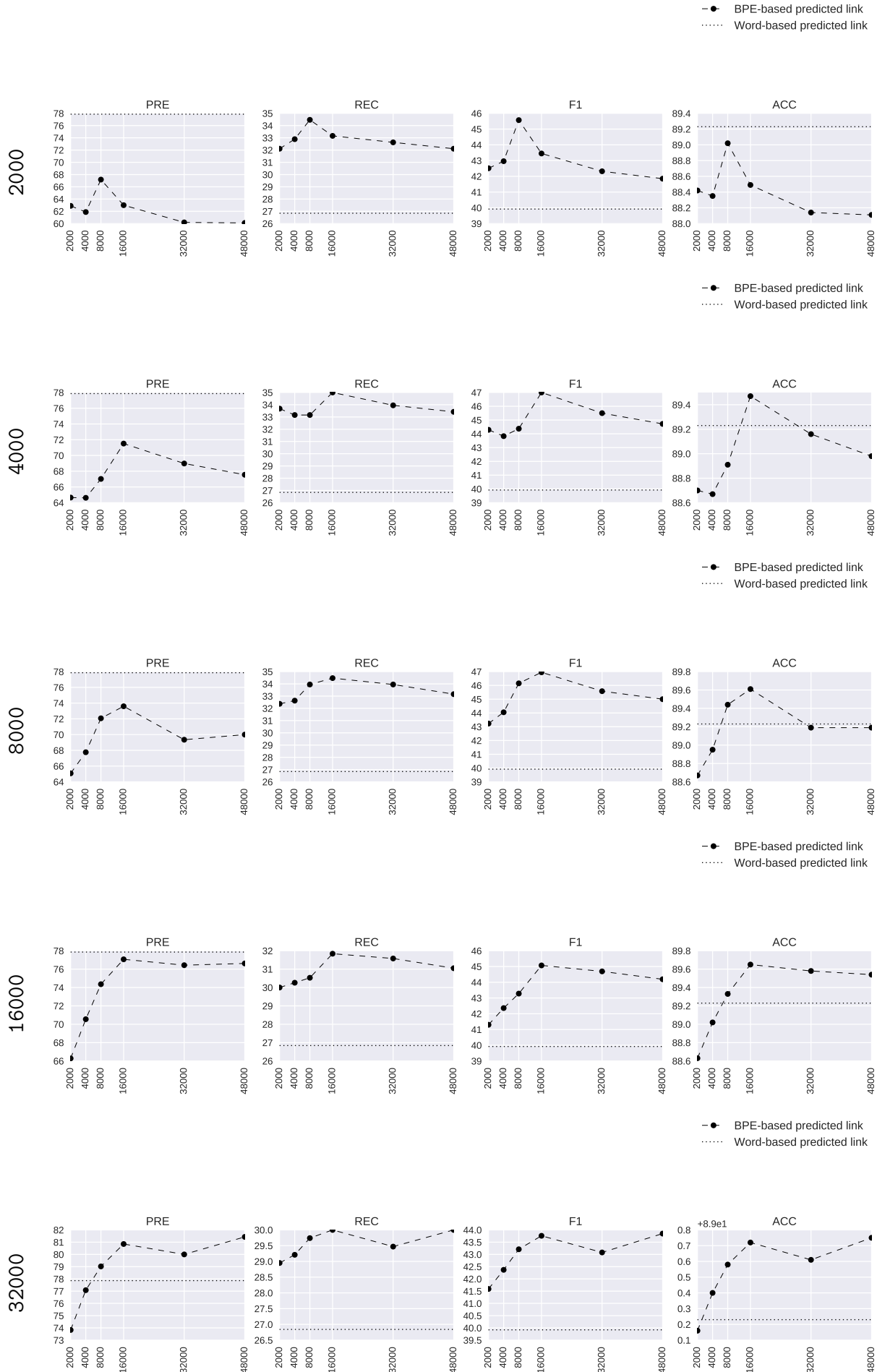


Figure E.66: BPE-based Fastalign for the direction Vietnamese-English: We observe the scores for rare source words (the top graphs) and for rare target words (the bottom graphs). For each source vocabulary size, we show the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of target vocabulary size.

E.5.2 Scores of BPE-based vs word-based alignment



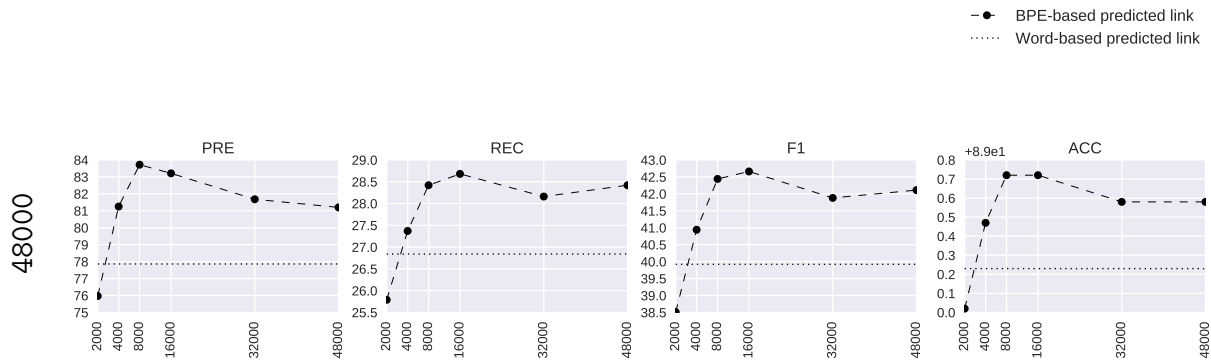
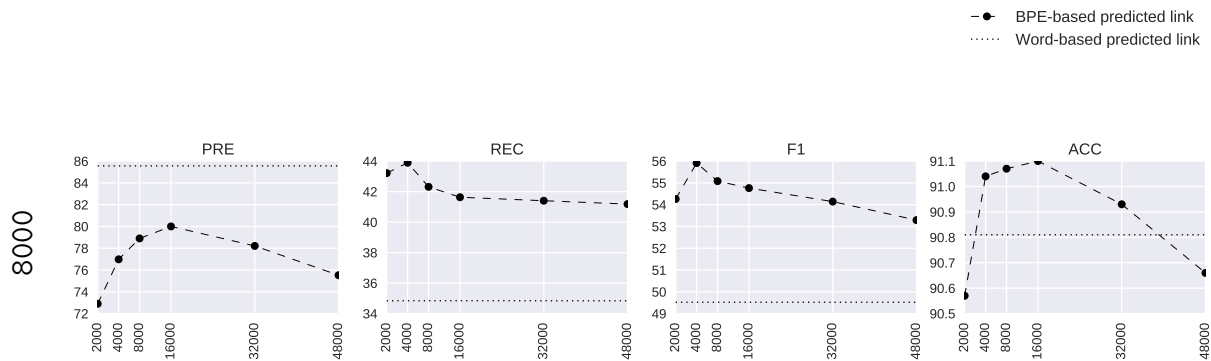
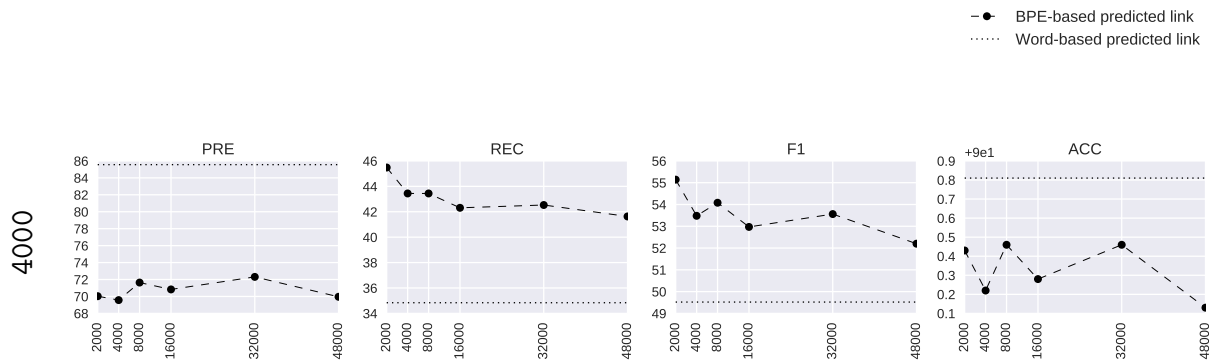
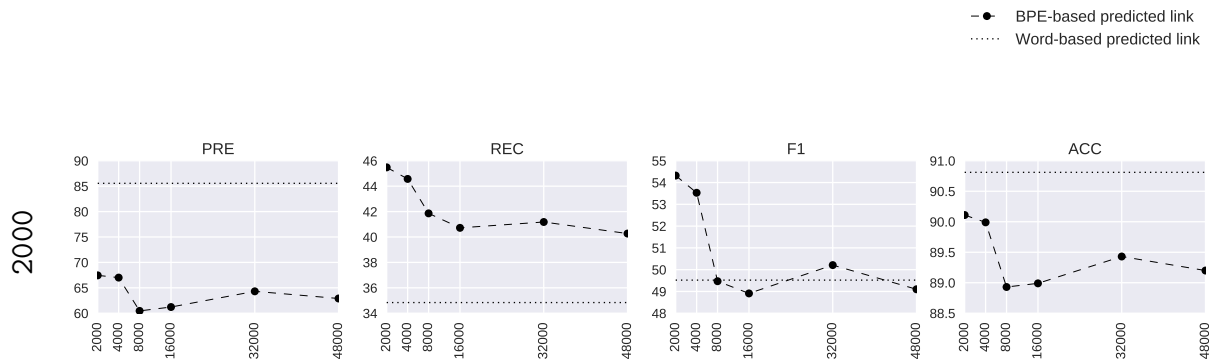


Figure E.67: BPE-based Fastalign for the direction English-French: We observe the scores for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.



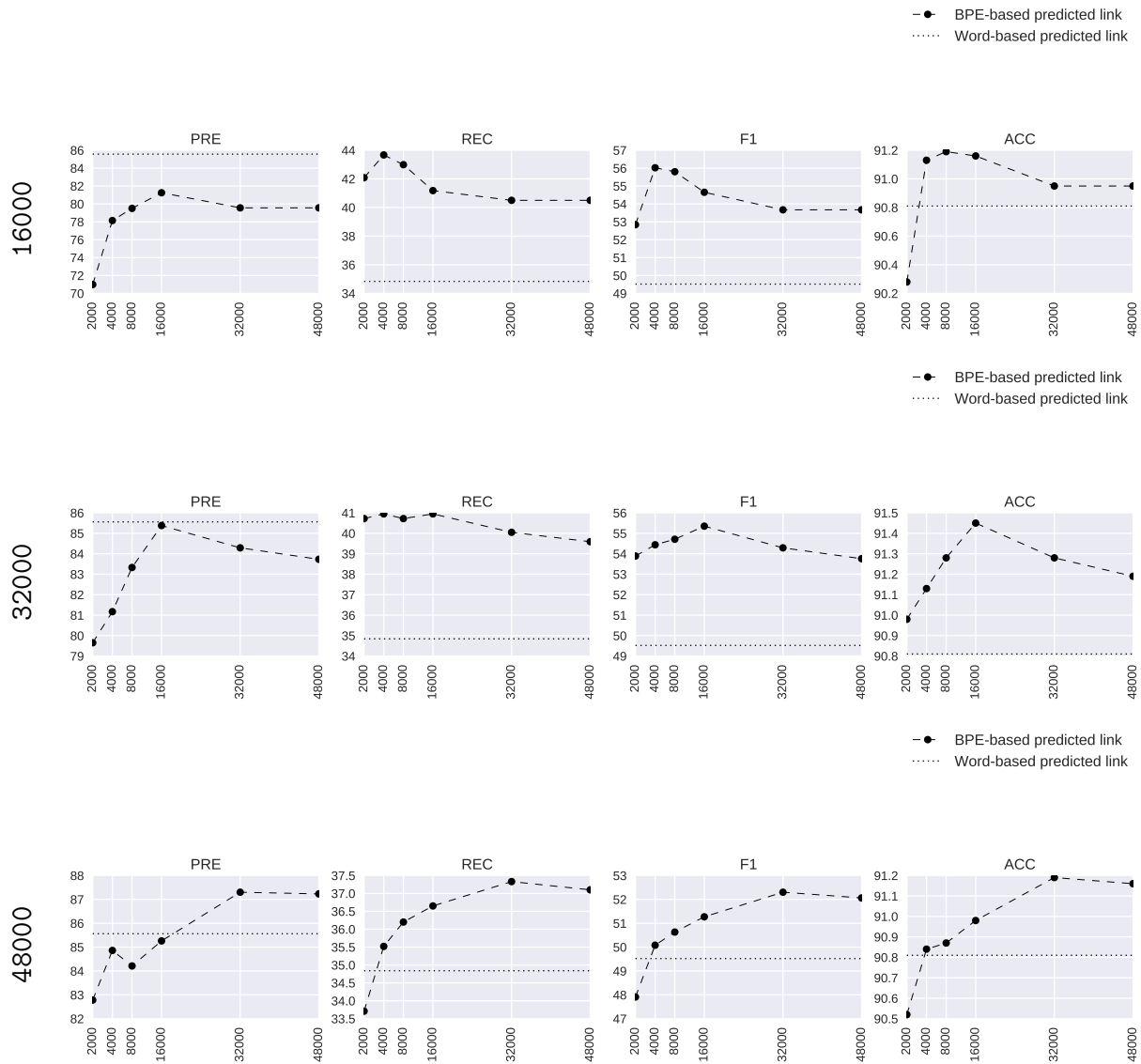
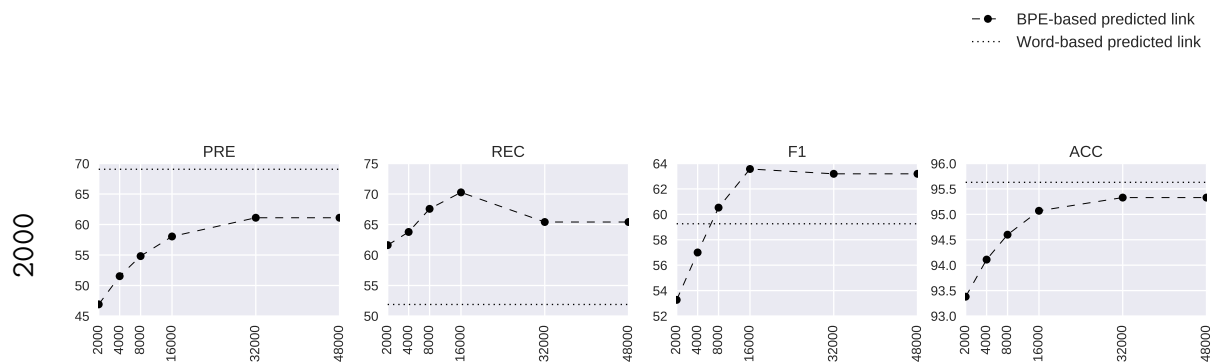


Figure E.68: BPE-based Fastalign for the direction French-English: We observe the scores for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.



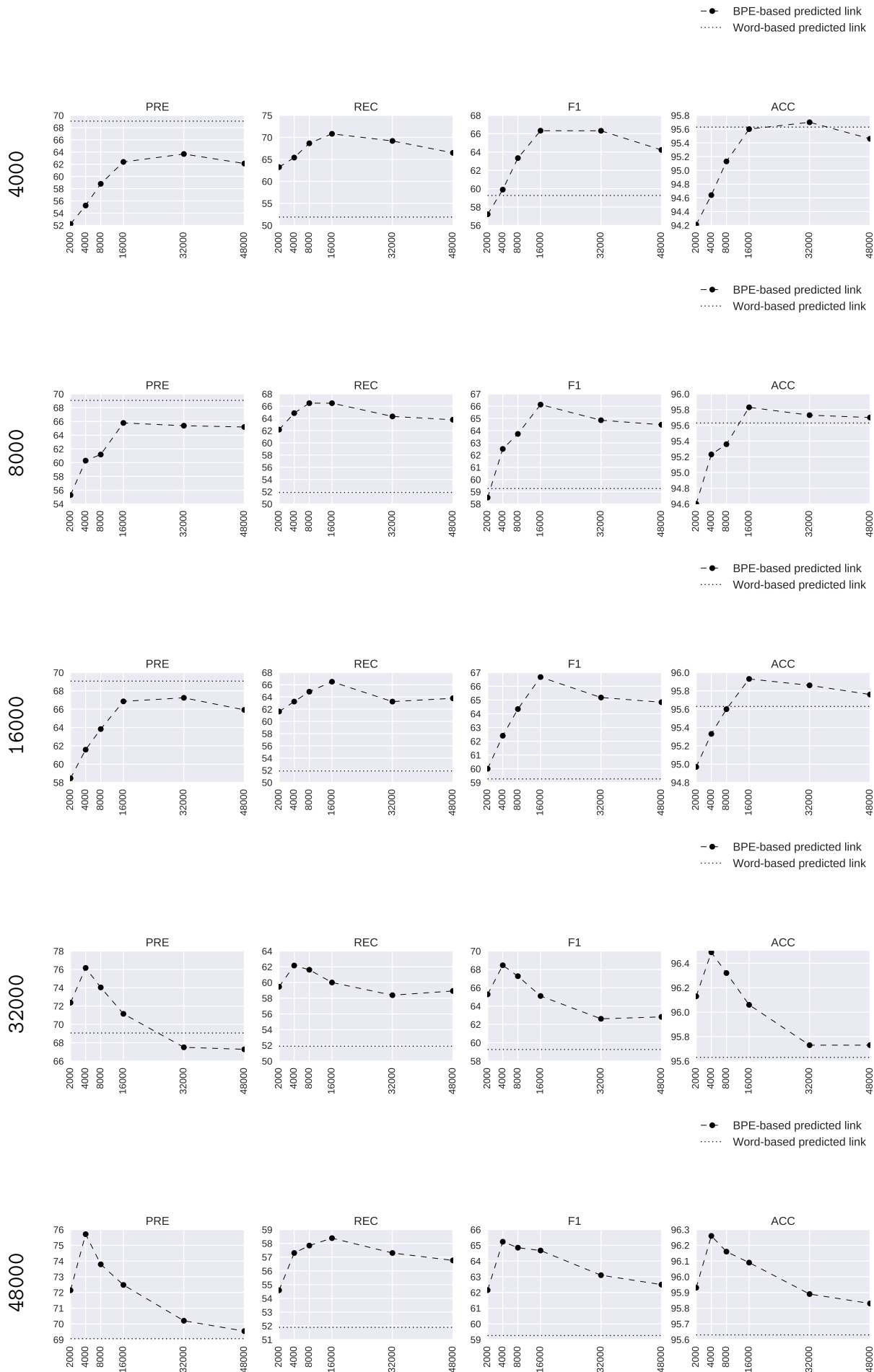
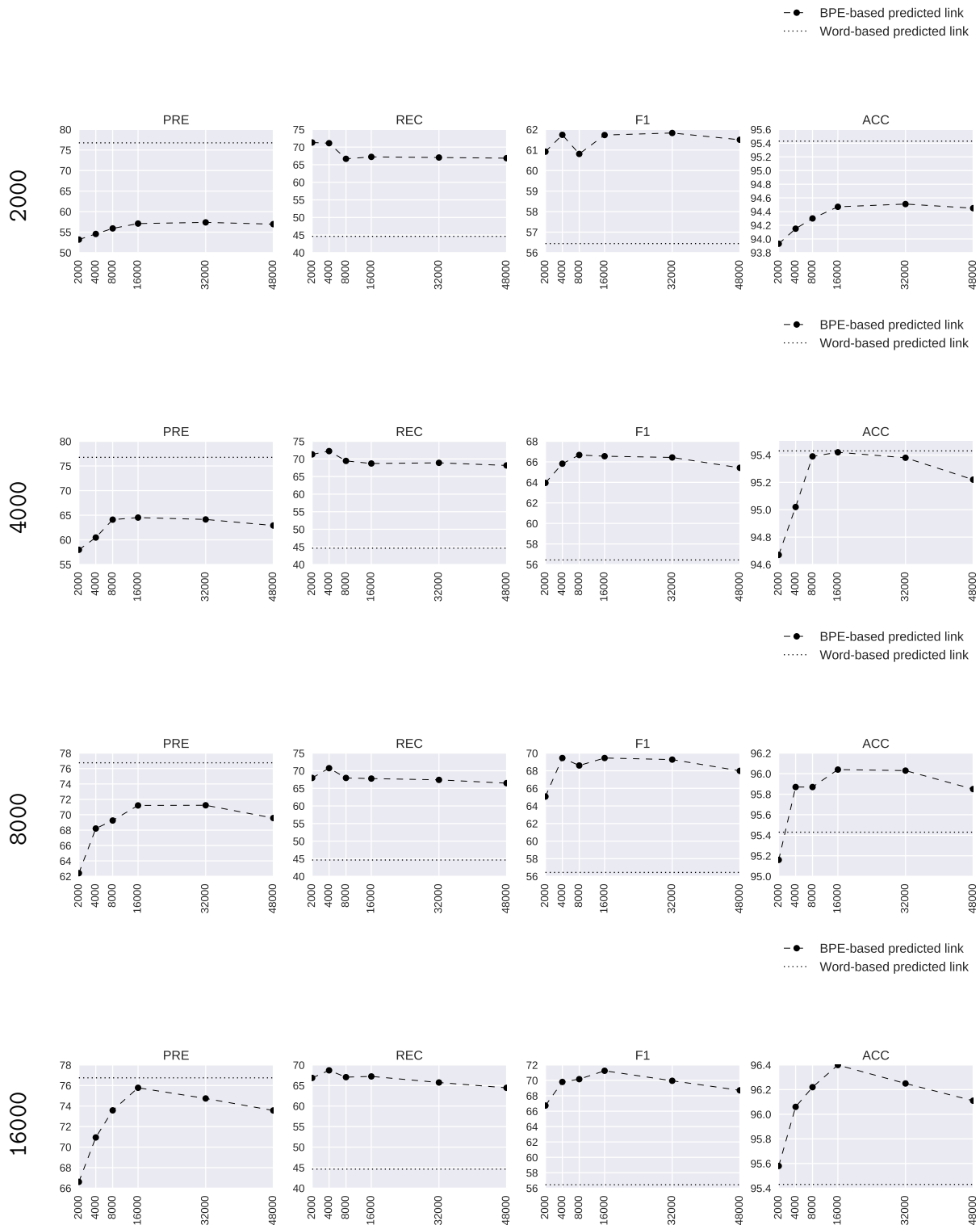


Figure E.69: BPE-based Fastalign for the direction English-German: We observe the scores

for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.



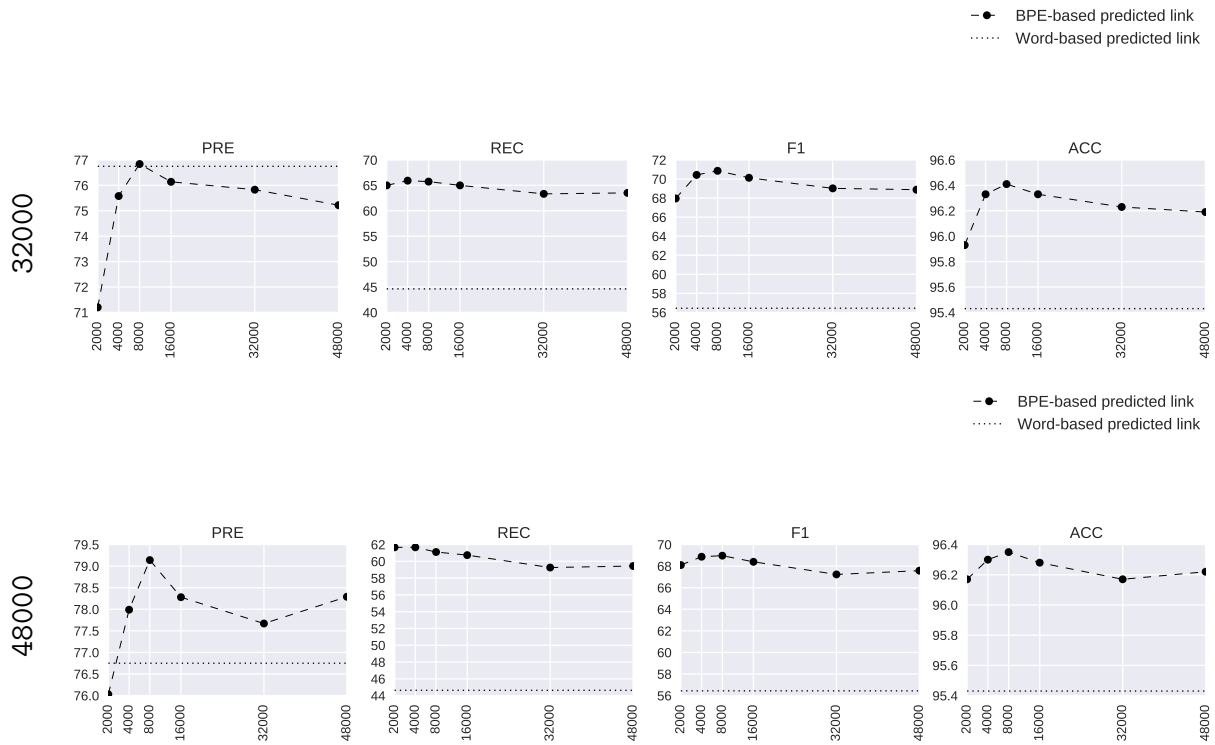
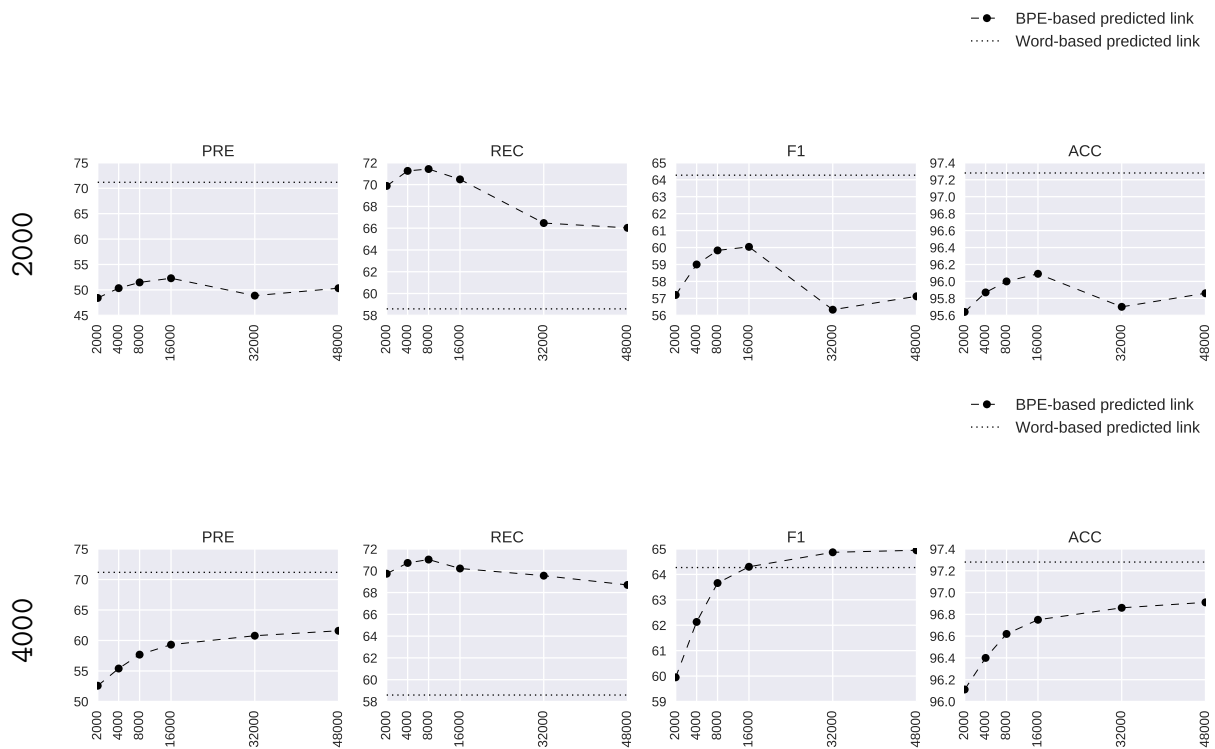


Figure E.70: BPE-based Fastalign for the direction German-English: We observe the scores for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.



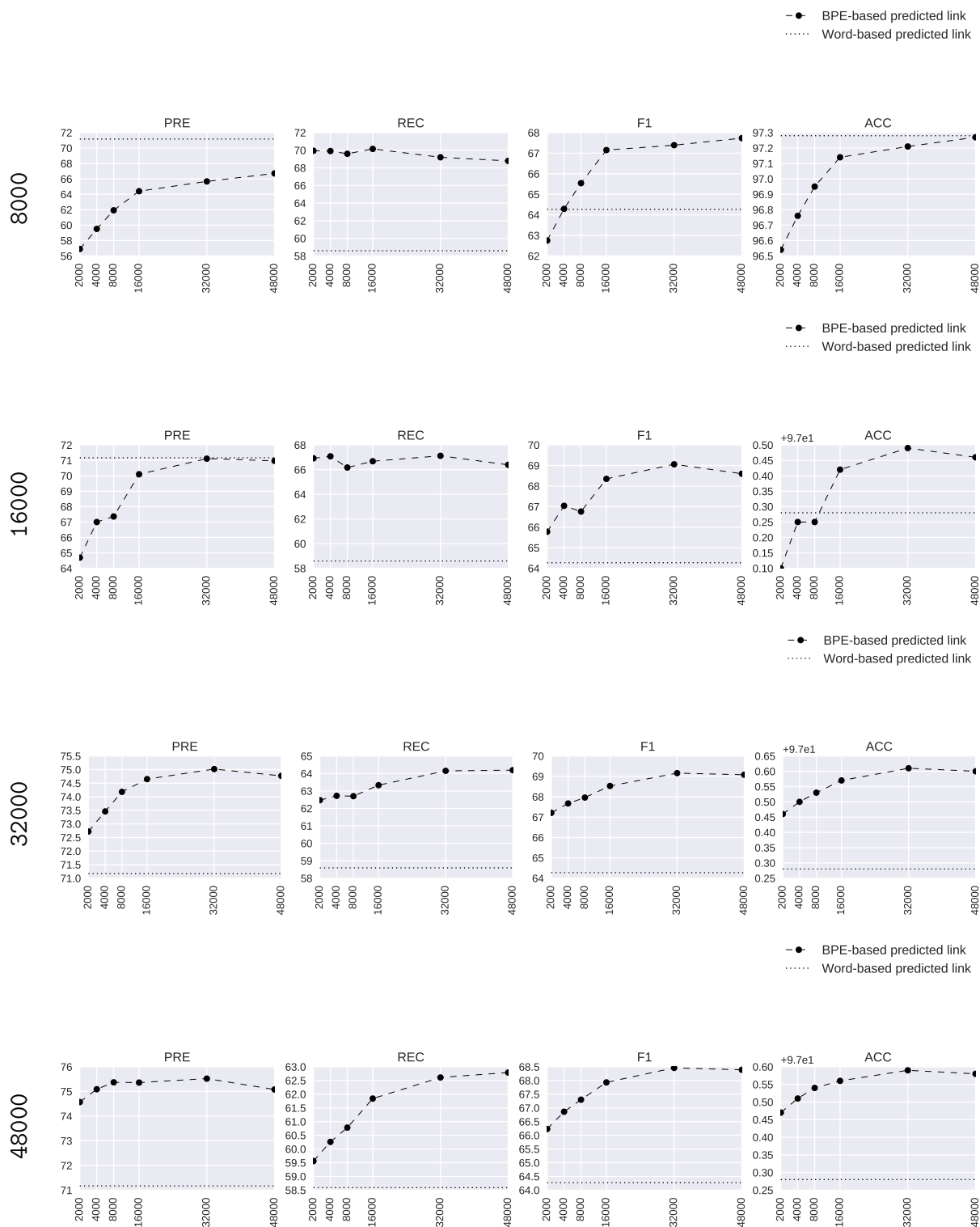
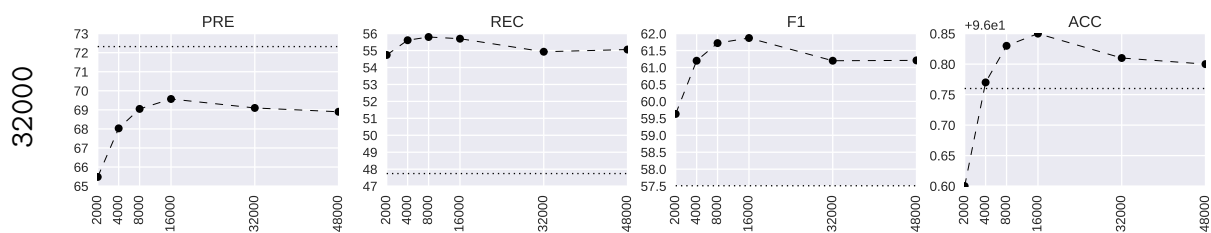
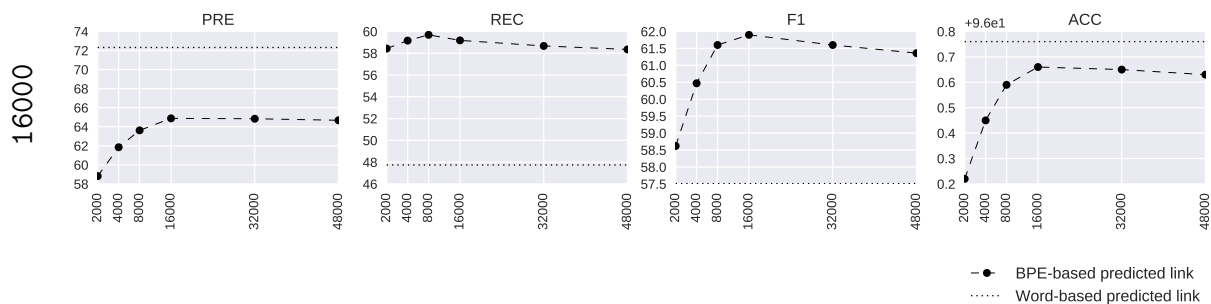
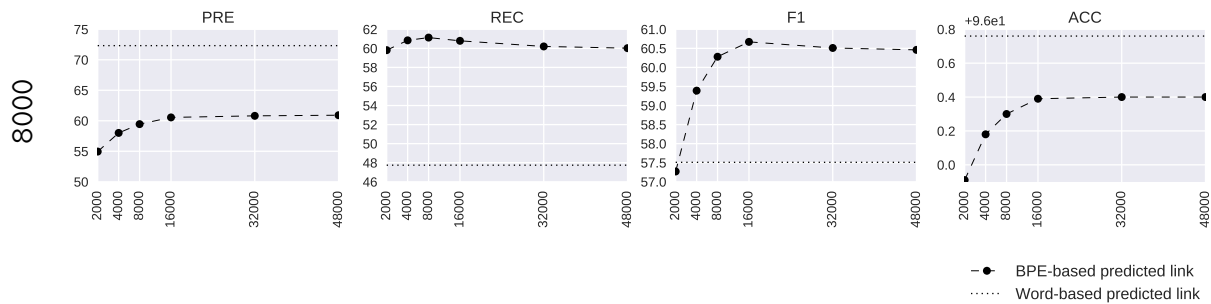
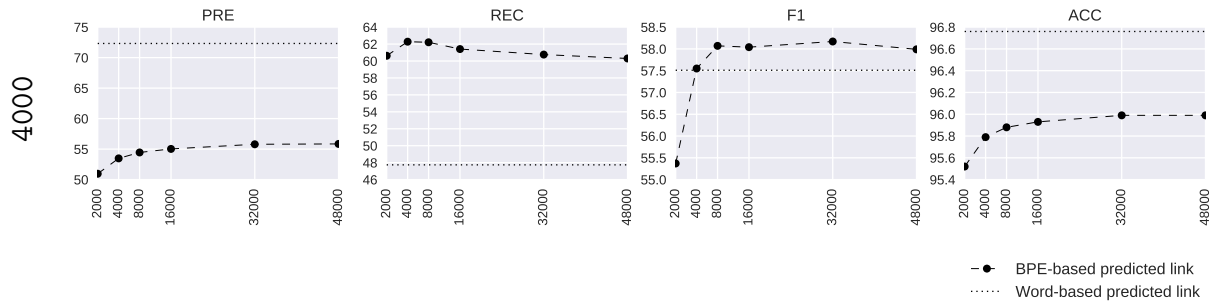
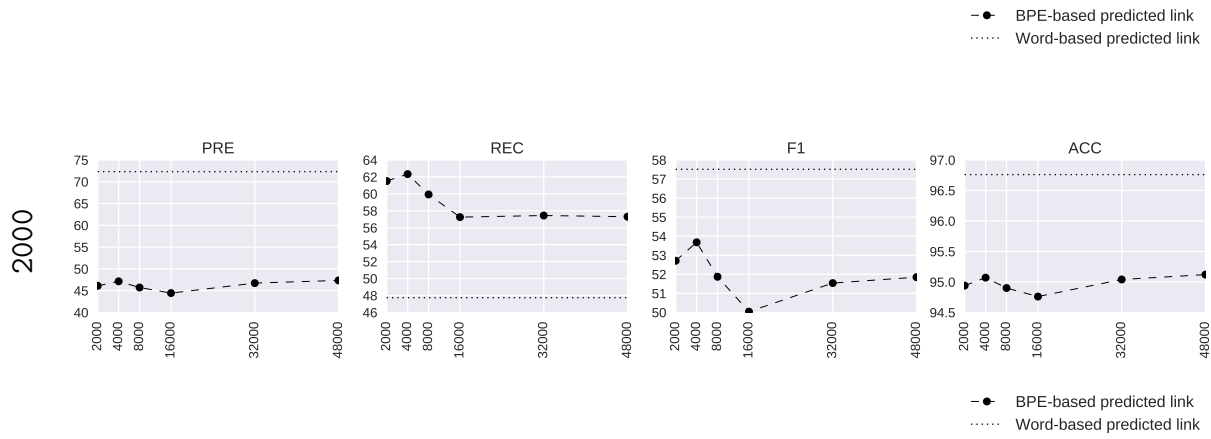


Figure E.71: BPE-based Fastalign for the direction English-Czech: We observe the scores for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.



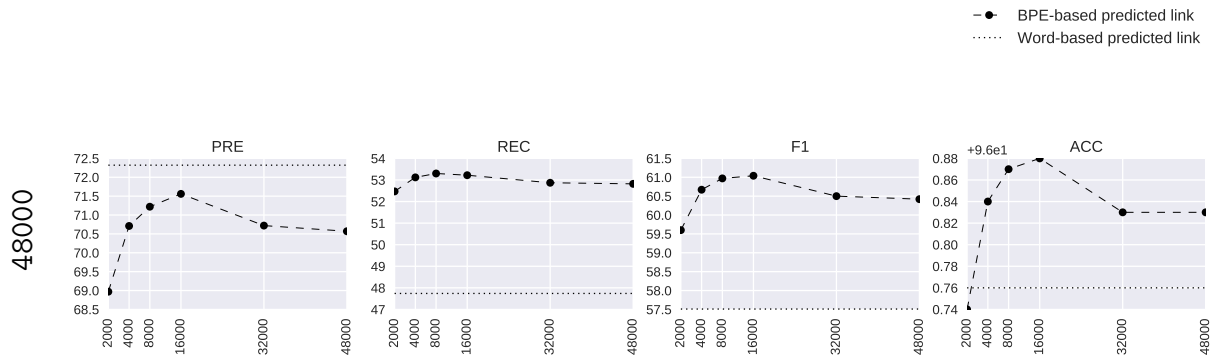
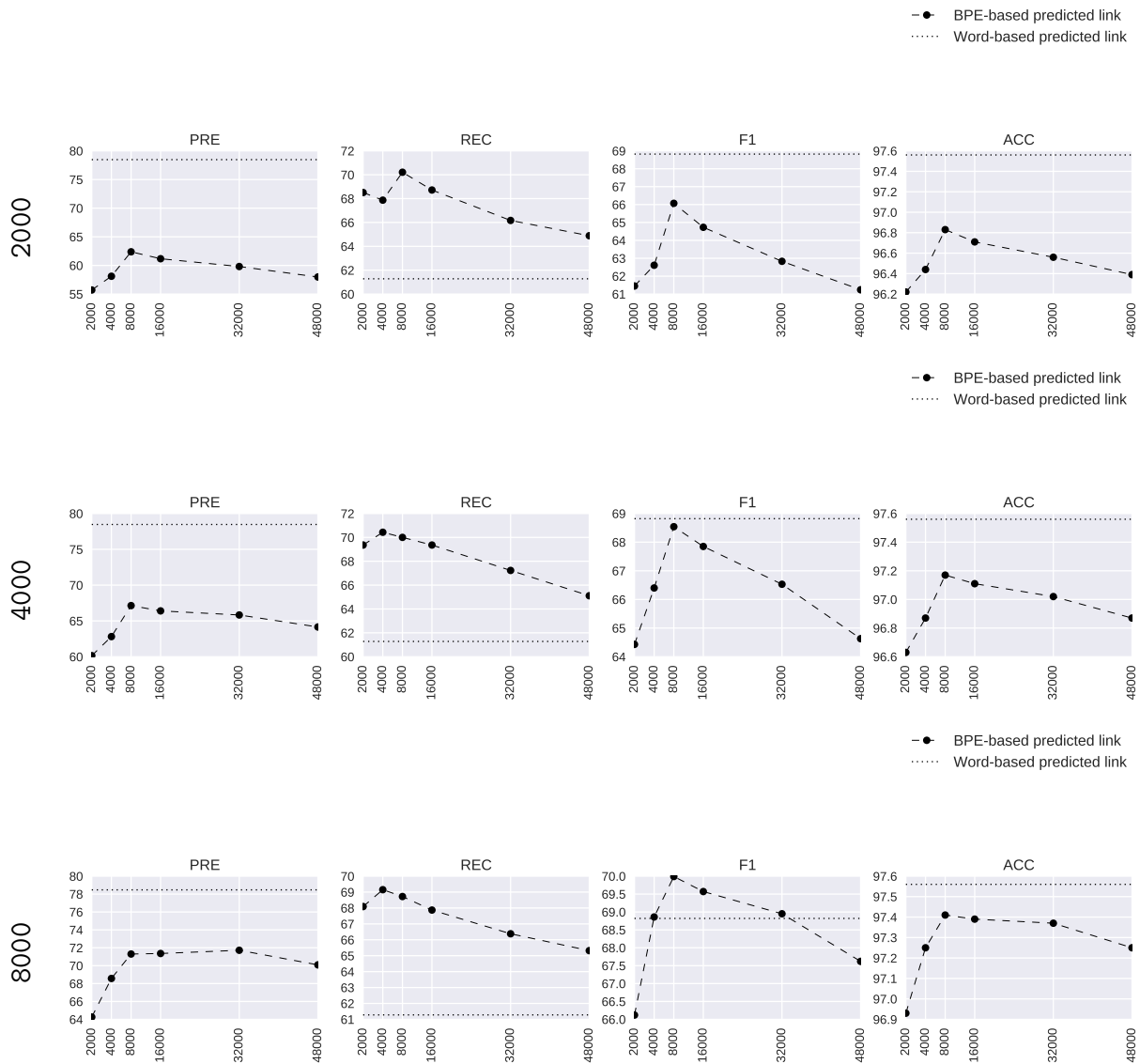


Figure E.72: BPE-based Fastalign for the direction Czech-English: We observe the scores for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.



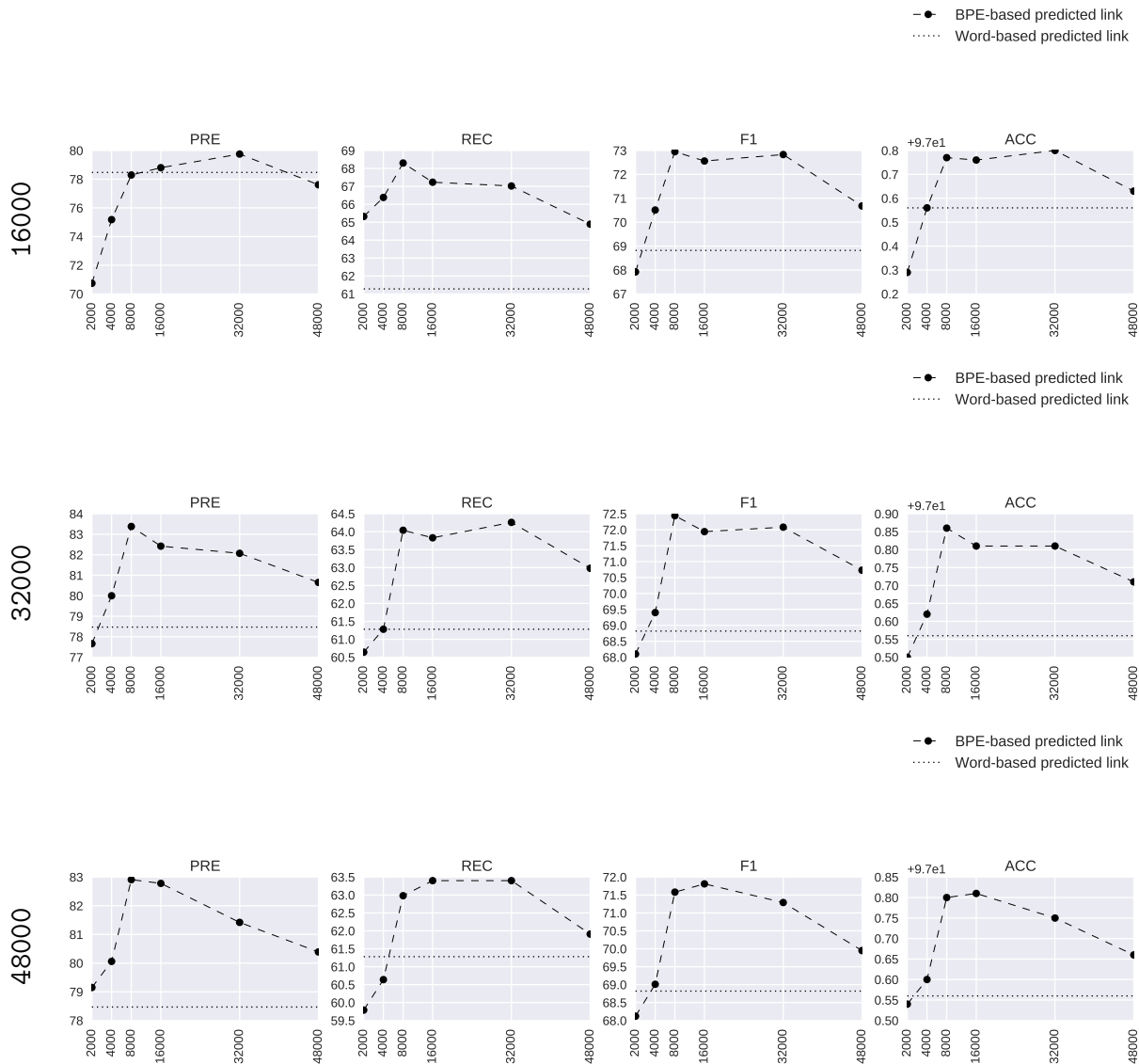
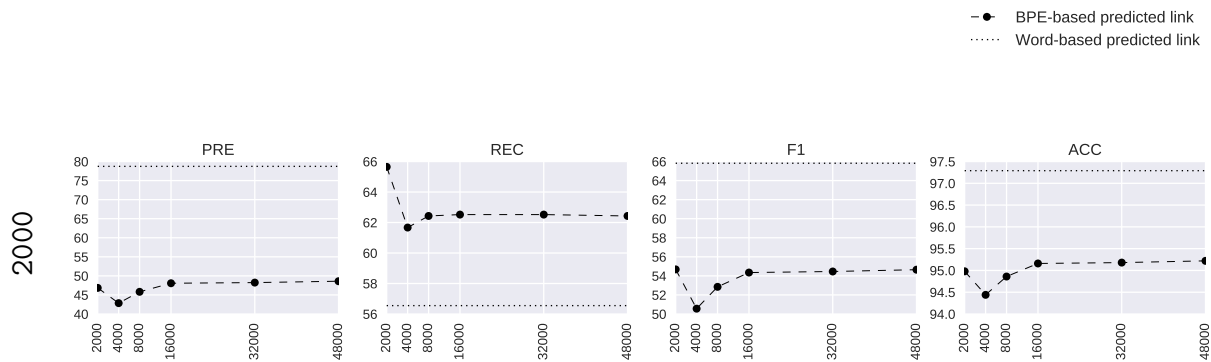


Figure E.73: BPE-based Fastalign for the direction English-Romanian: We observe the scores for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.



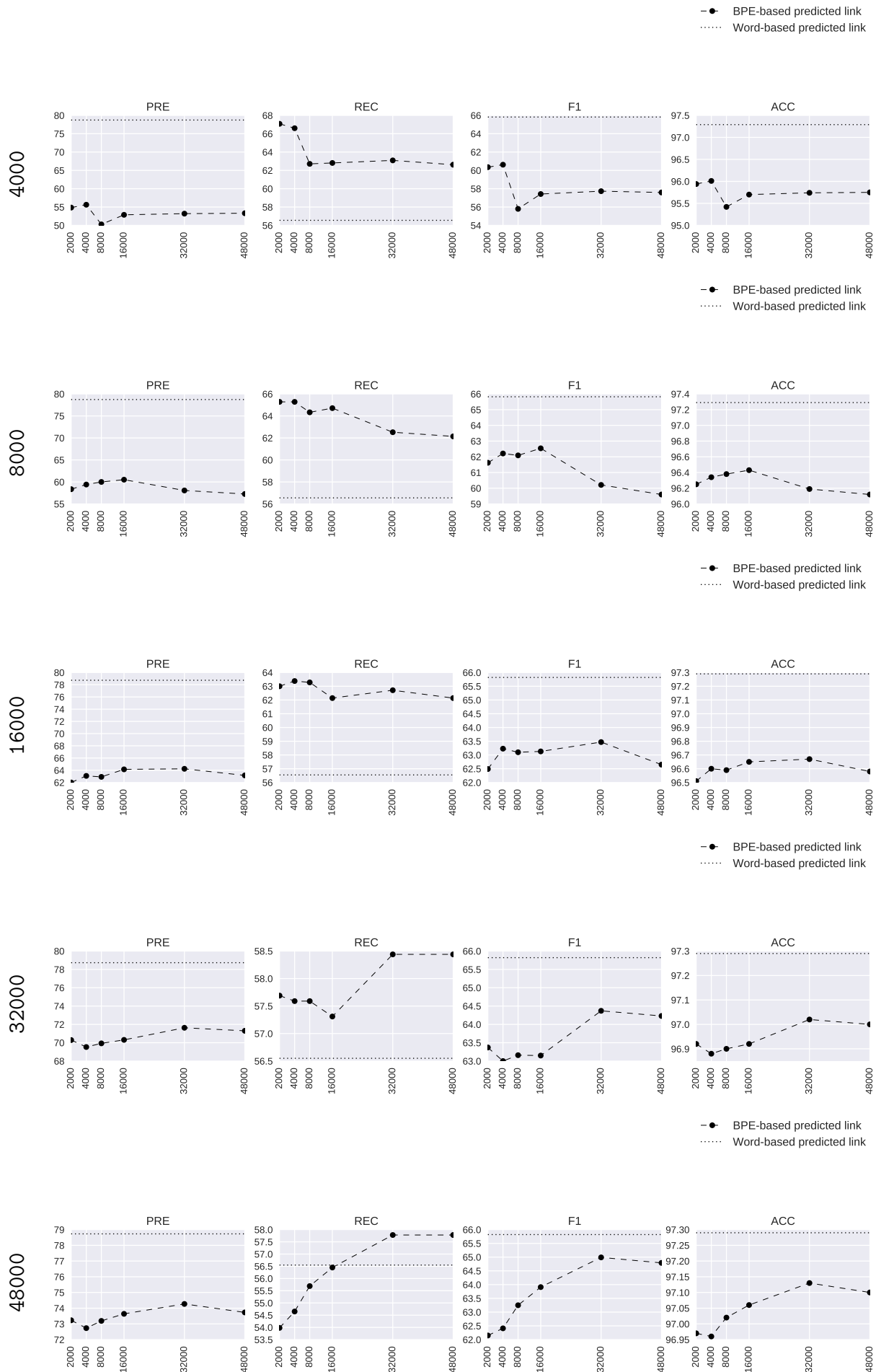
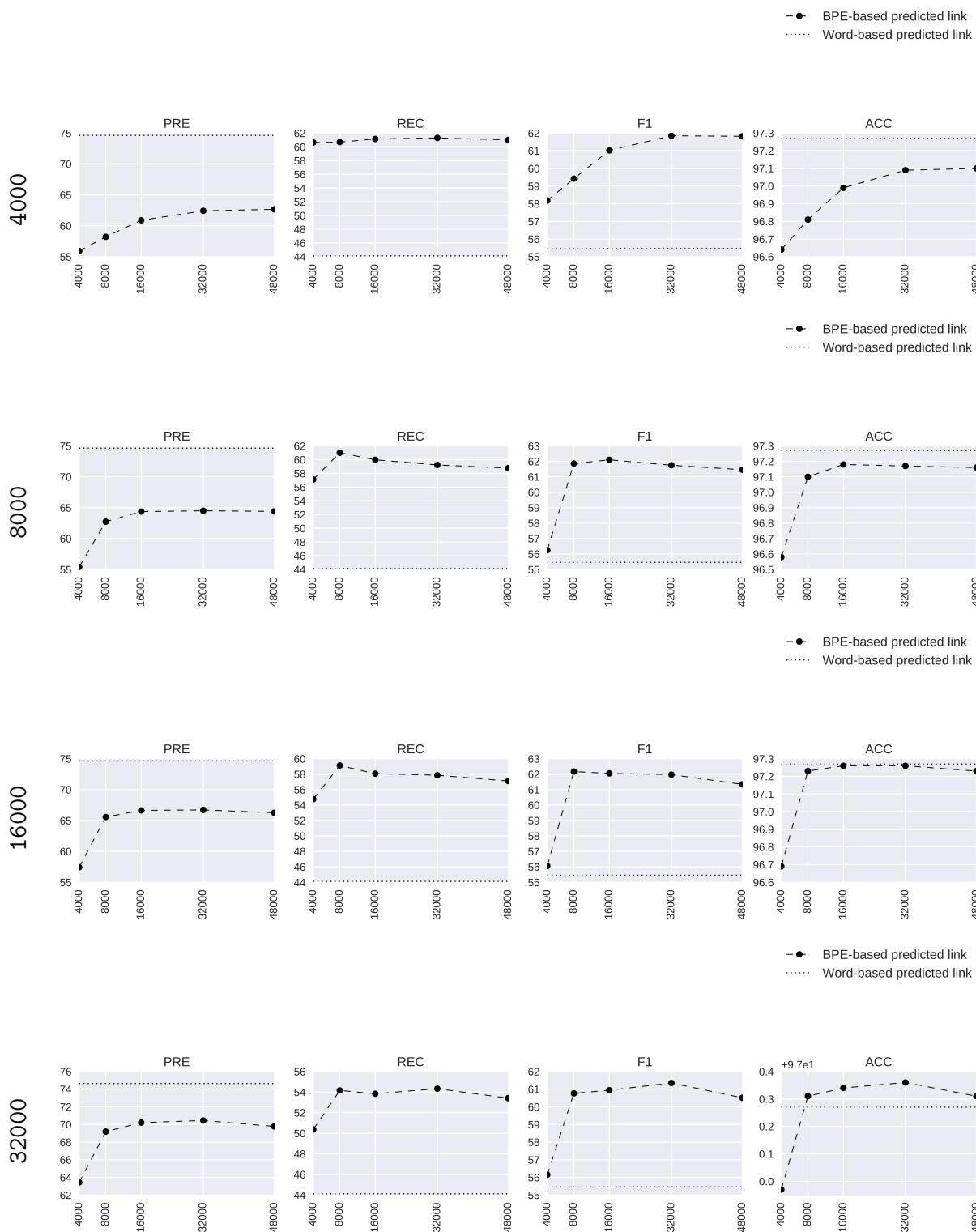


Figure E.74: BPE-based Fastalign for the direction Romanian-English: We observe the scores

for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.



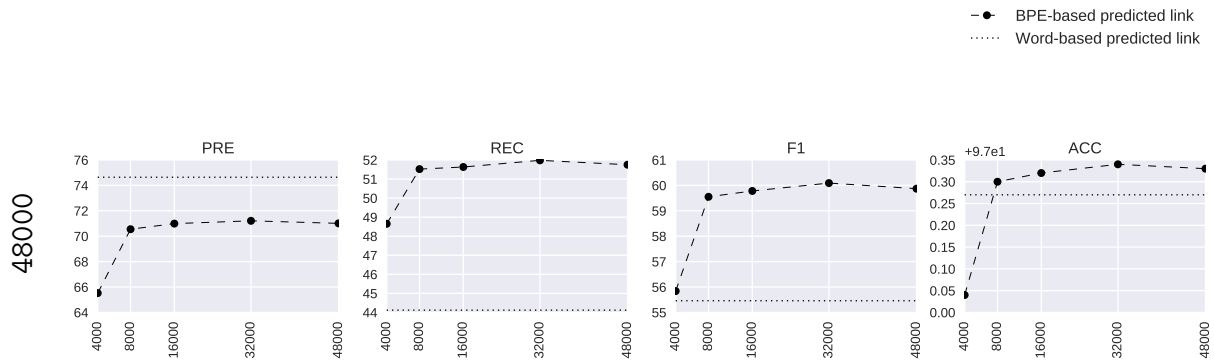
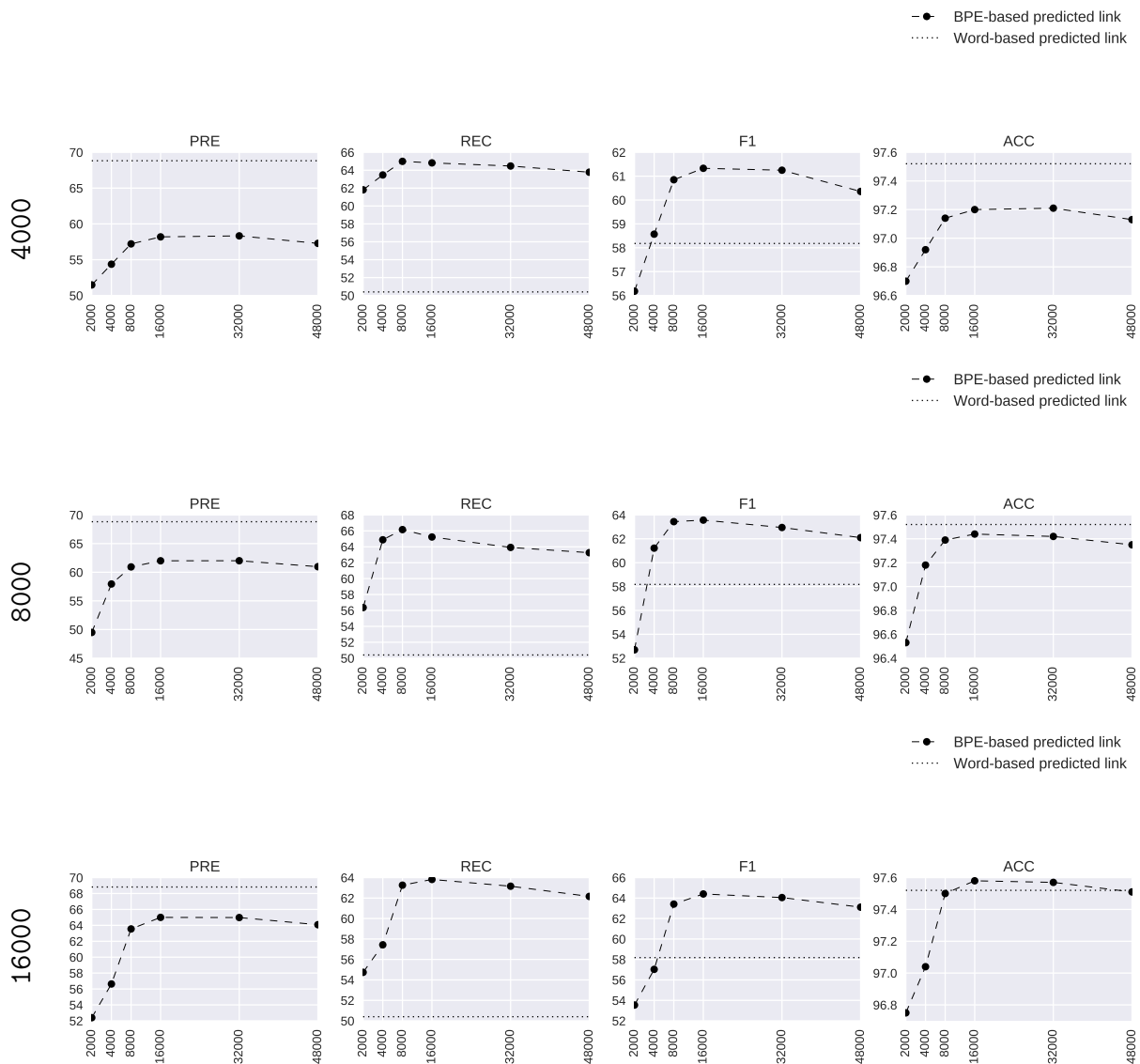


Figure E.75: BPE-based Fastalign for the direction English-Japanese: We observe the scores for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.



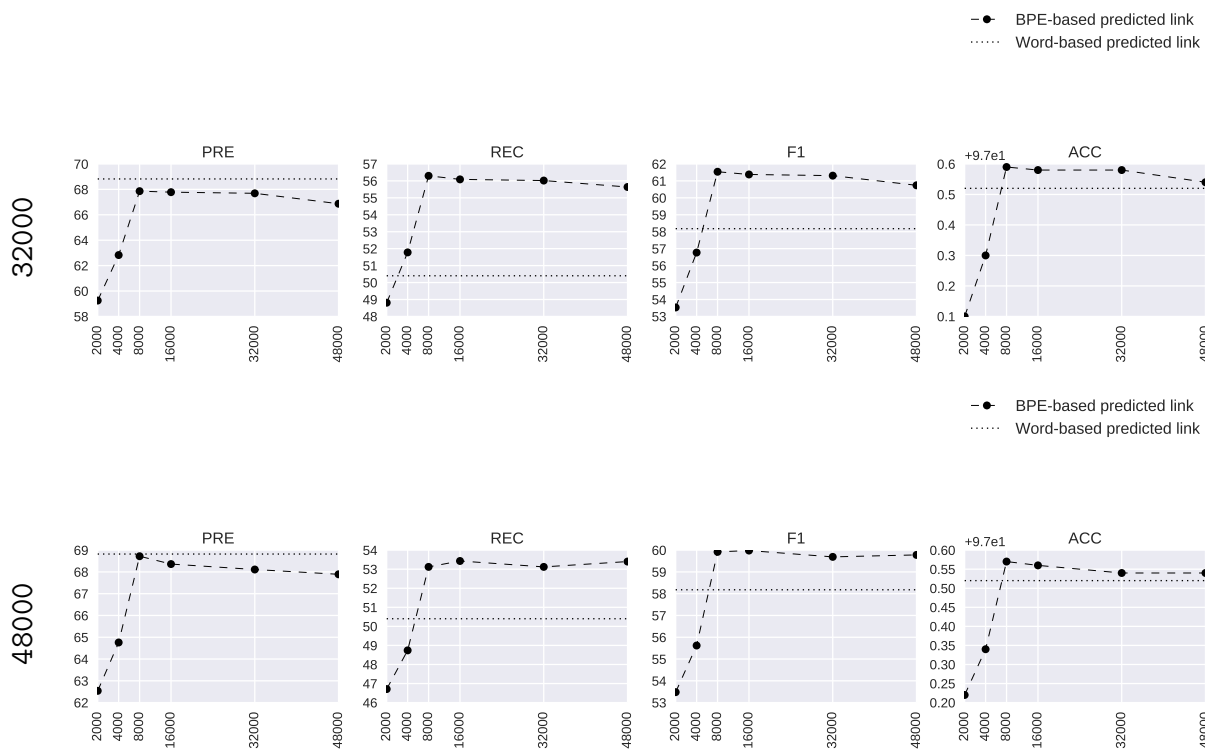
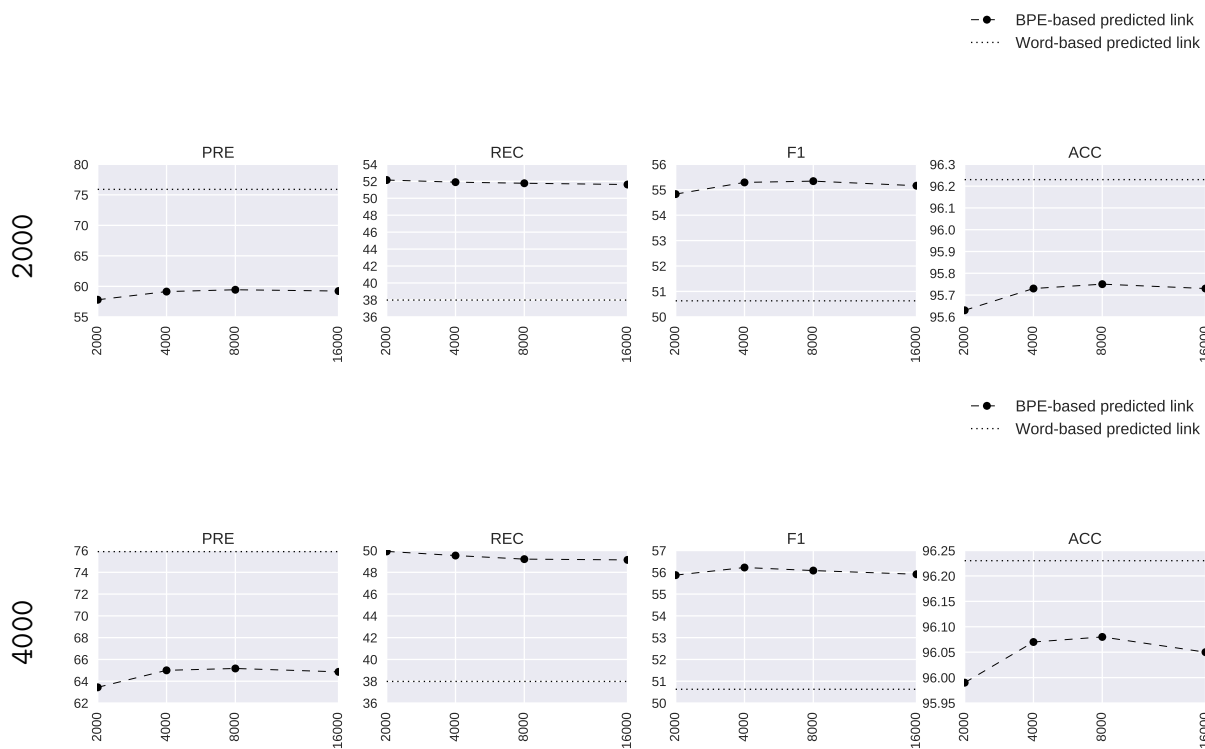


Figure E.76: BPE-based Fastalign for the direction Japanese-English: We observe the scores for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.



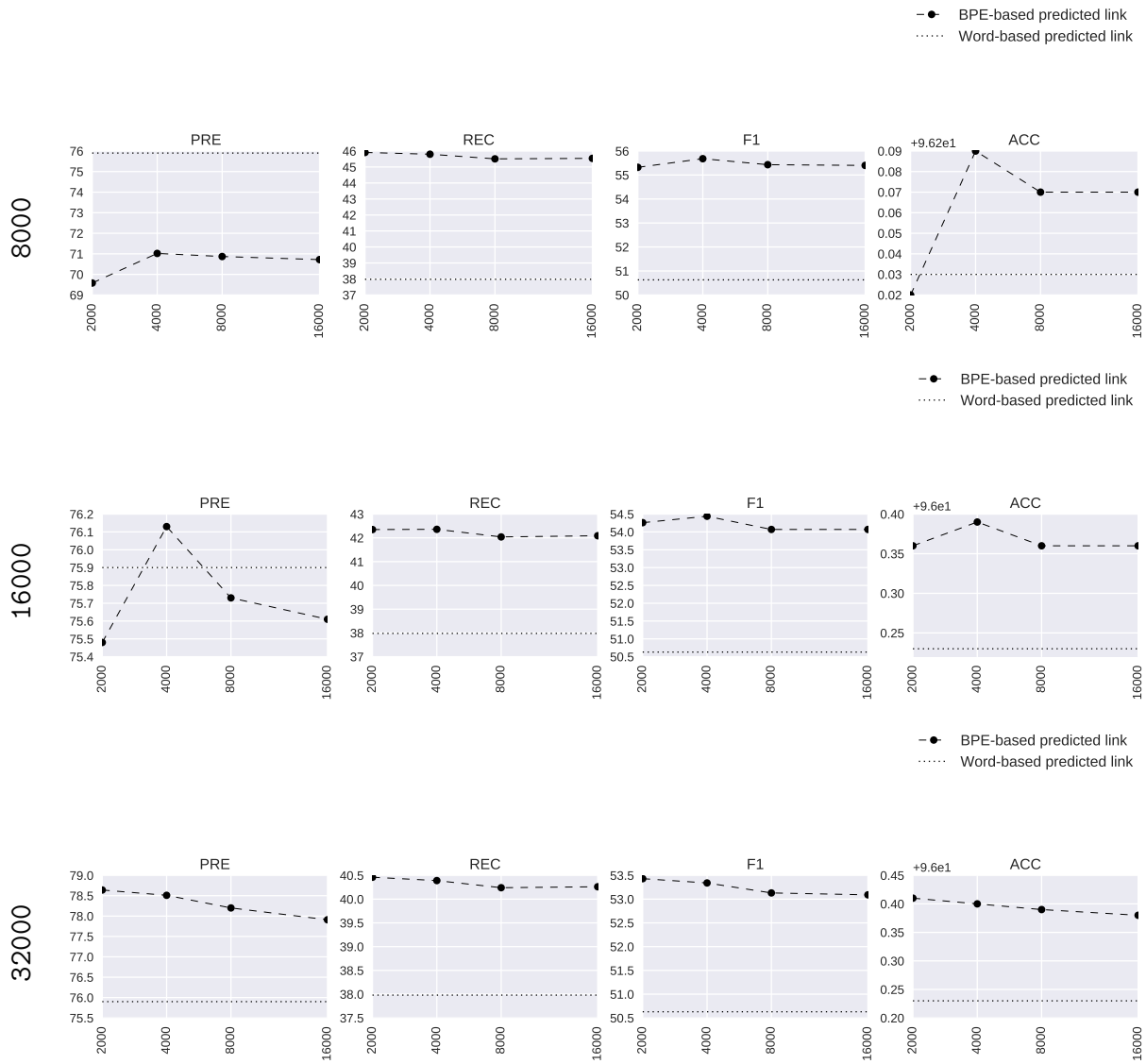
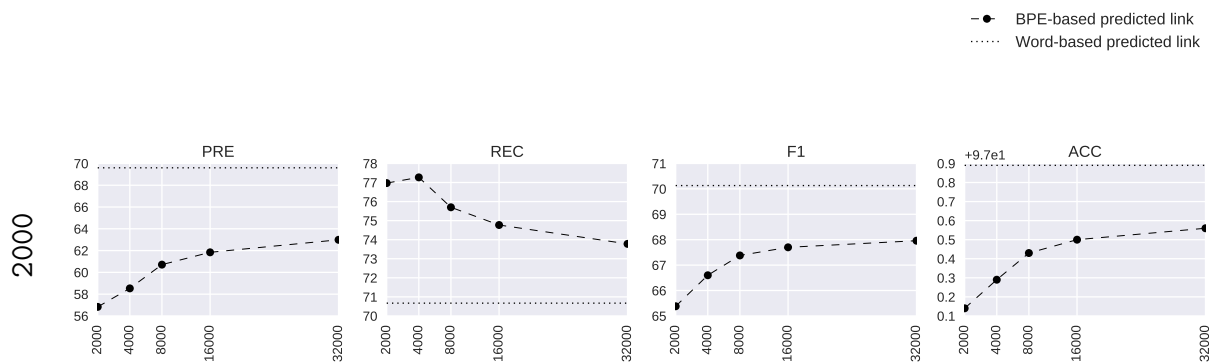


Figure E.77: BPE-based Fastalign for the direction English-Vietnamese: We observe the scores for rare source words. For each source vocabulary size, we report accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.



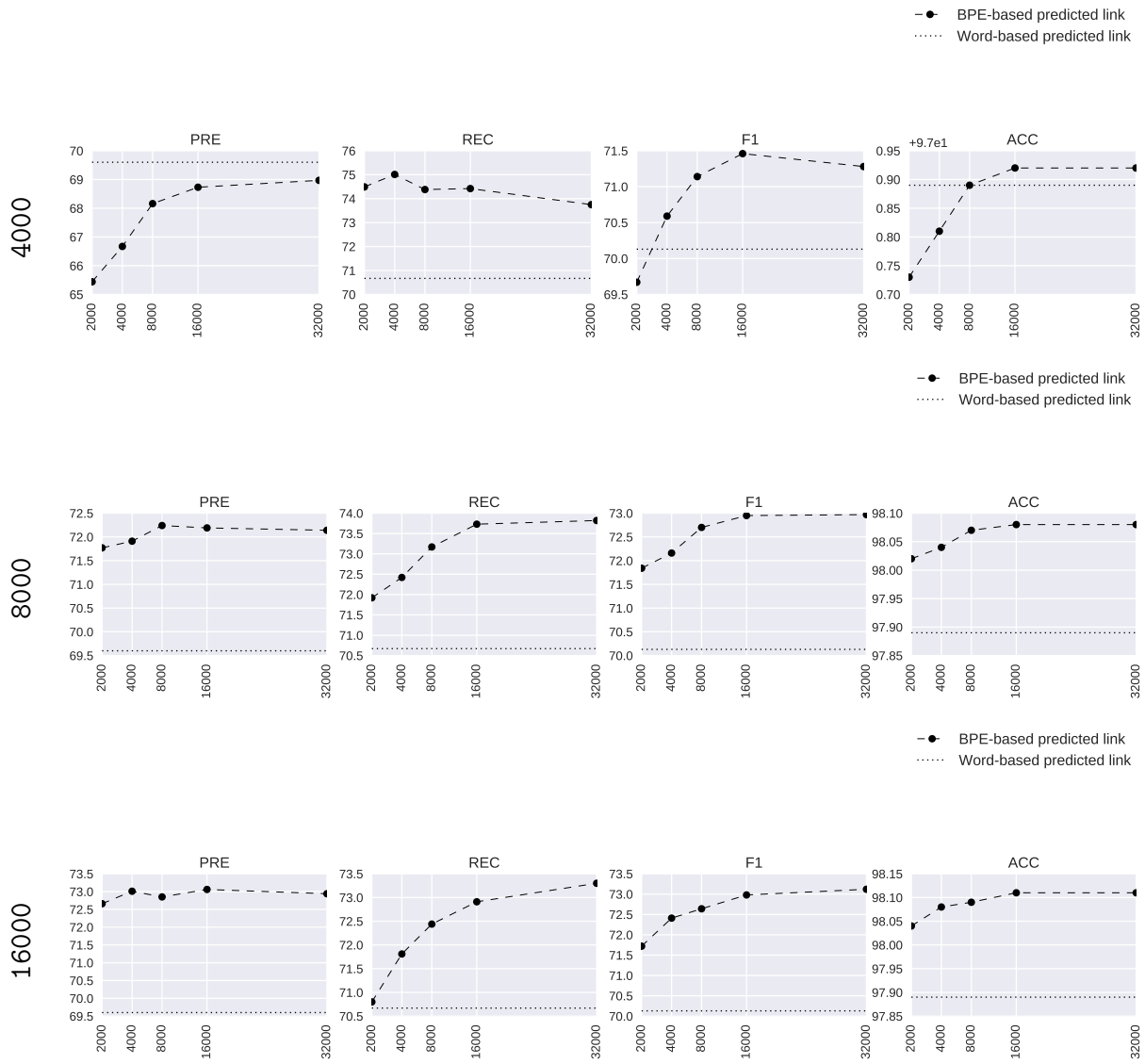
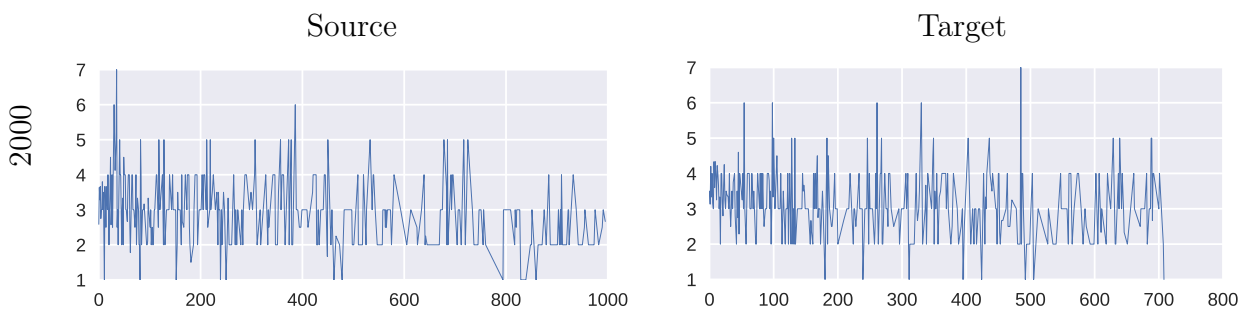


Figure E.78: BPE-based `Fastalign` for the direction Vietnamese-English: We observe the scores for rare source words. For each source vocabulary size, we report the accuracy (ACC), F-score (F1), precision (PRE) and recall (REC) as a function of the target vocabulary size.

E.6 Number of fragments as a function of word occurrence



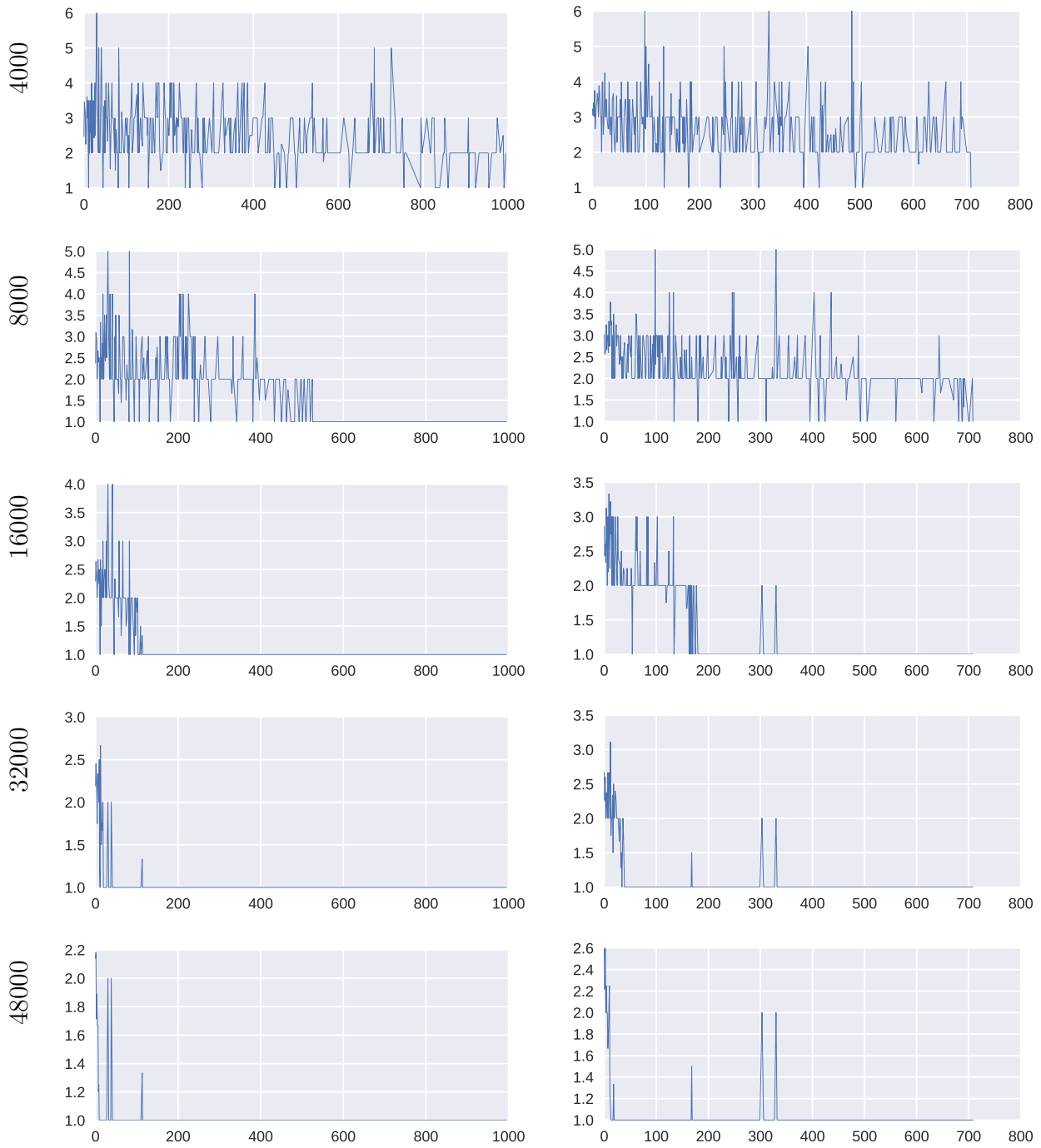
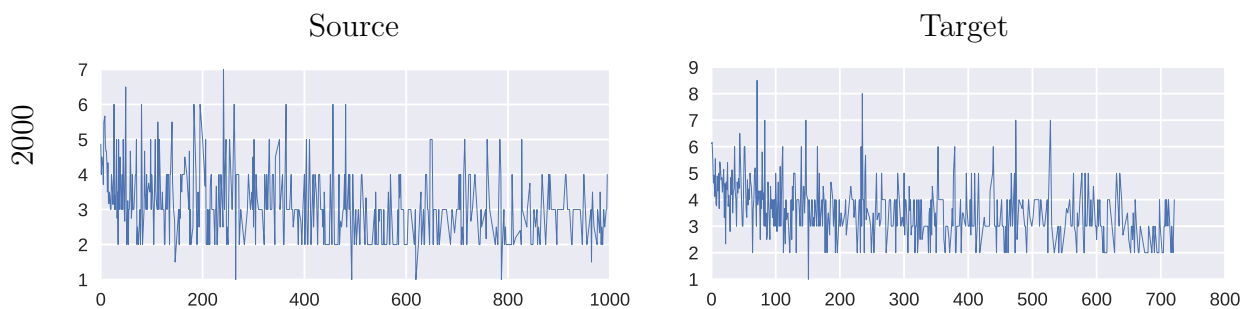


Figure E.79: The direction English-French: Average number of BPE-based fragments as a function of word occurrence.



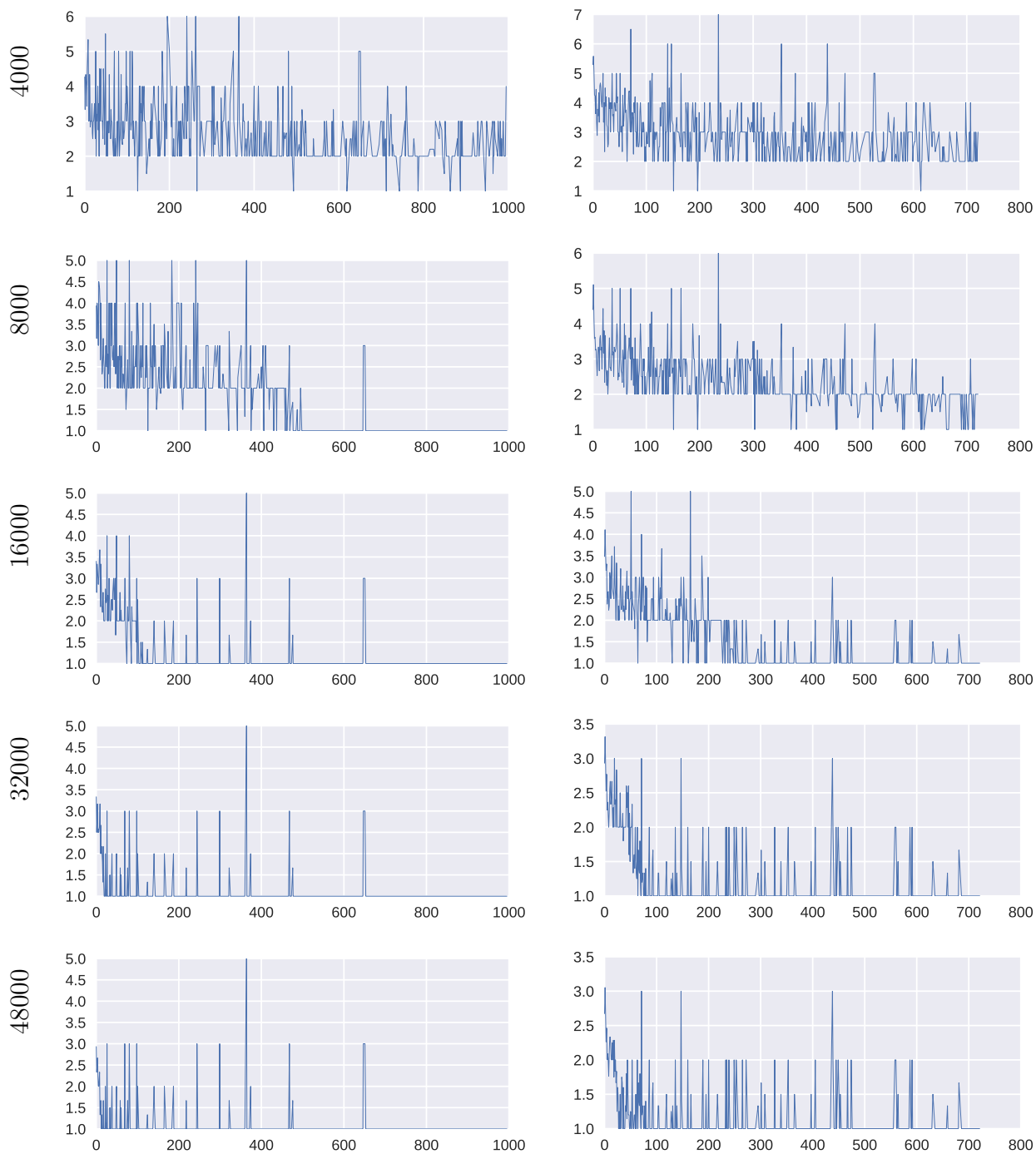
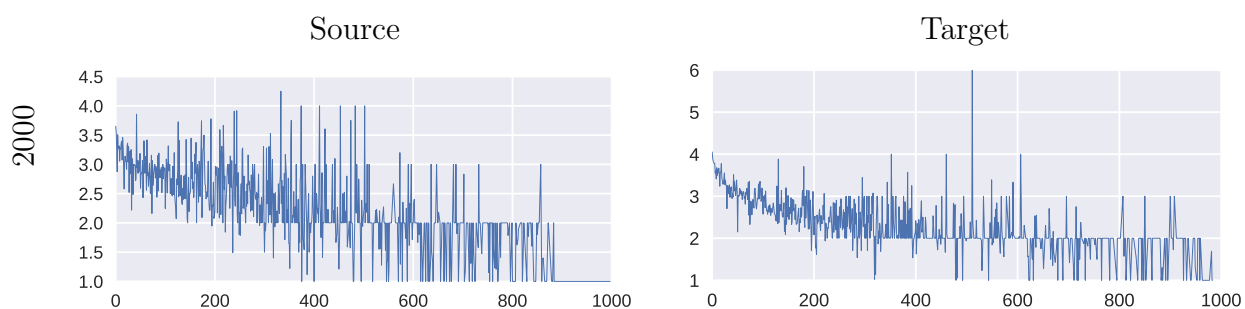


Figure E.80: The direction English-German: Average number of BPE-based fragments as a function of word occurrence.



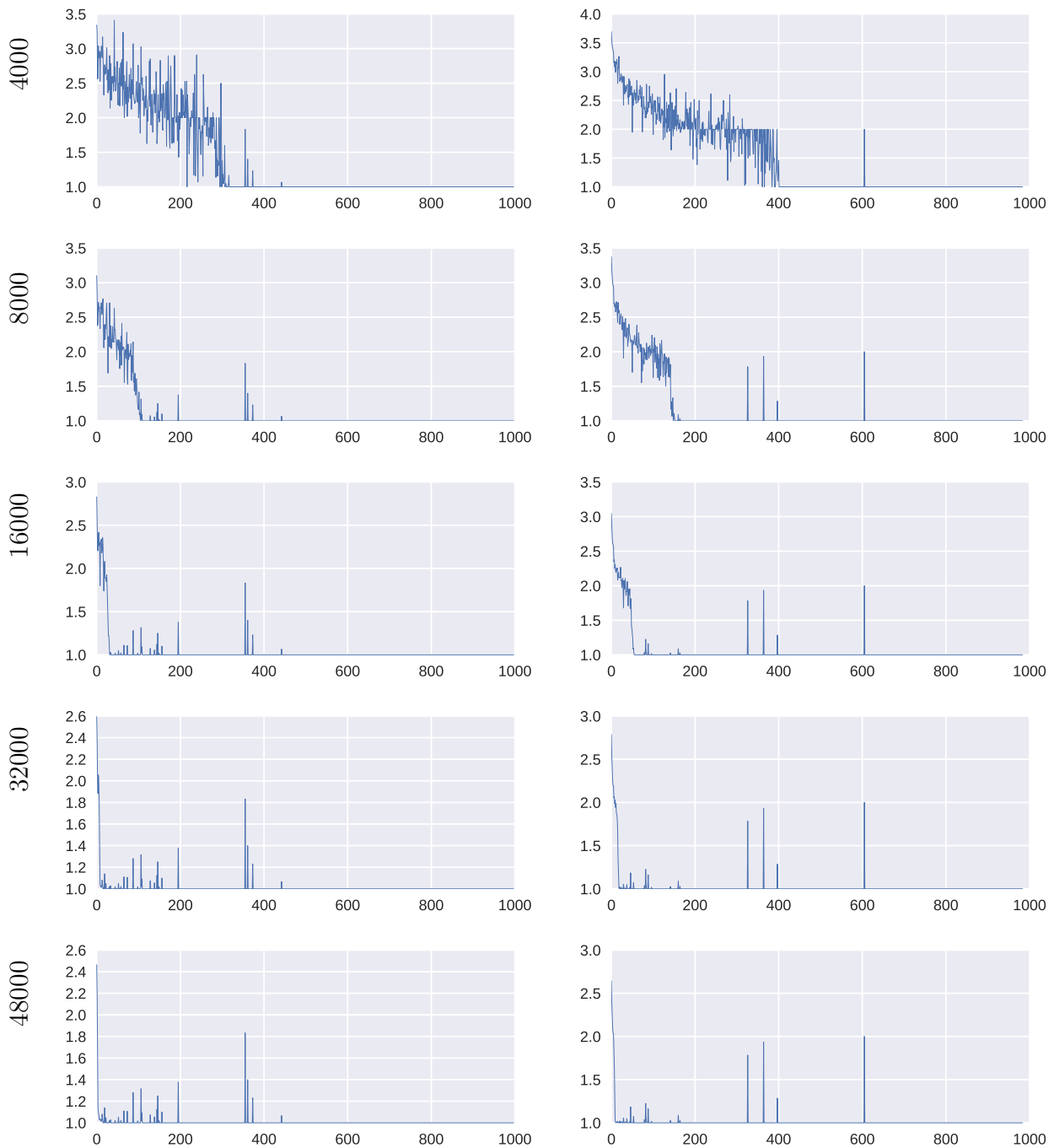
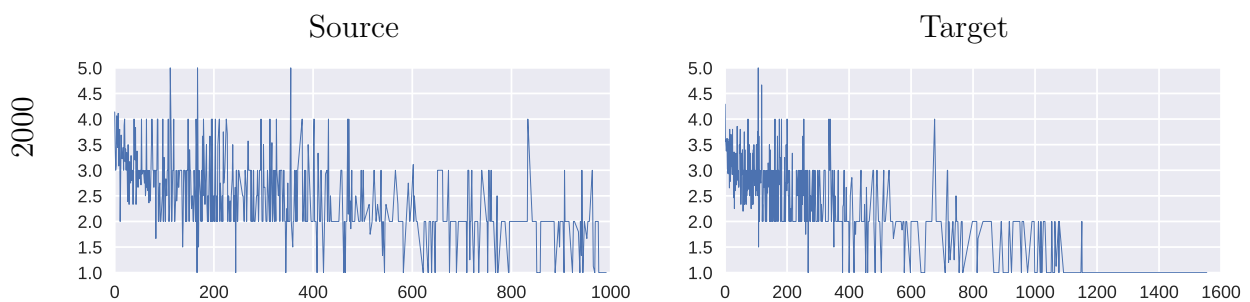


Figure E.81: The direction English-Czech: Average number of BPE-based fragments as a function of word occurrence.



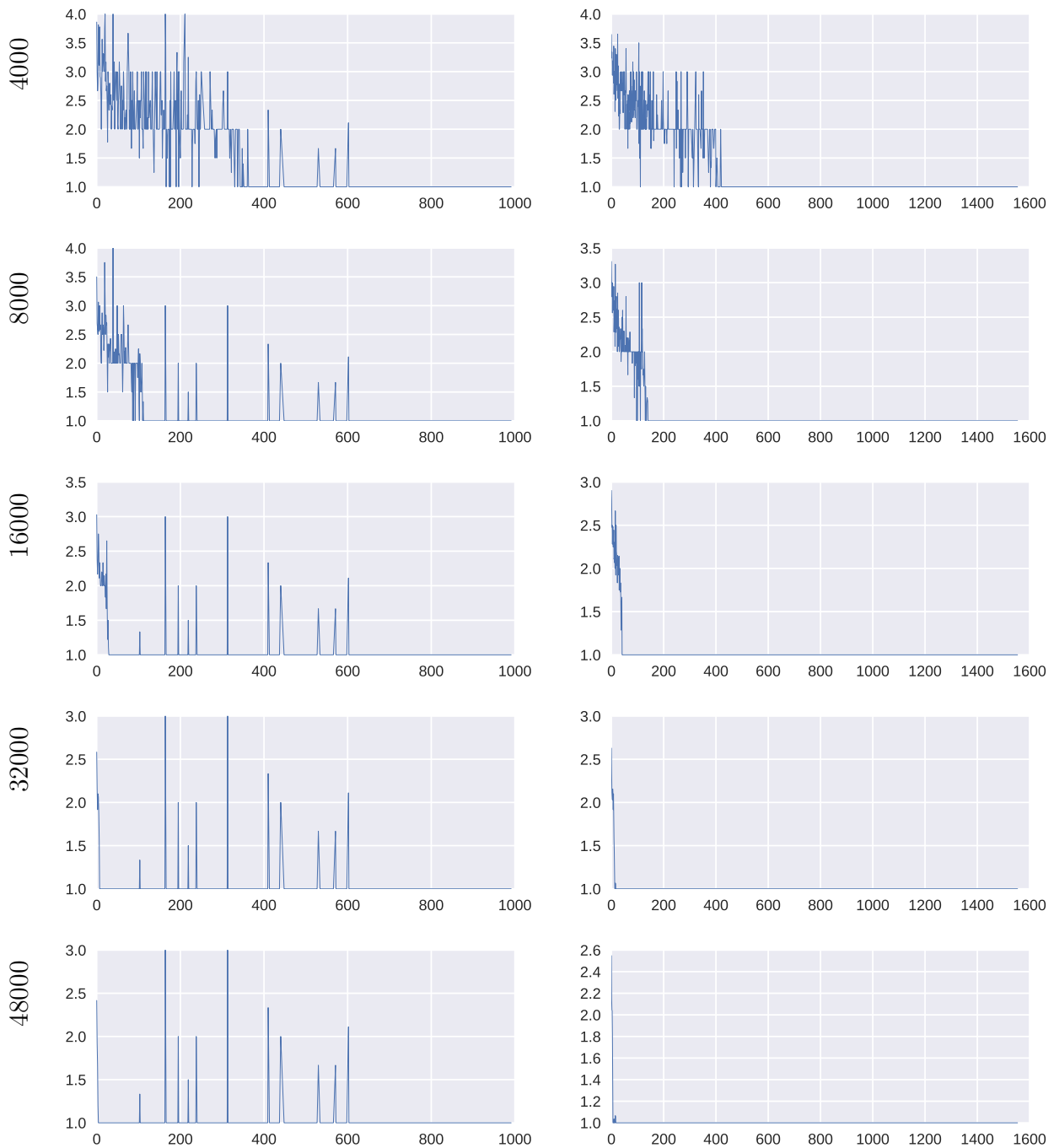
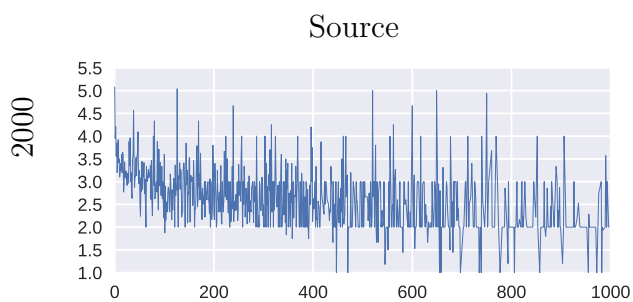


Figure E.82: The direction English-Romanian: Average number of BPE-based fragments as a function of word occurrence.



Target

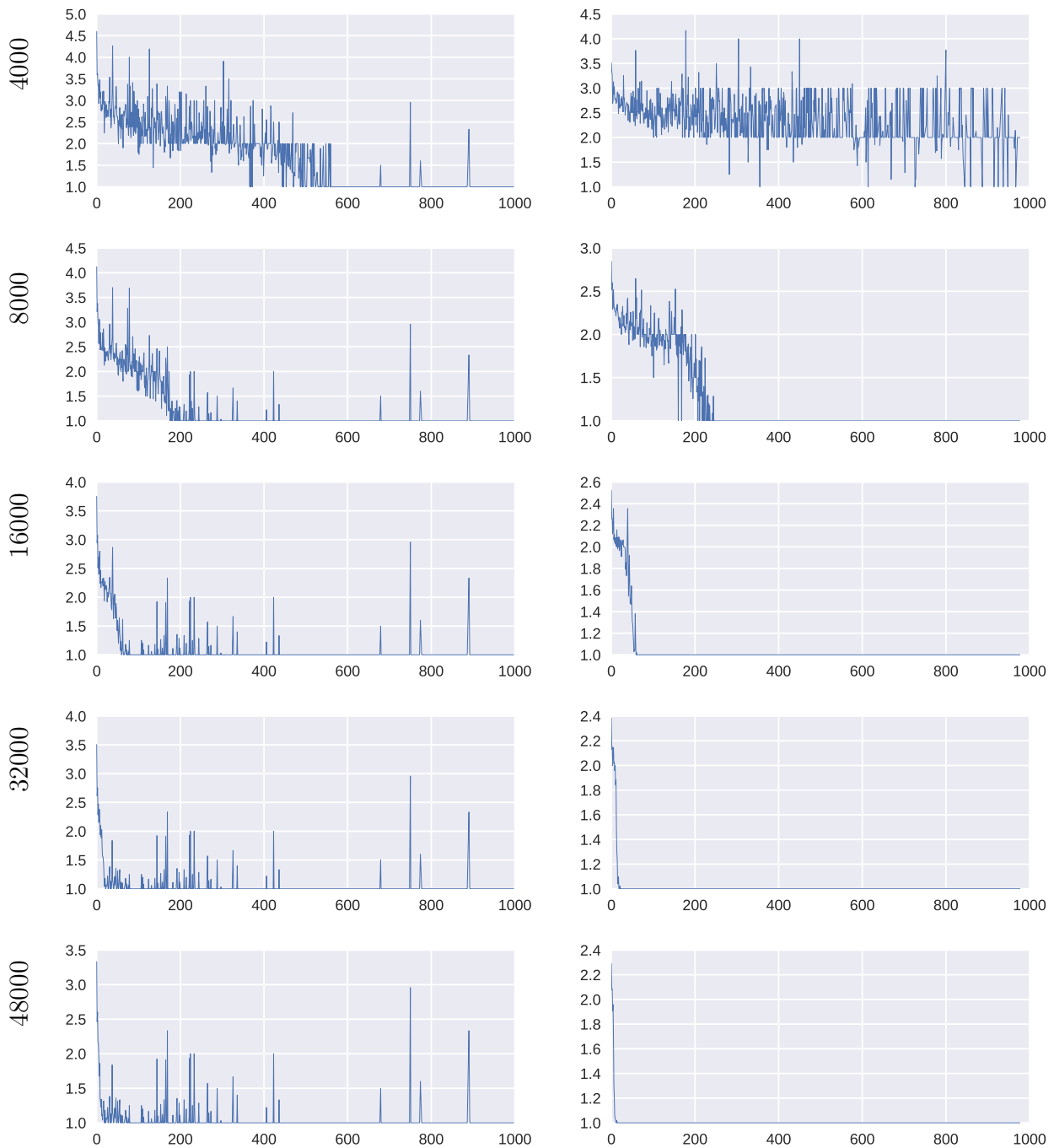
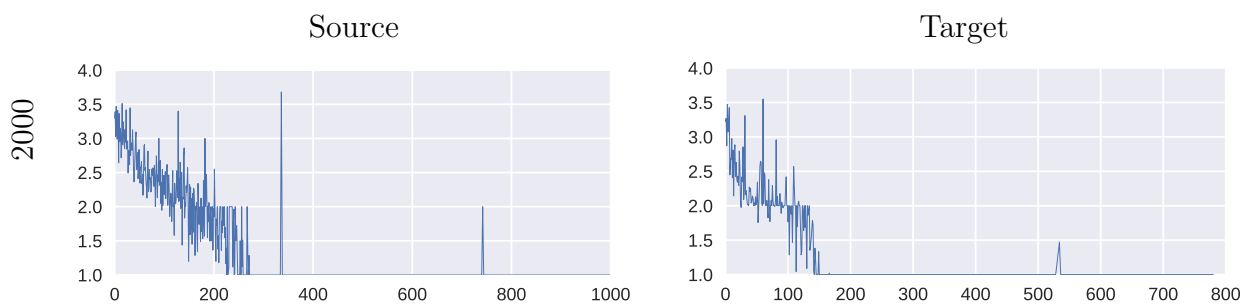


Figure E.83: The direction English-Japanese: Average number of BPE-based fragments as a function of word occurrence.



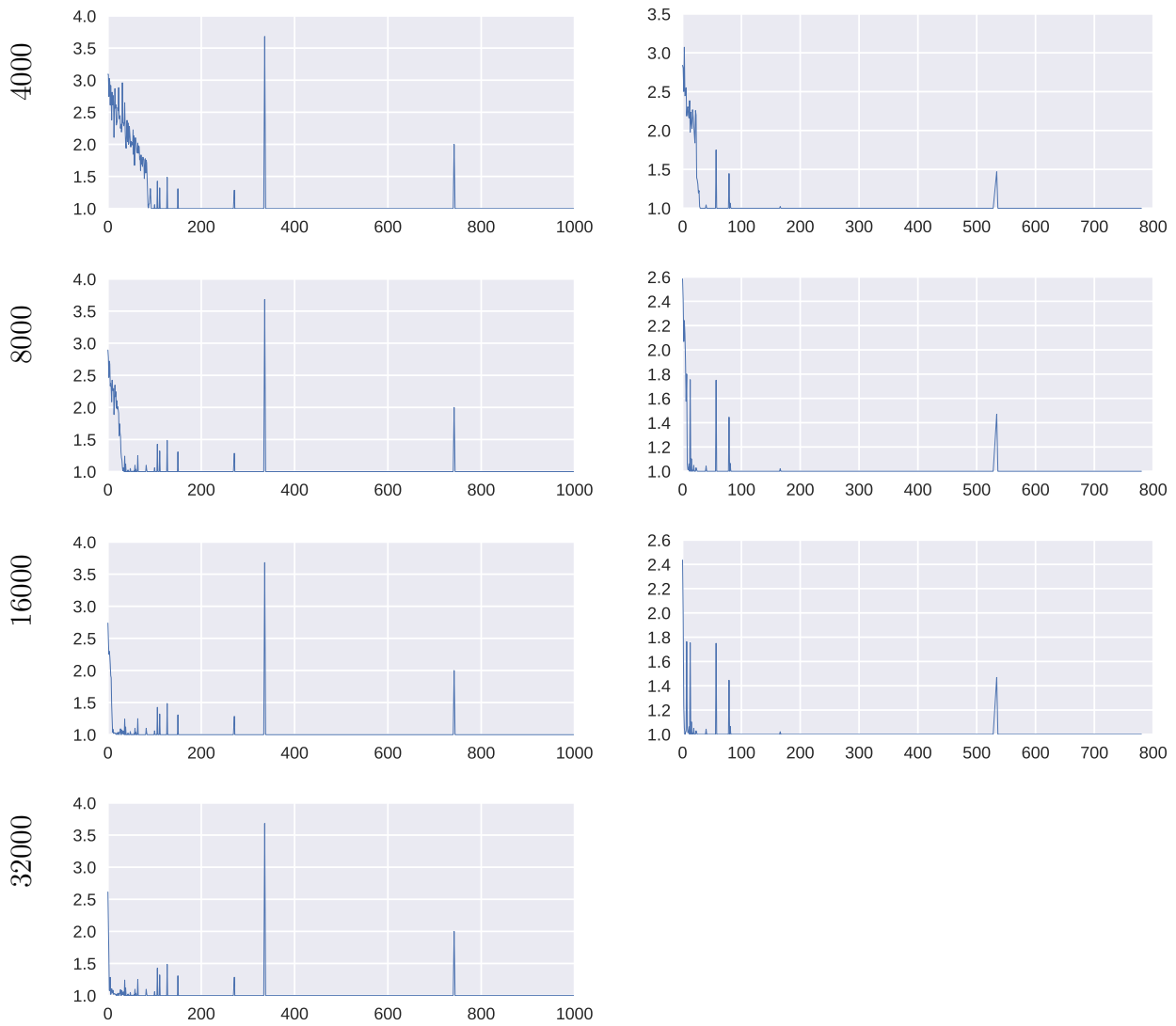
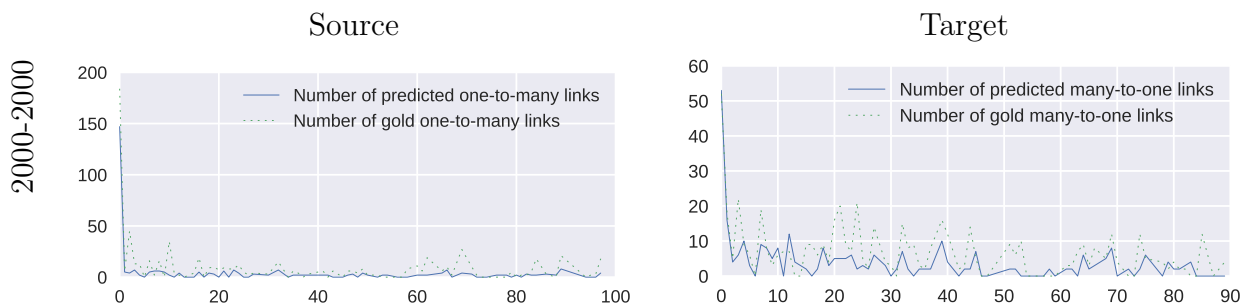
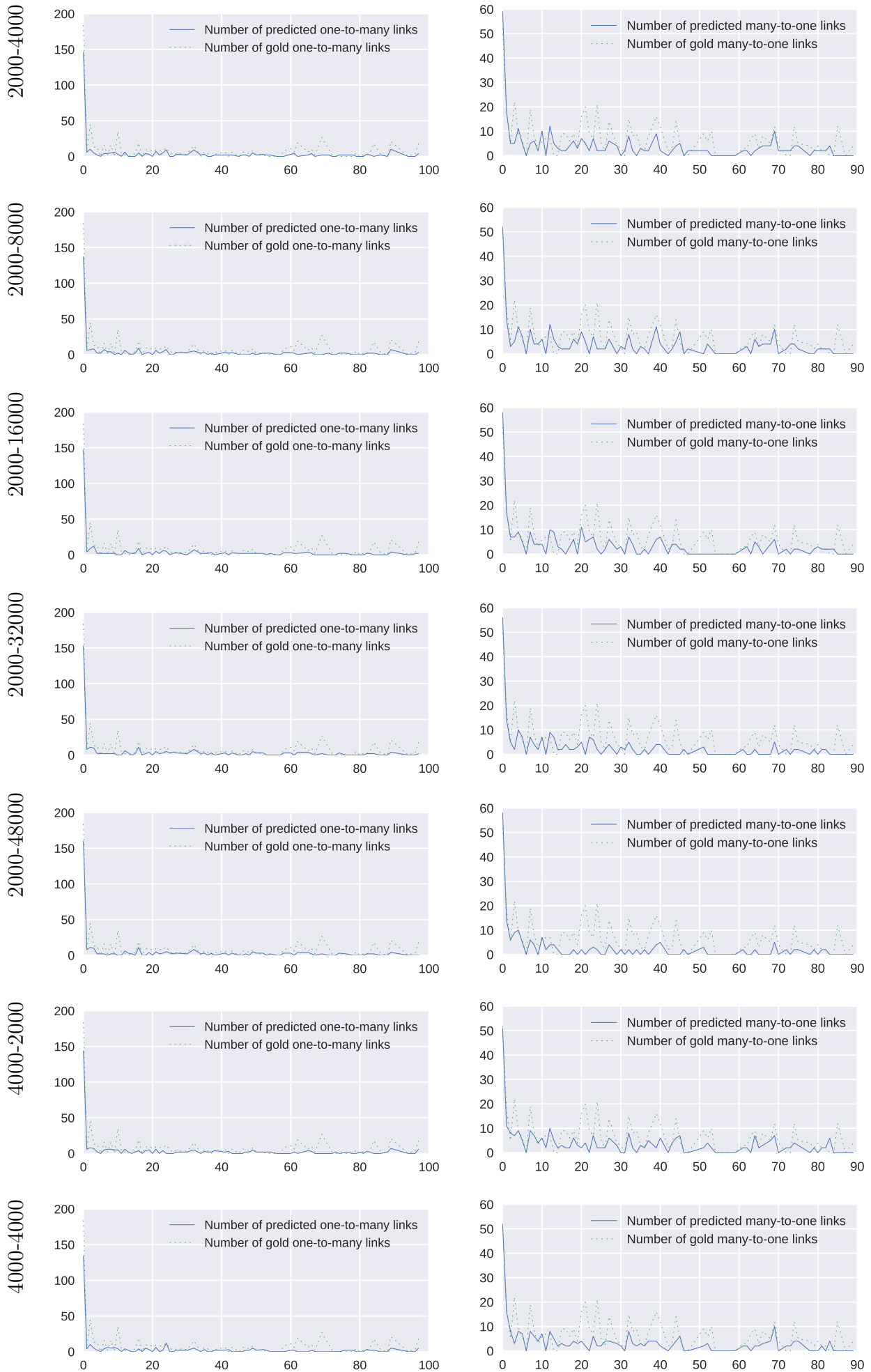


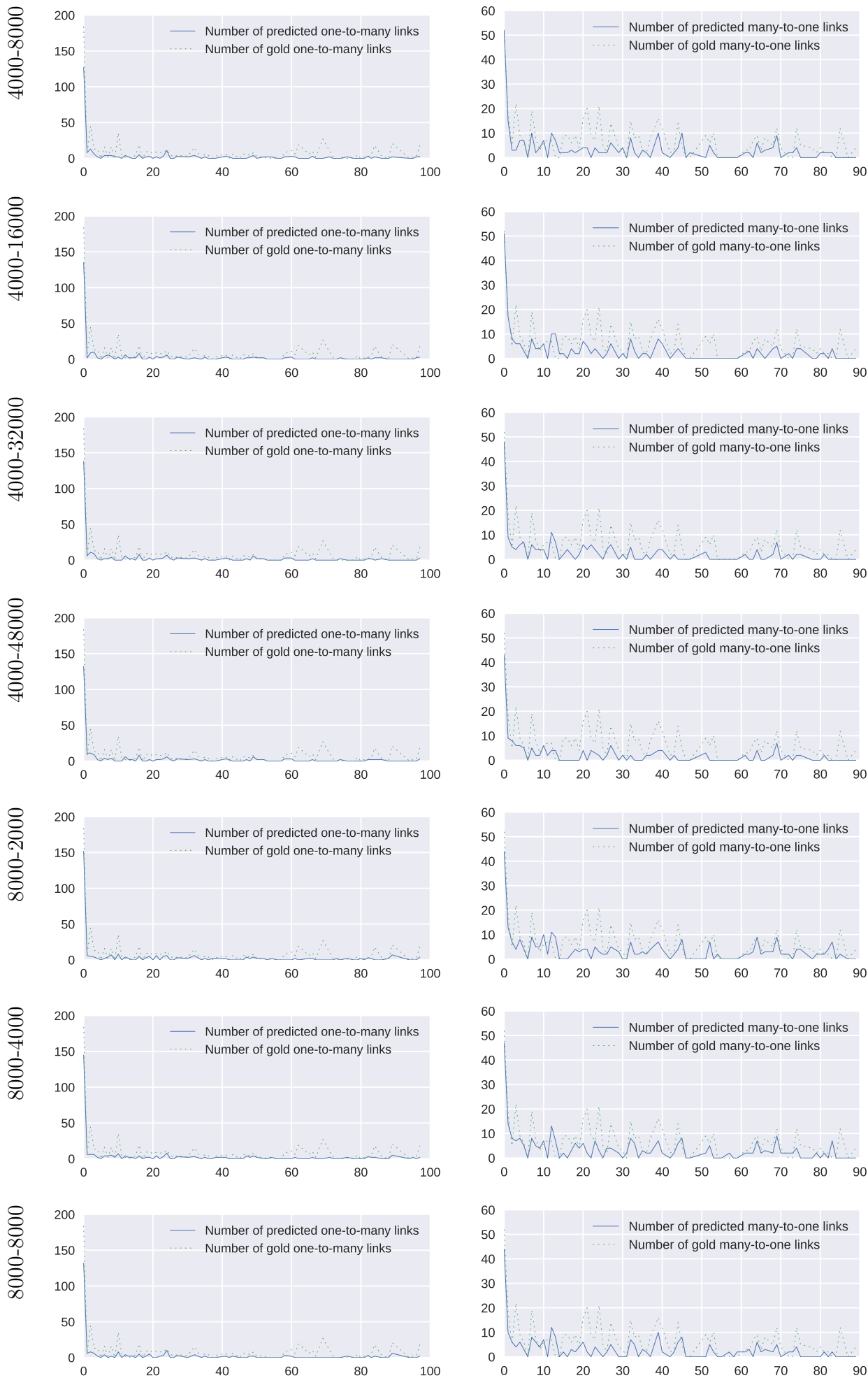
Figure E.84: The direction English-Vietnamese: Average number of BPE-based fragments as a function of word occurrence.

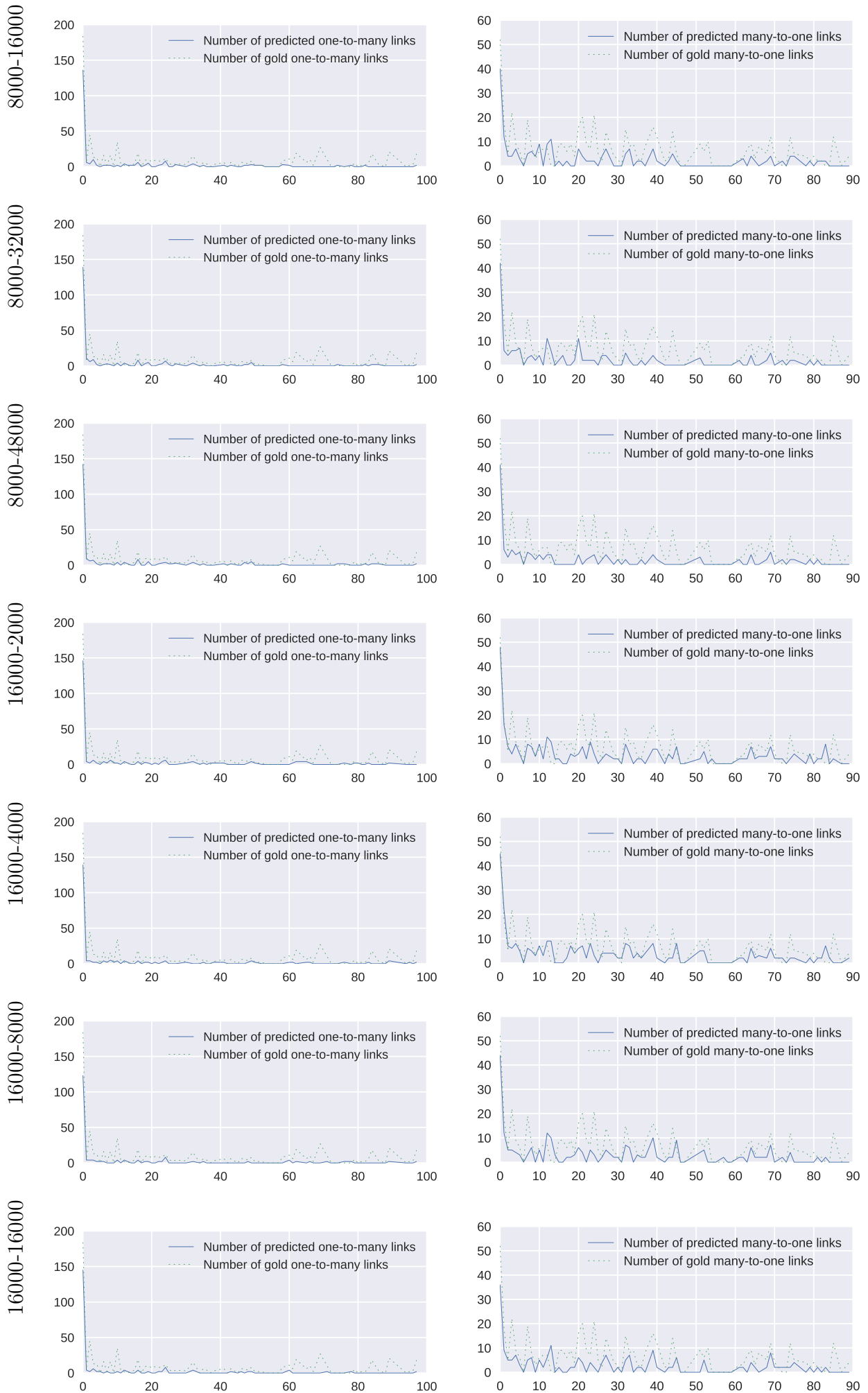
E.7 Number of one-to-many/many-to-one as a function of word occurrence



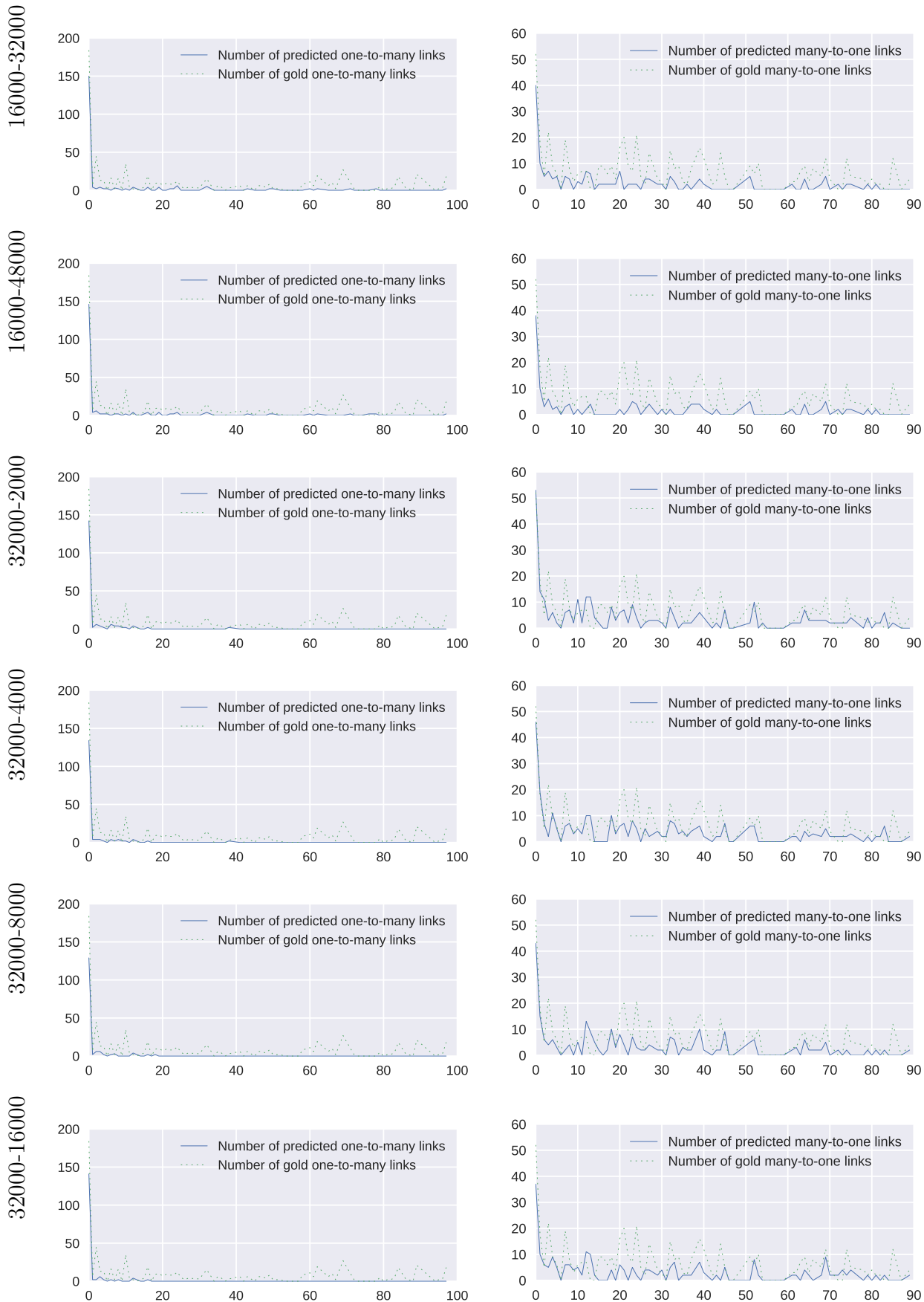


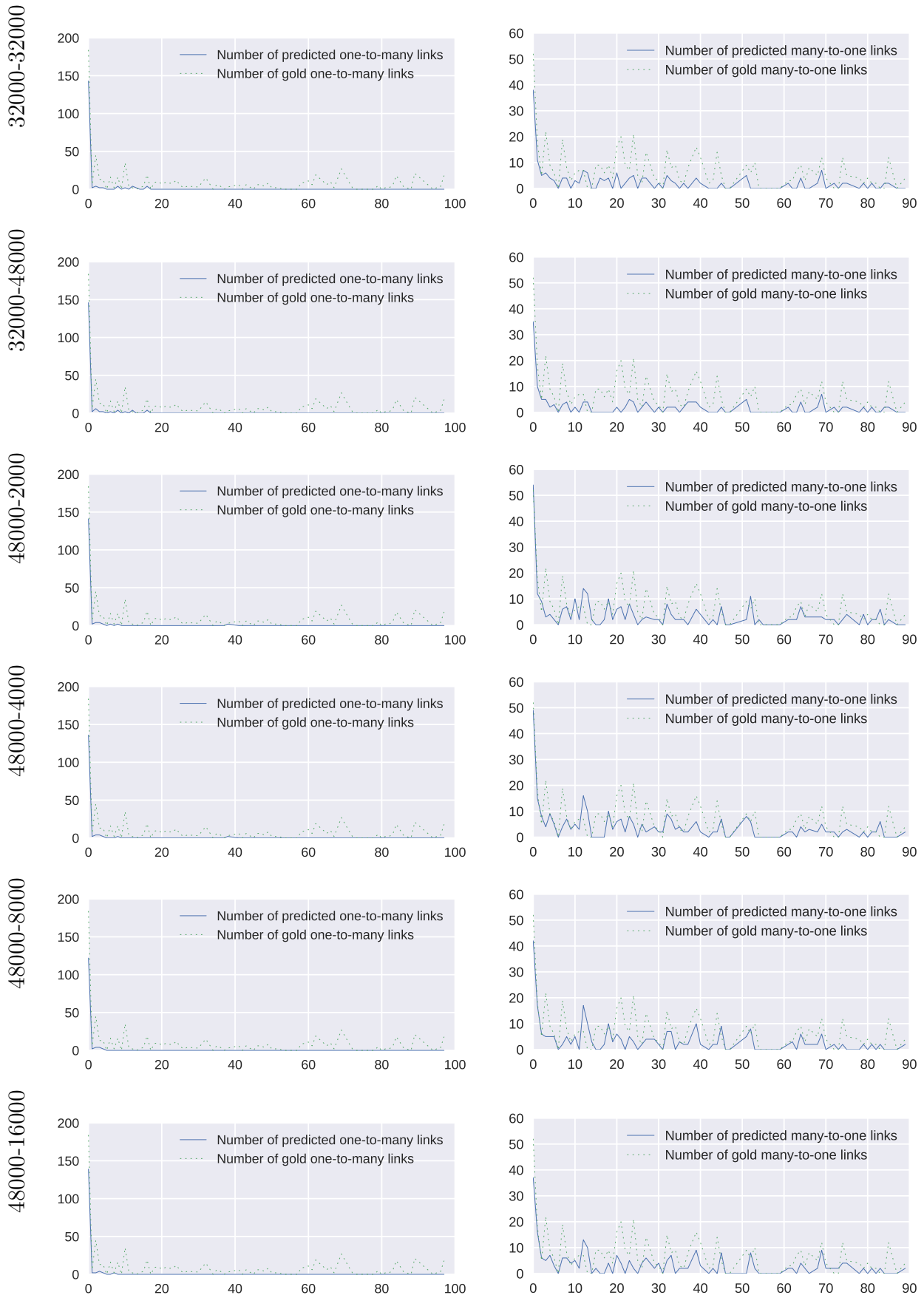
E.7. NUMBER OF ONE-TO-MANY/MANY-TO-ONE AS A FUNCTION OF WORD OCCURRENCE48





E.7. NUMBER OF ONE-TO-MANY/MANY-TO-ONE AS A FUNCTION OF WORD OCCURRENCE48





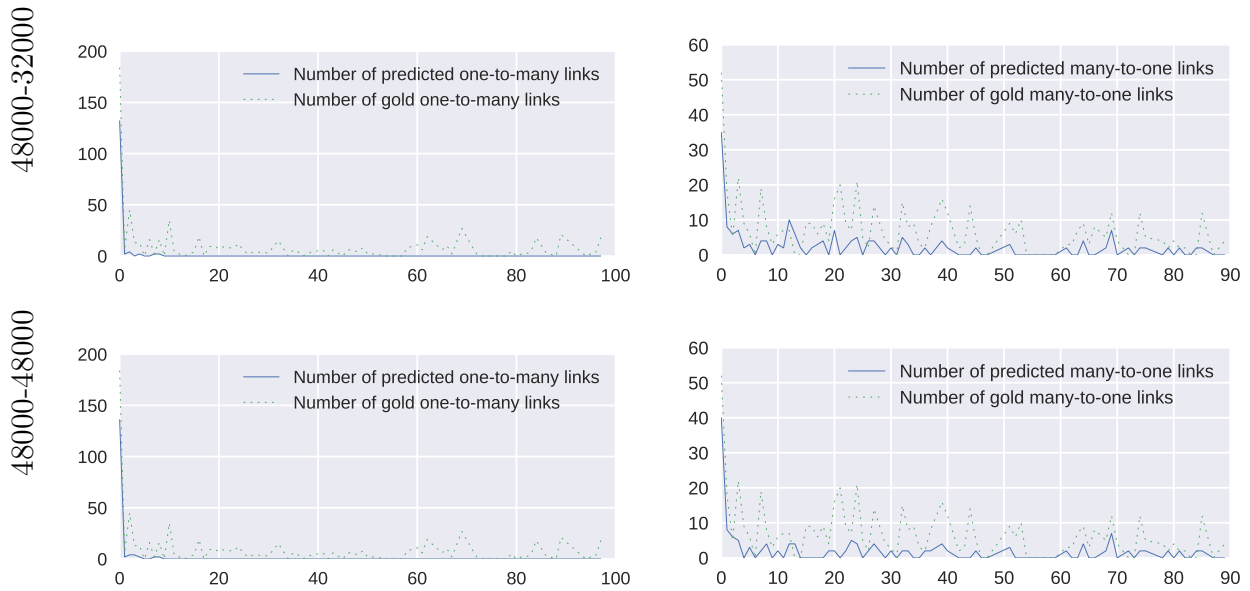
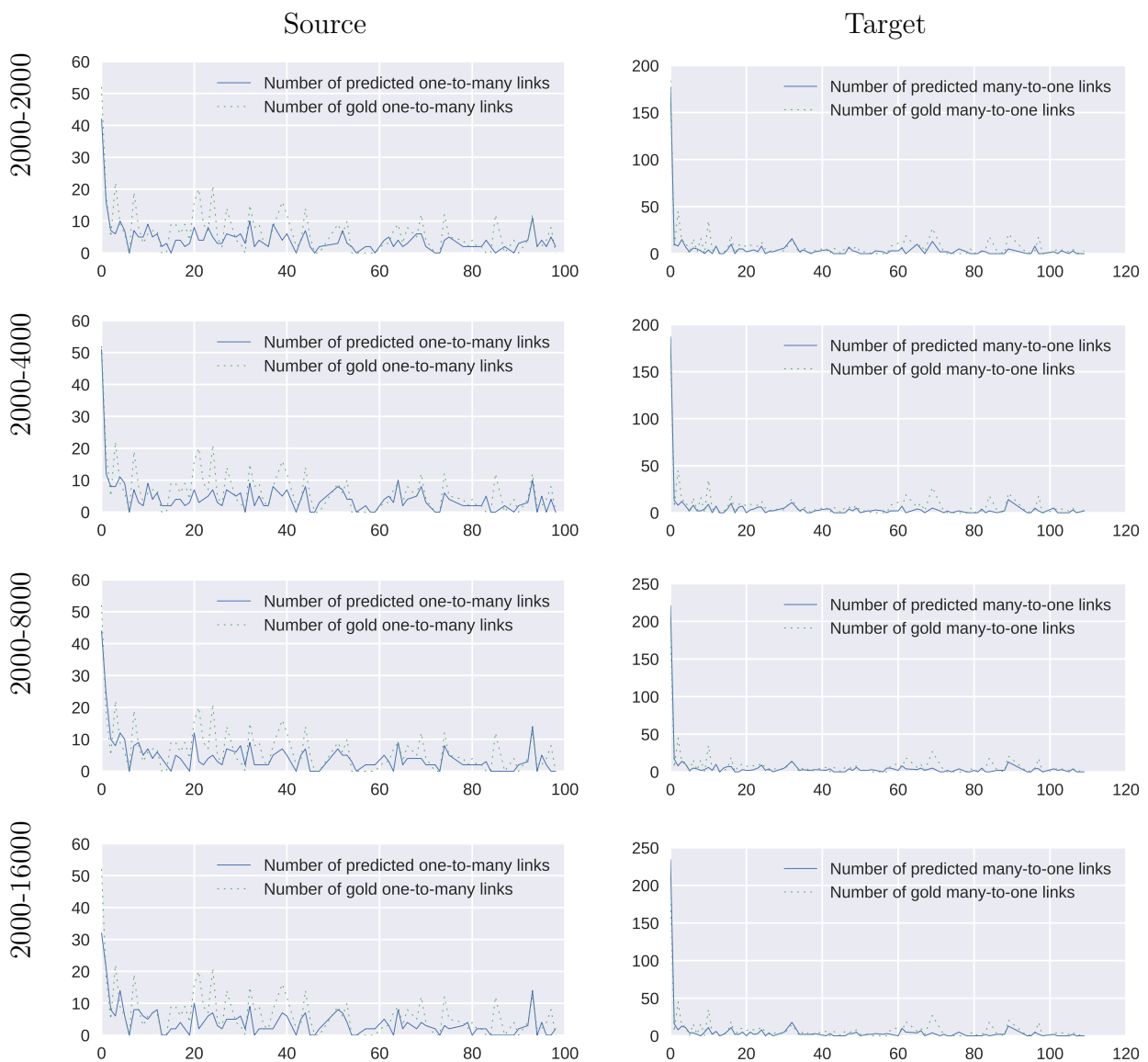
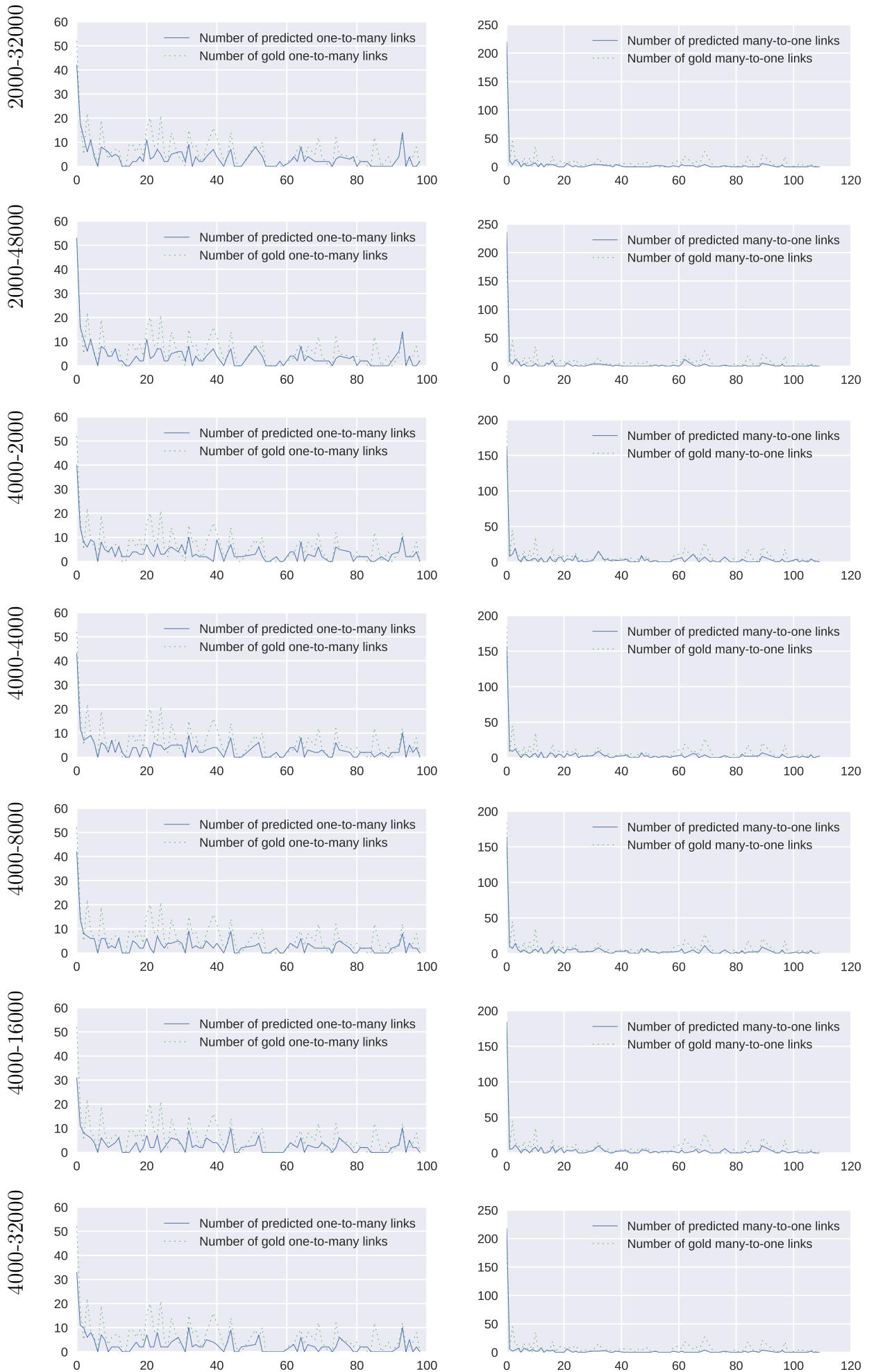
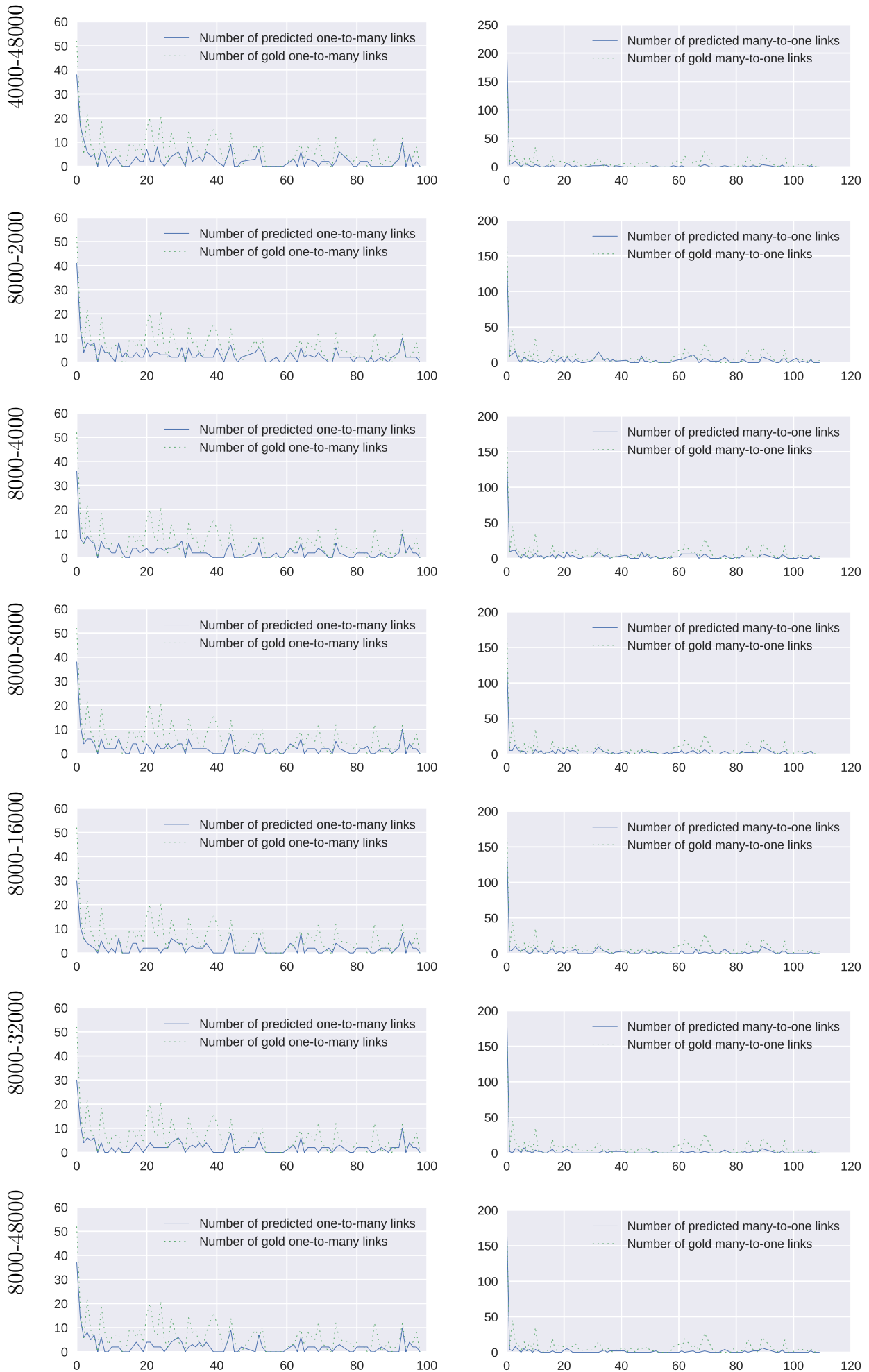


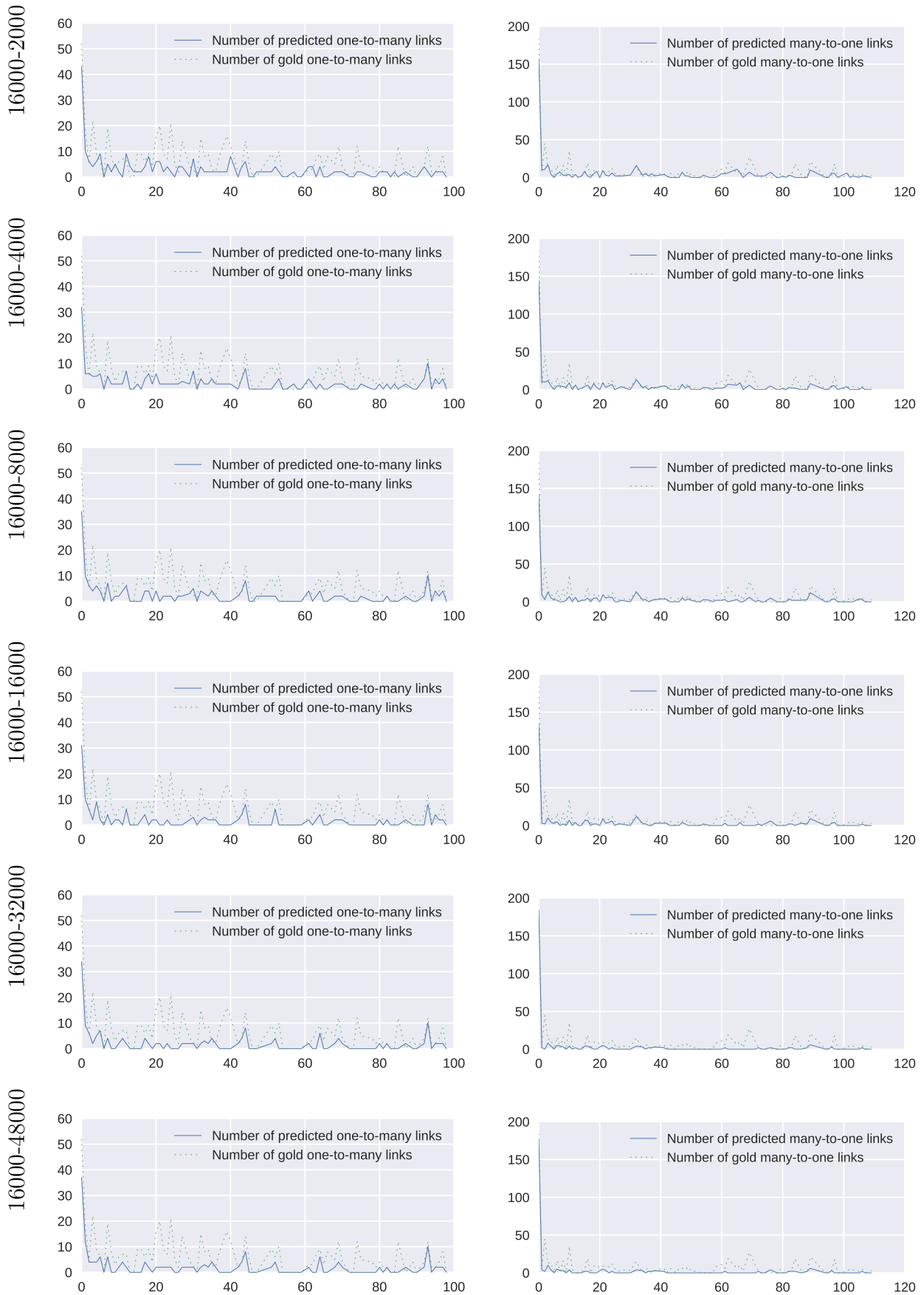
Figure E.85: The direction English-French: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.



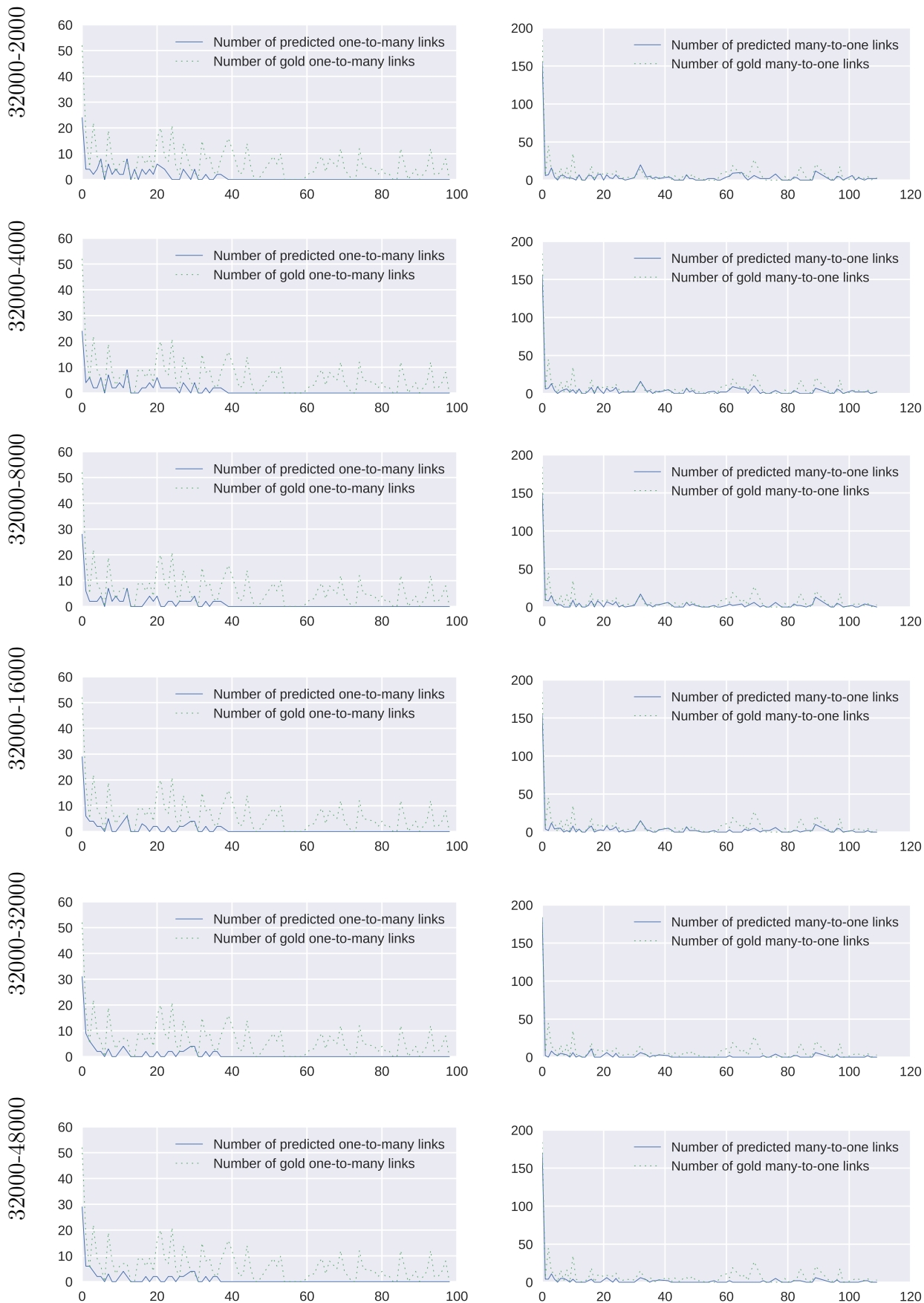


E.7. NUMBER OF ONE-TO-MANY/MANY-TO-ONE AS A FUNCTION OF WORD OCCURRENCE49





E.7. NUMBER OF ONE-TO-MANY/MANY-TO-ONE AS A FUNCTION OF WORD OCCURRENCE49



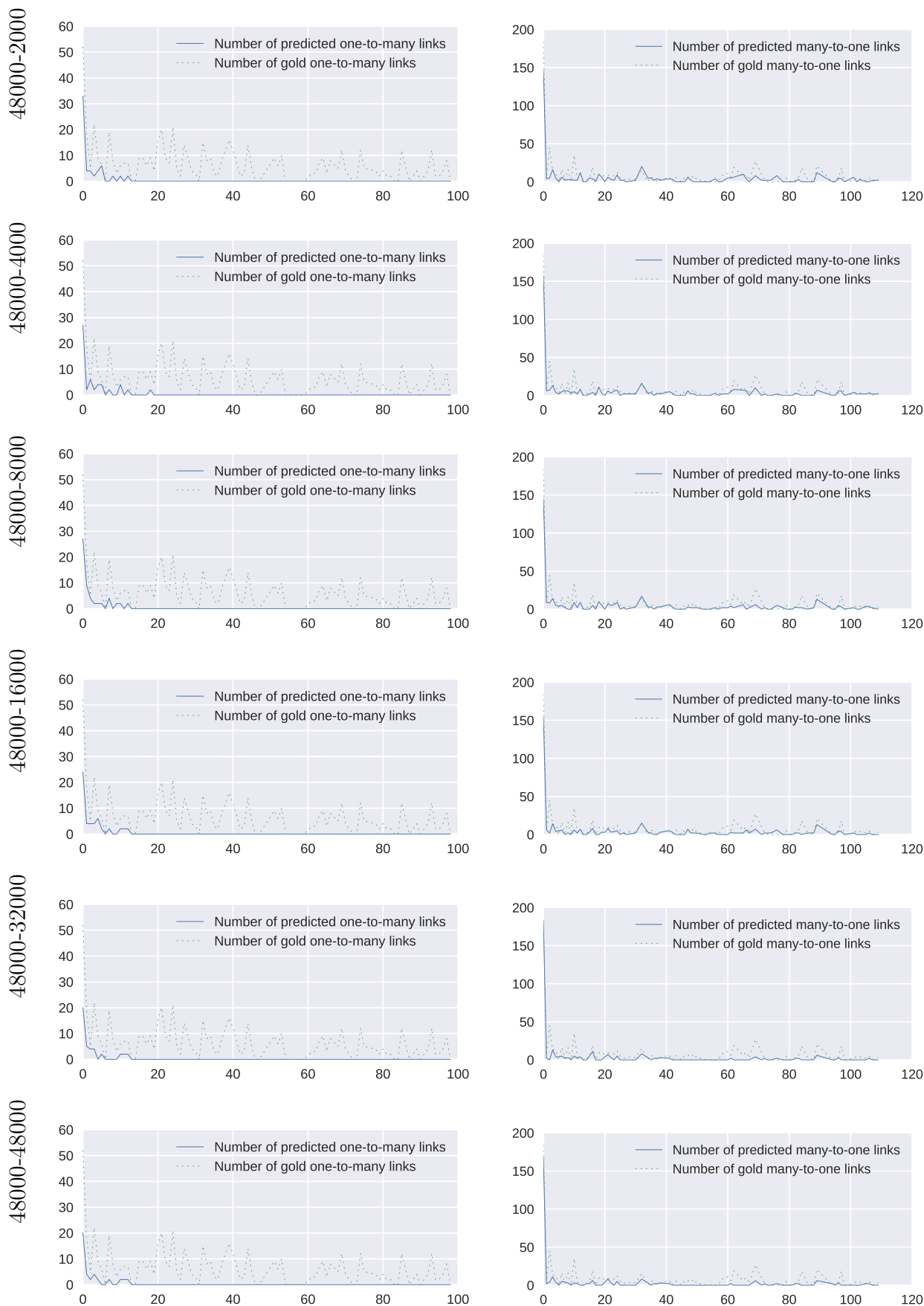
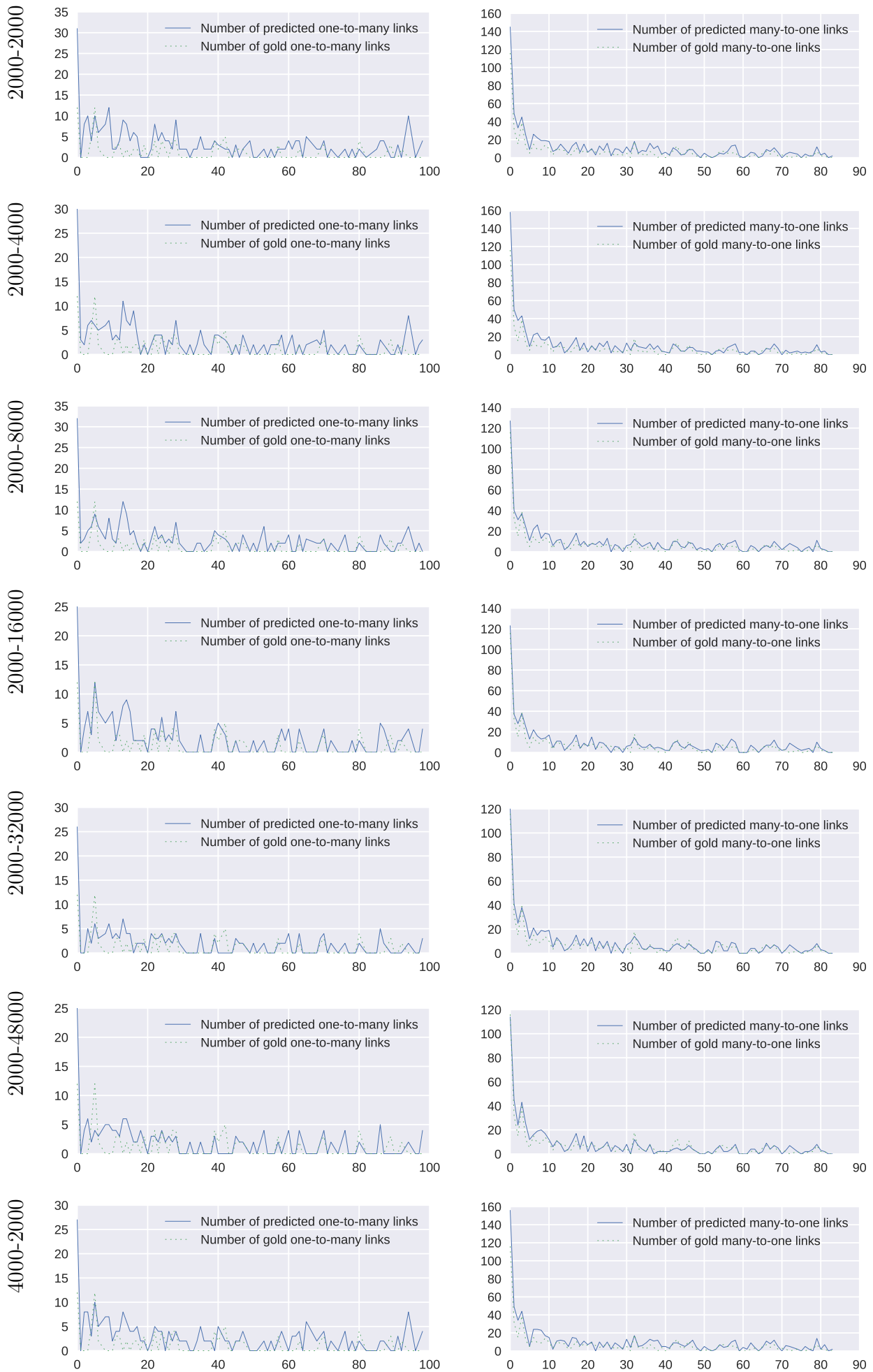


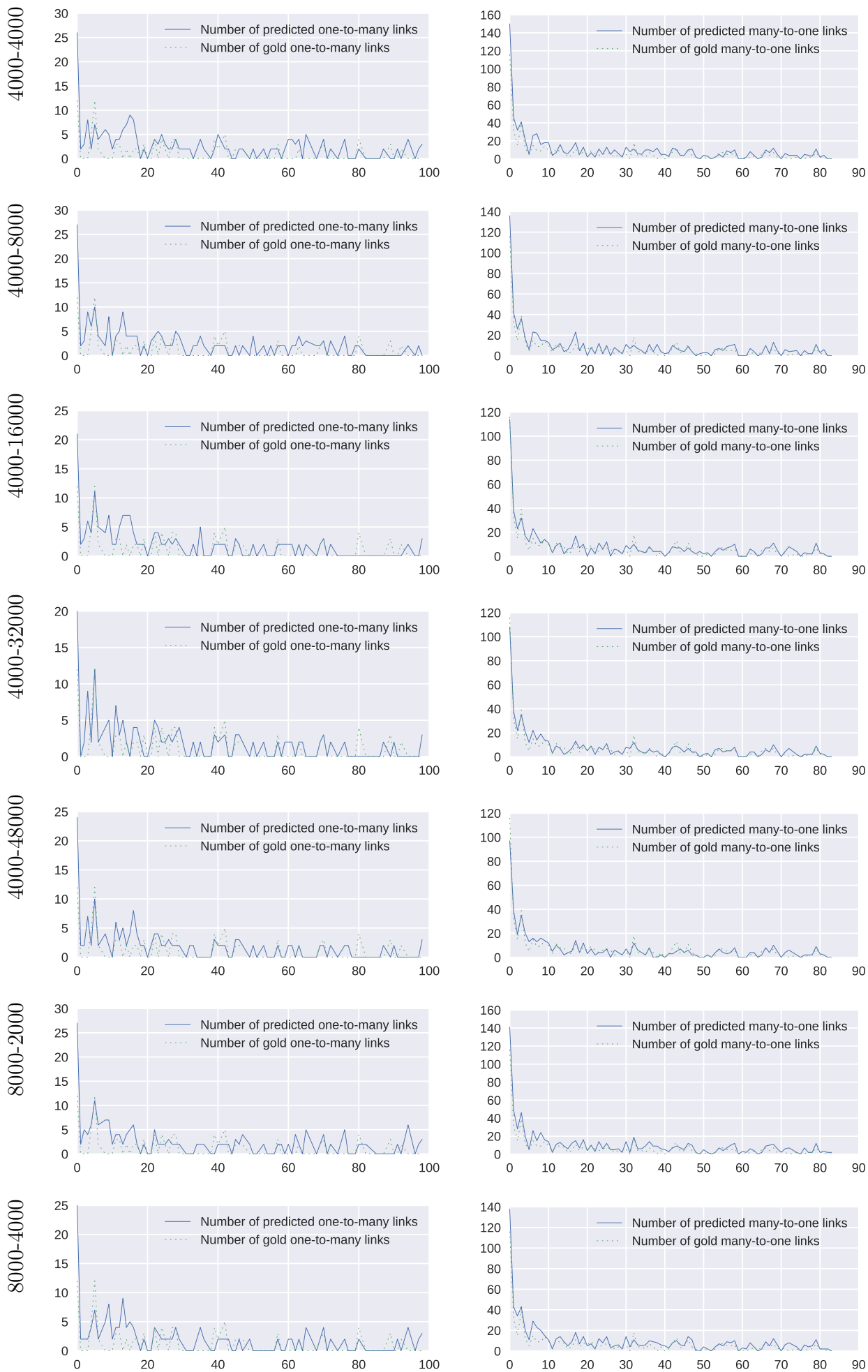
Figure E.86: The direction French-English: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.

Source

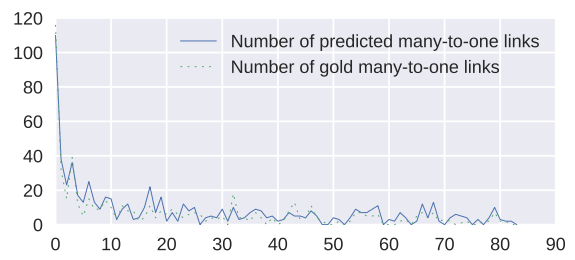
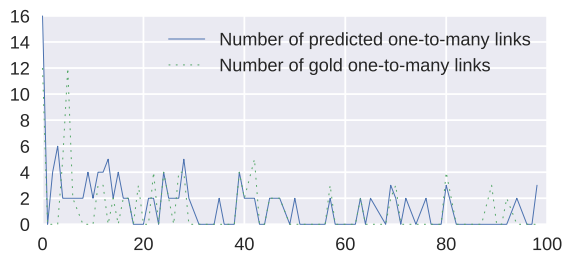
Target

E.7. NUMBER OF ONE-TO-MANY/MANY-TO-ONE AS A FUNCTION OF WORD OCCURRENCE49

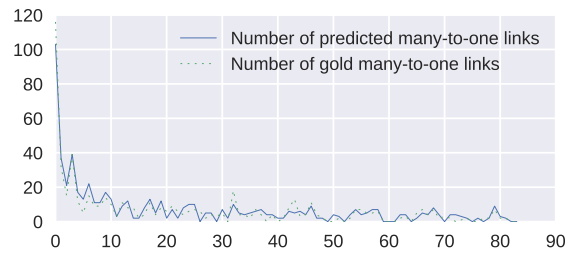
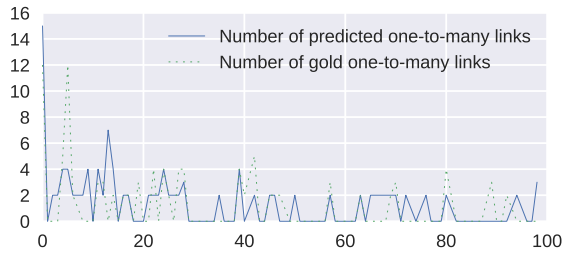




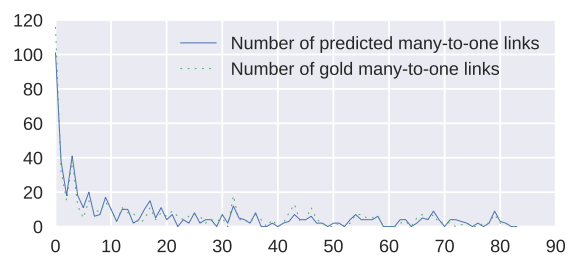
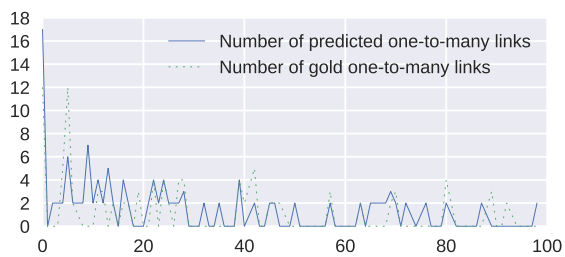
16000-16000



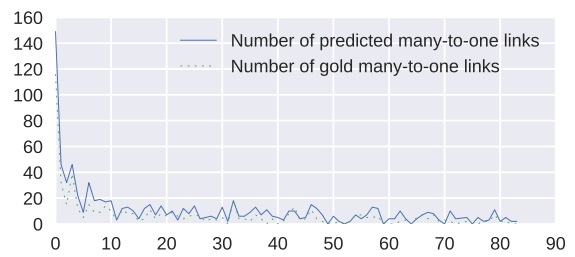
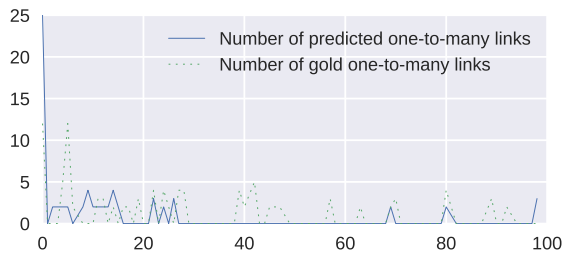
16000-32000



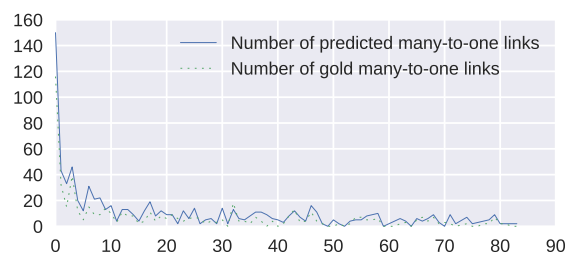
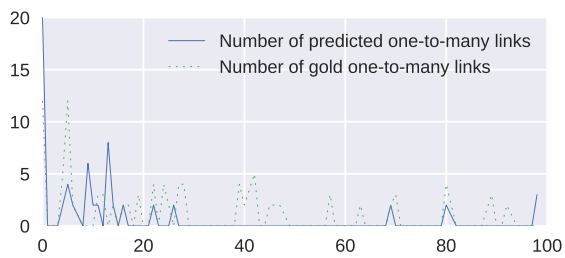
16000-48000



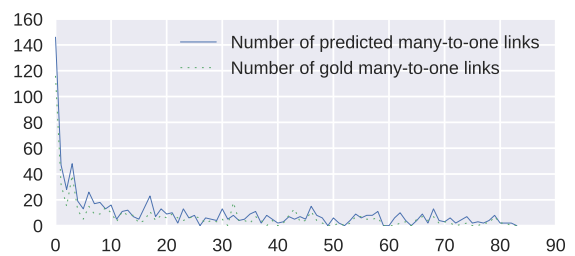
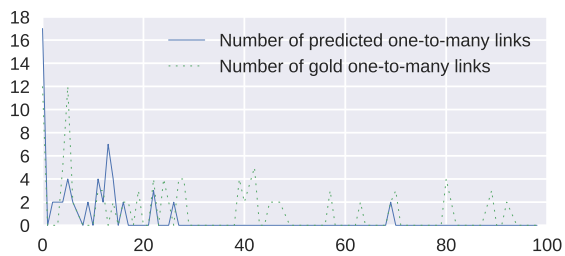
32000-2000



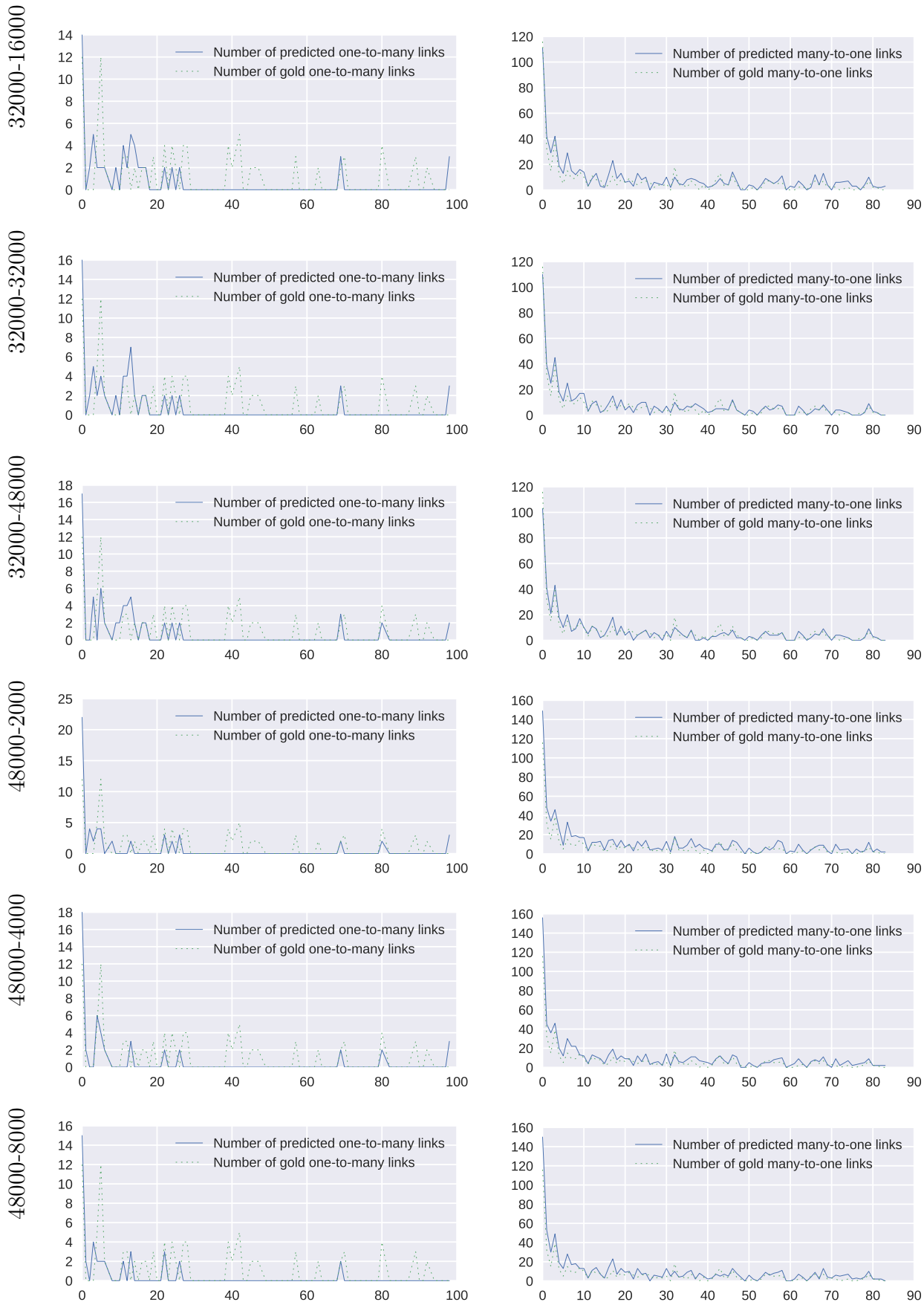
32000-4000



32000-8000



E.7. NUMBER OF ONE-TO-MANY/MANY-TO-ONE AS A FUNCTION OF WORD OCCURRENCE50



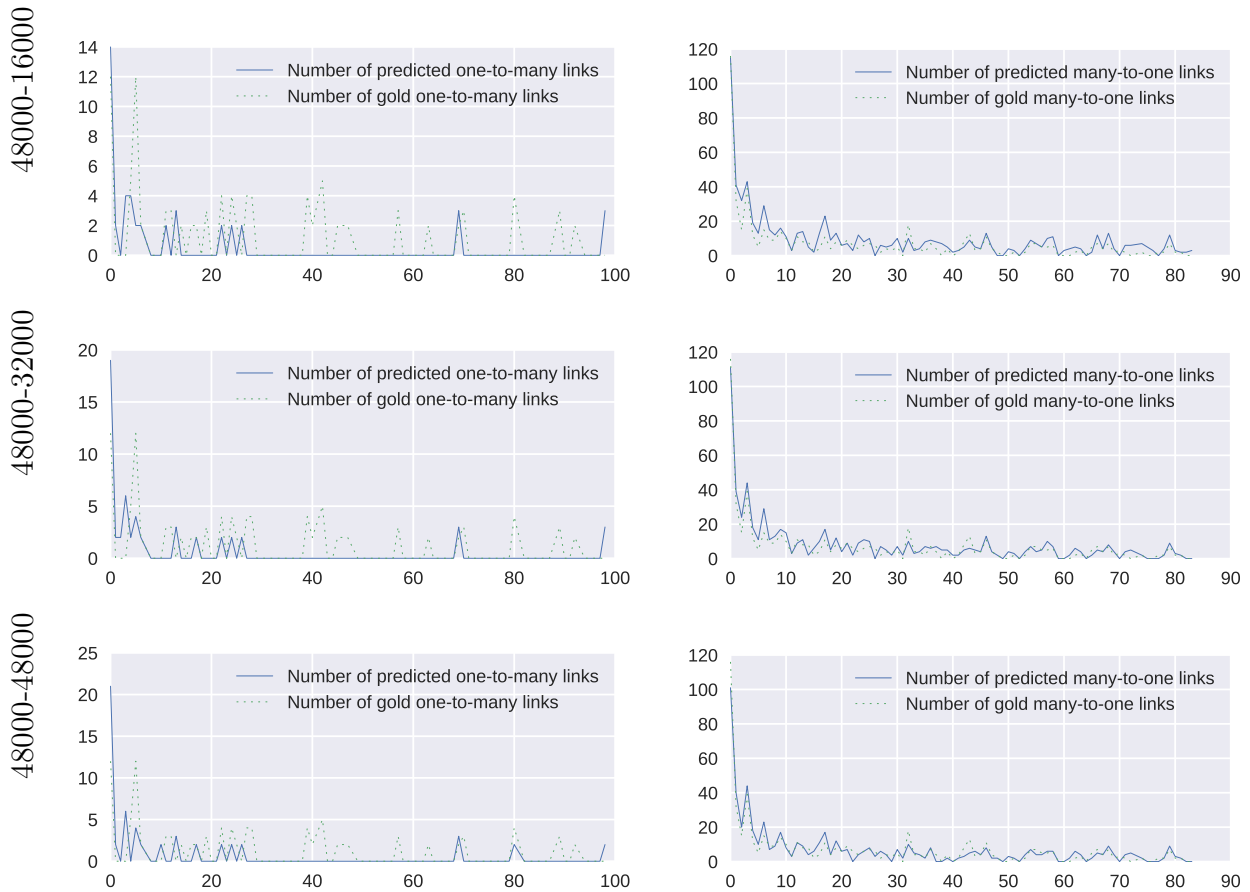
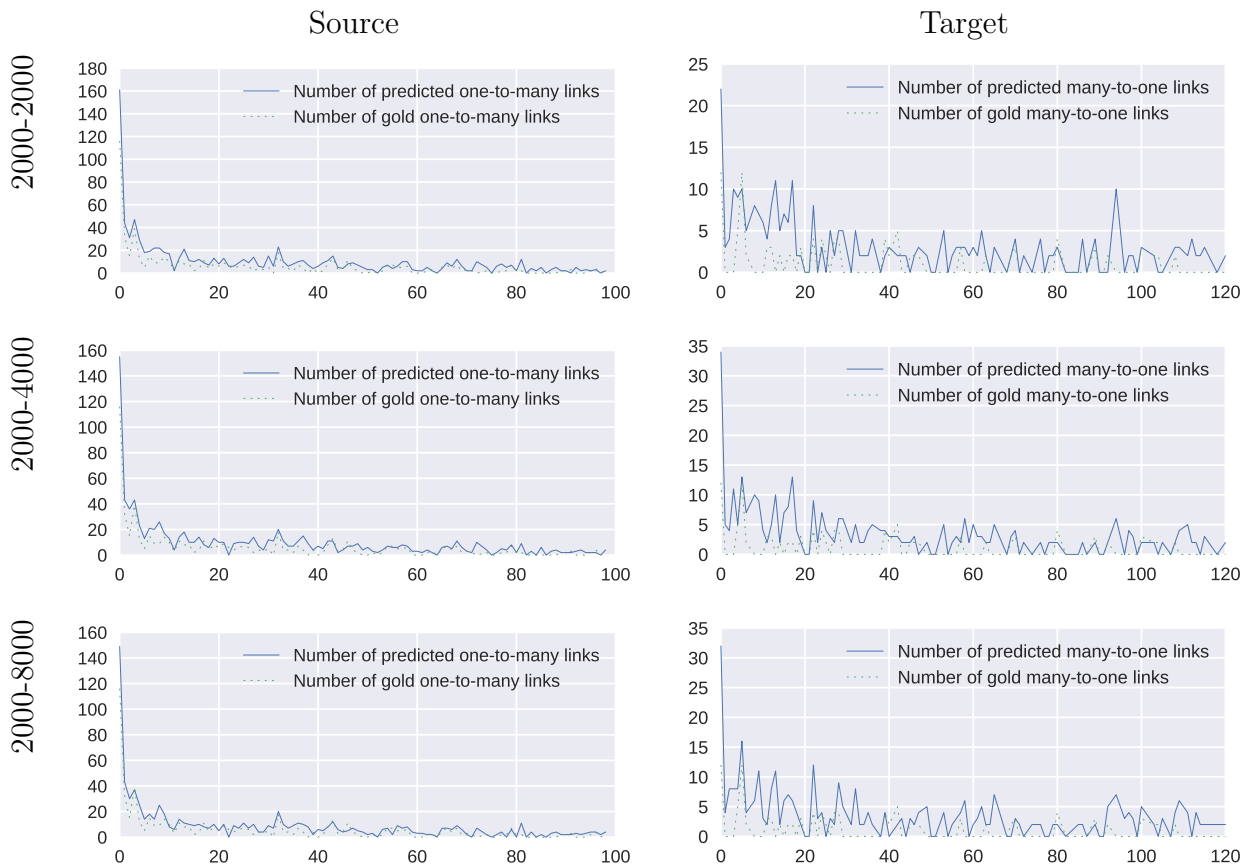
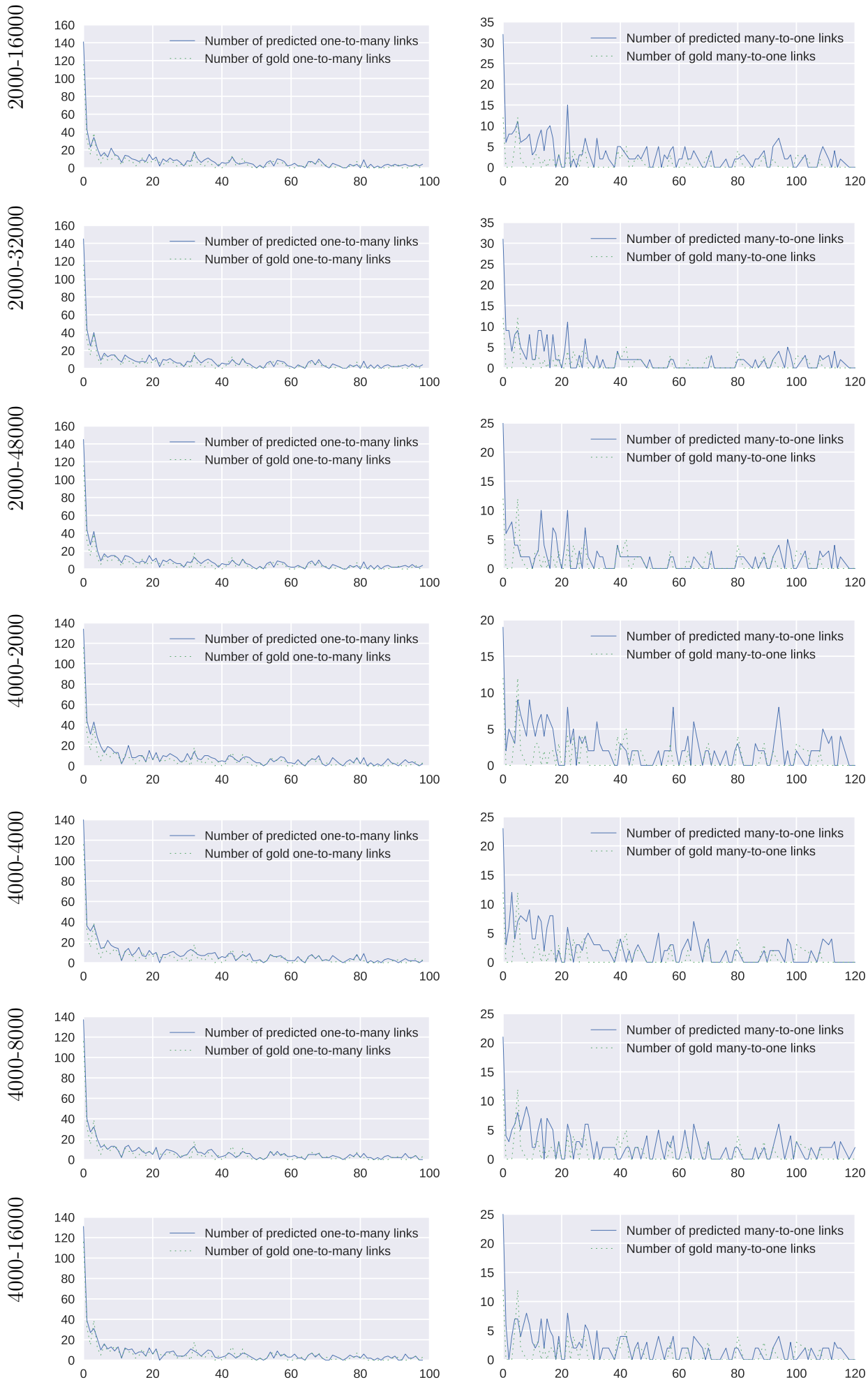
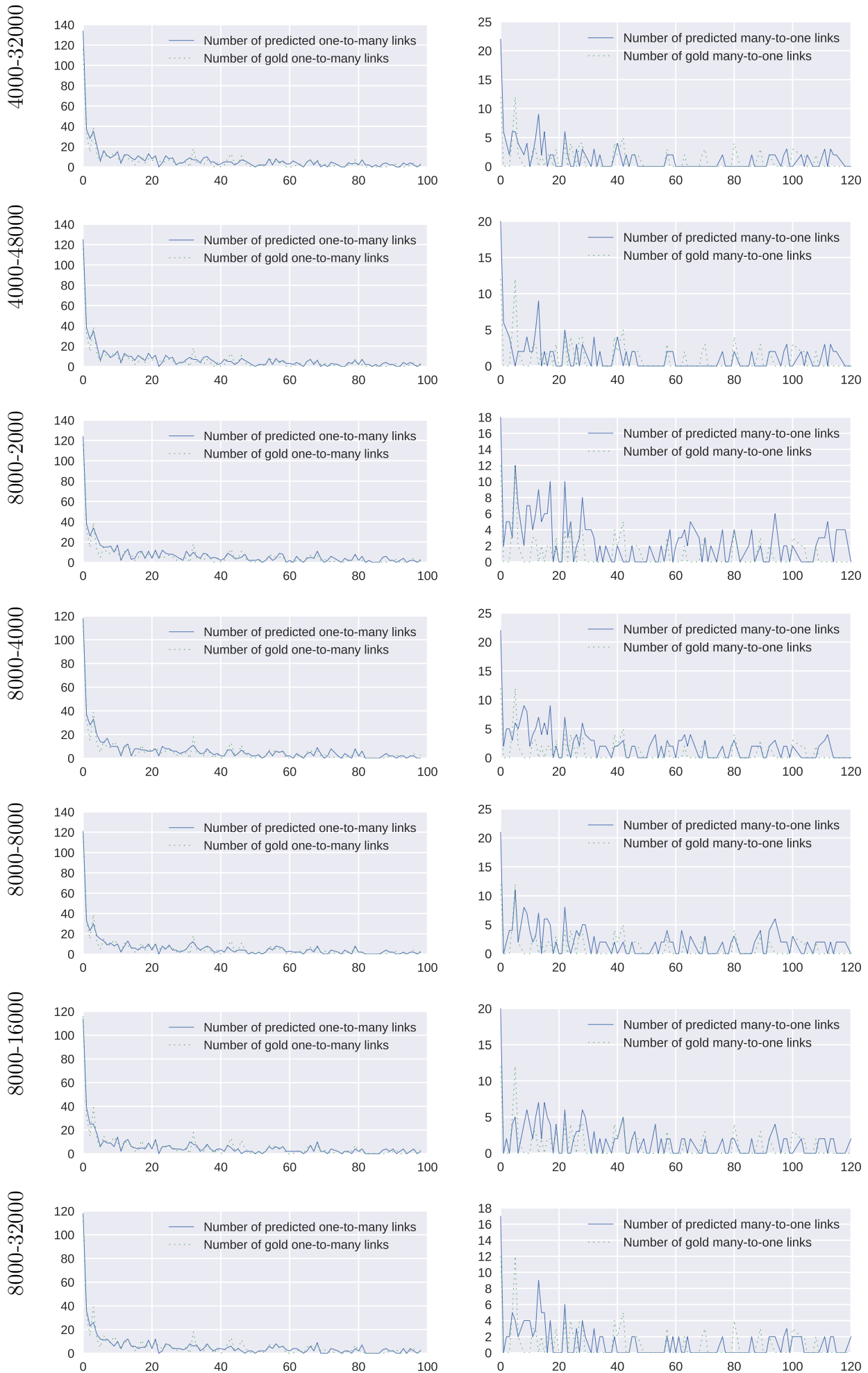


Figure E.87: The direction English-German: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.

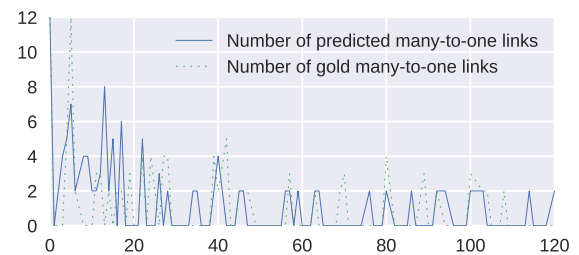
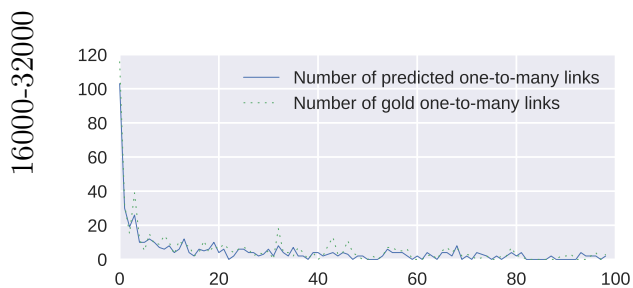
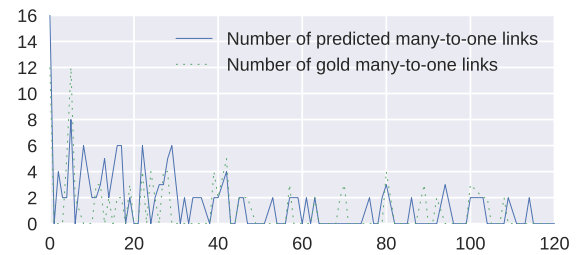
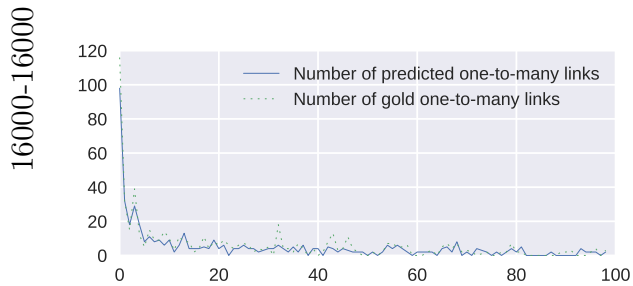
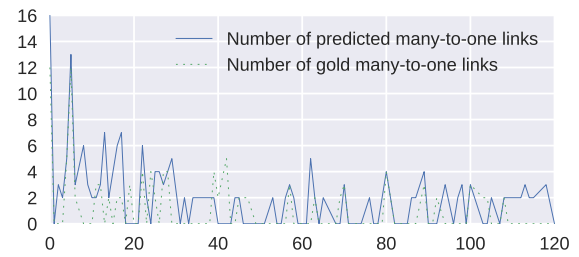
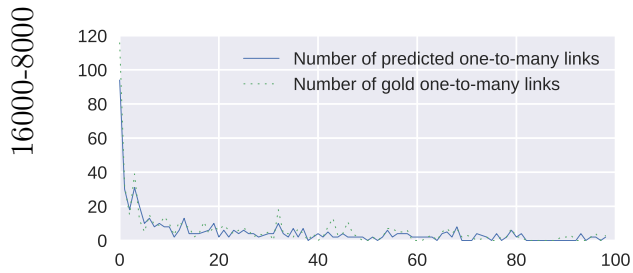
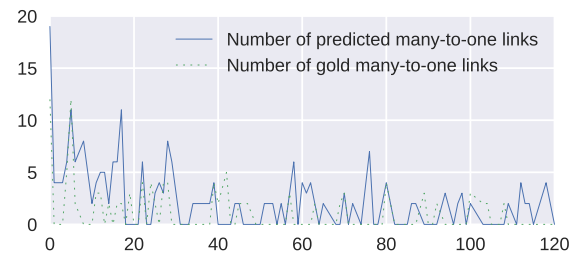
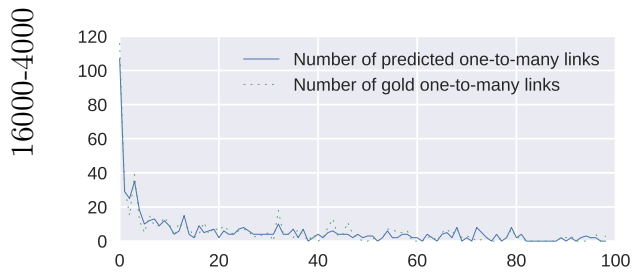
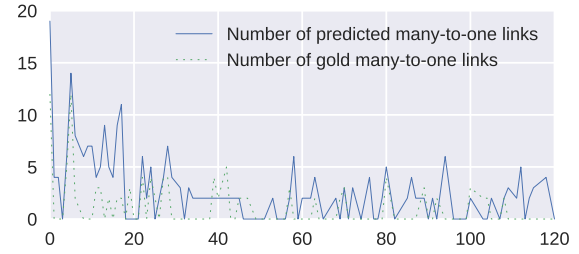
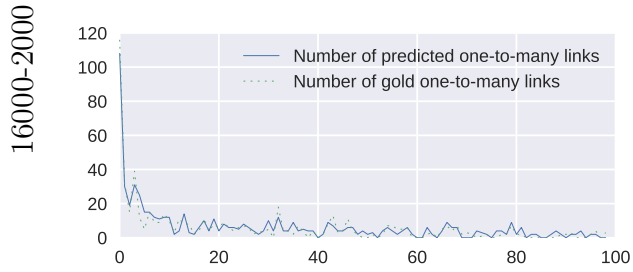
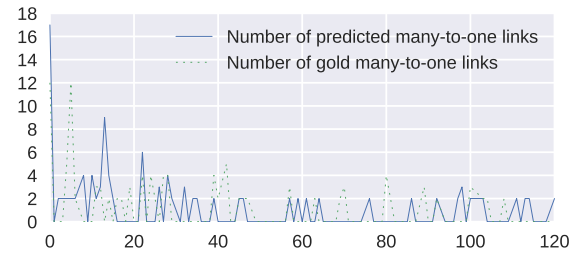
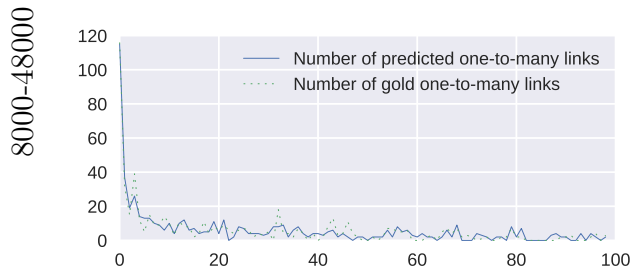


E.7. NUMBER OF ONE-TO-MANY/MANY-TO-ONE AS A FUNCTION OF WORD OCCURRENCE50

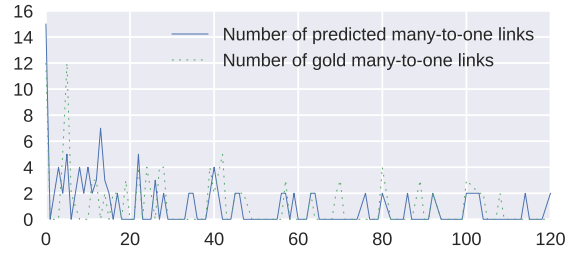
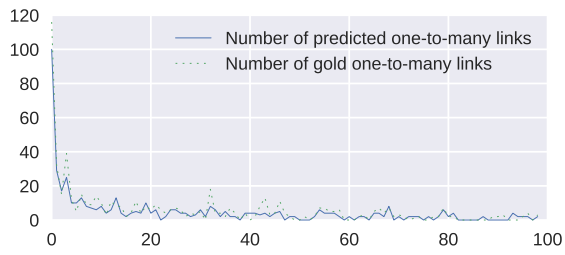




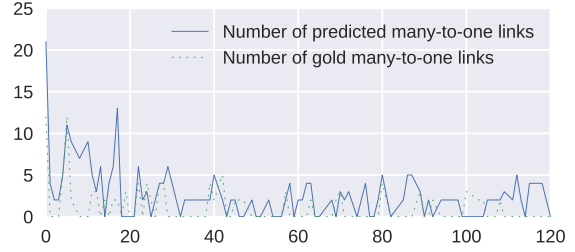
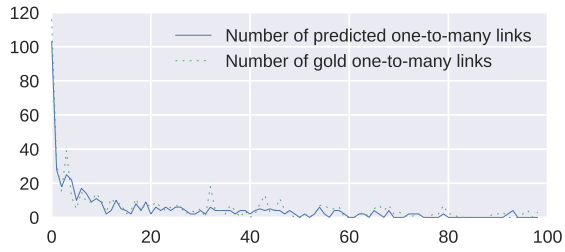
E.7. NUMBER OF ONE-TO-MANY/MANY-TO-ONE AS A FUNCTION OF WORD OCCURRENCE50



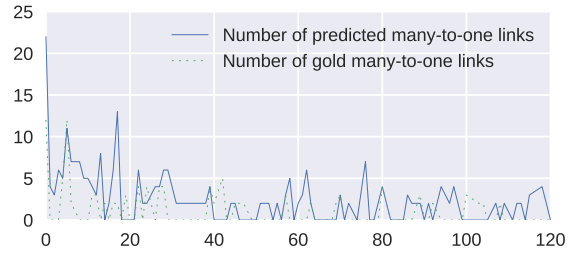
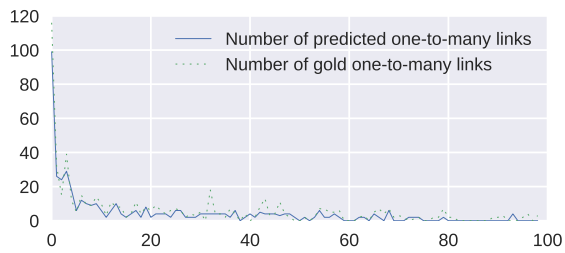
16000-48000



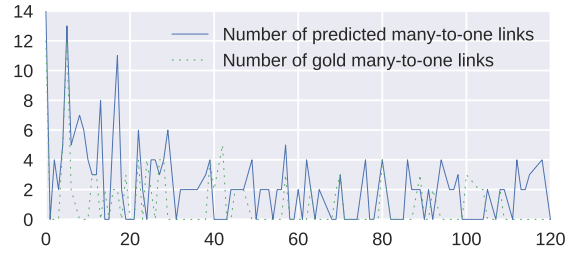
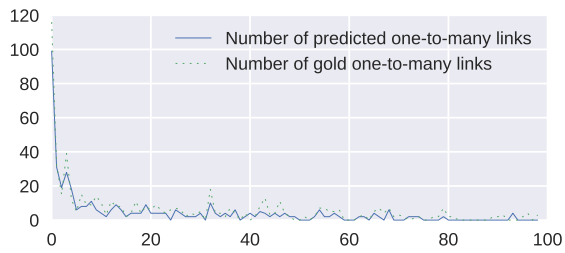
32000-20000



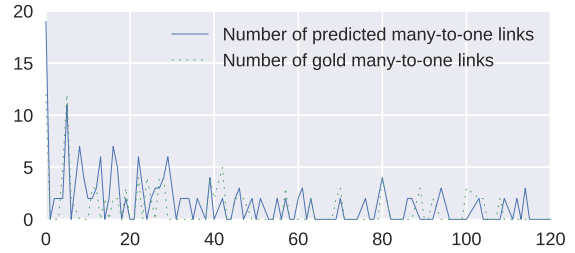
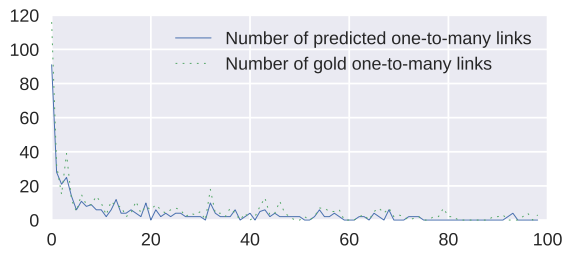
32000-4000



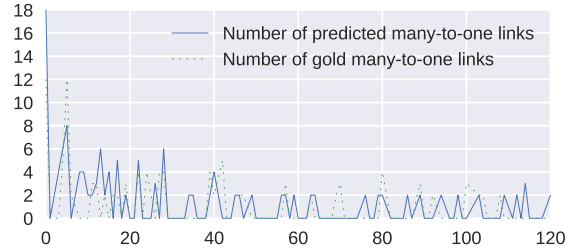
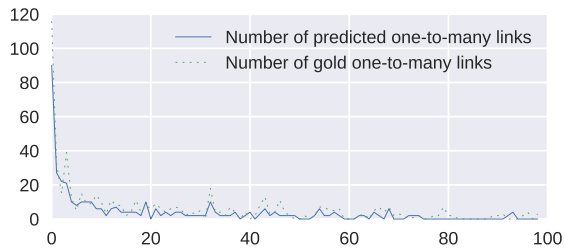
32000-8000



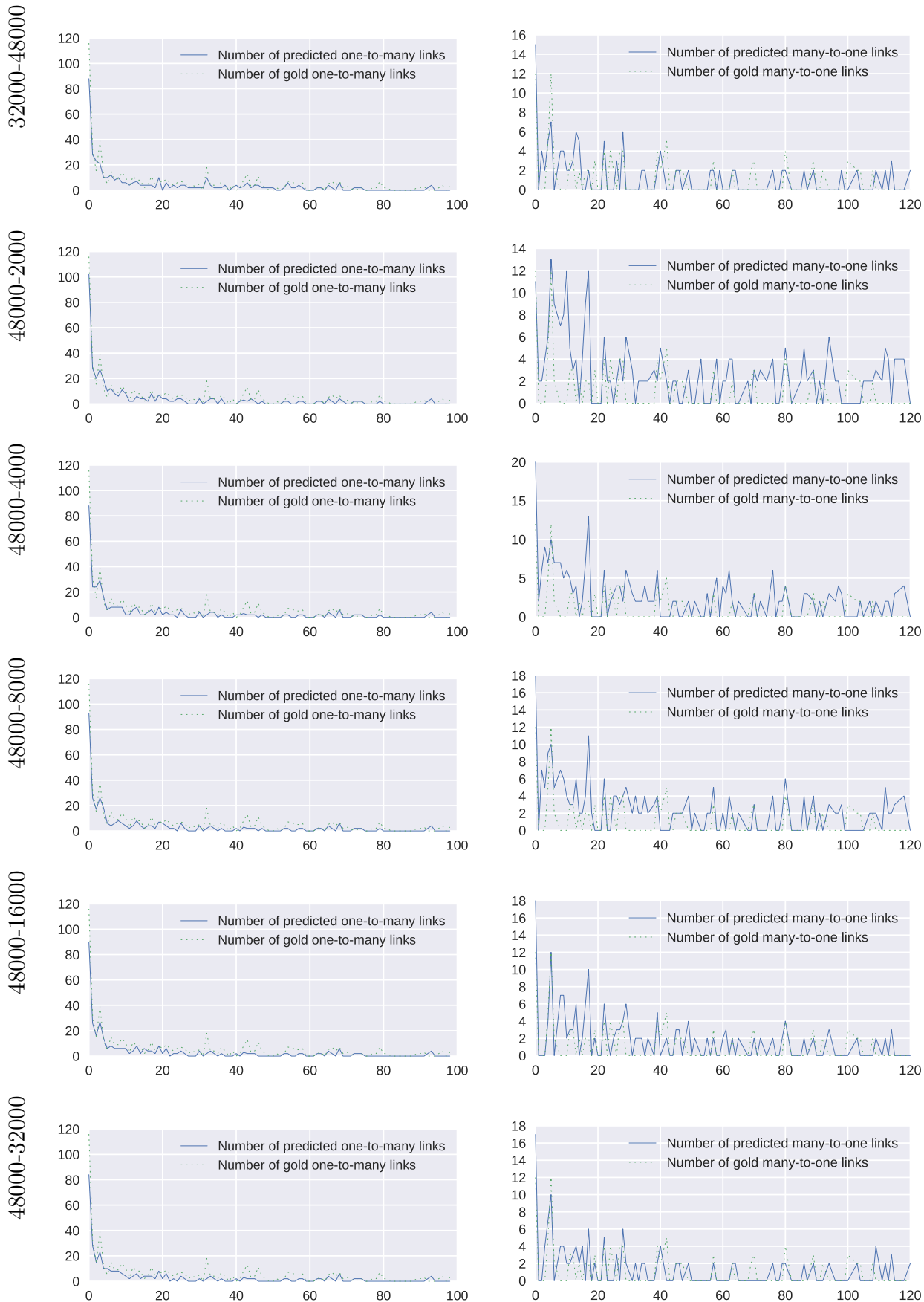
32000-16000



32000-32000



E.7. NUMBER OF ONE-TO-MANY/MANY-TO-ONE AS A FUNCTION OF WORD OCCURRENCE50



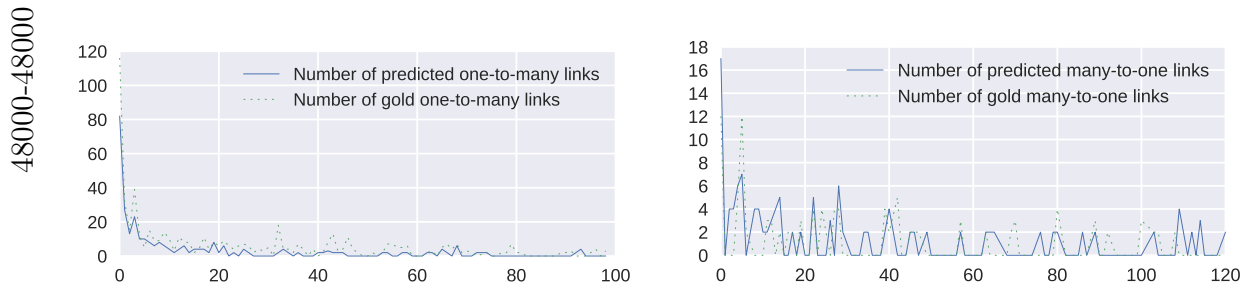
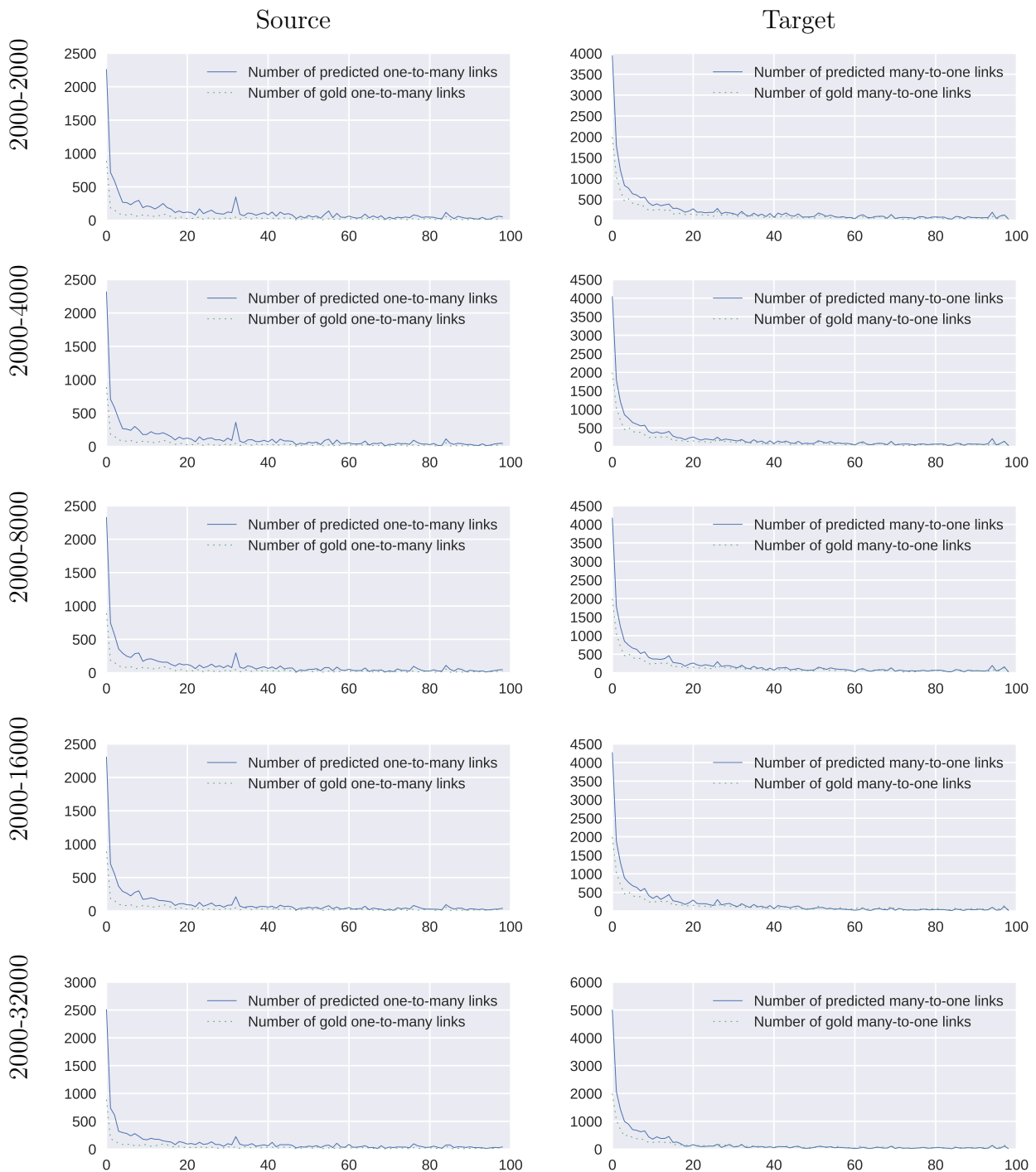
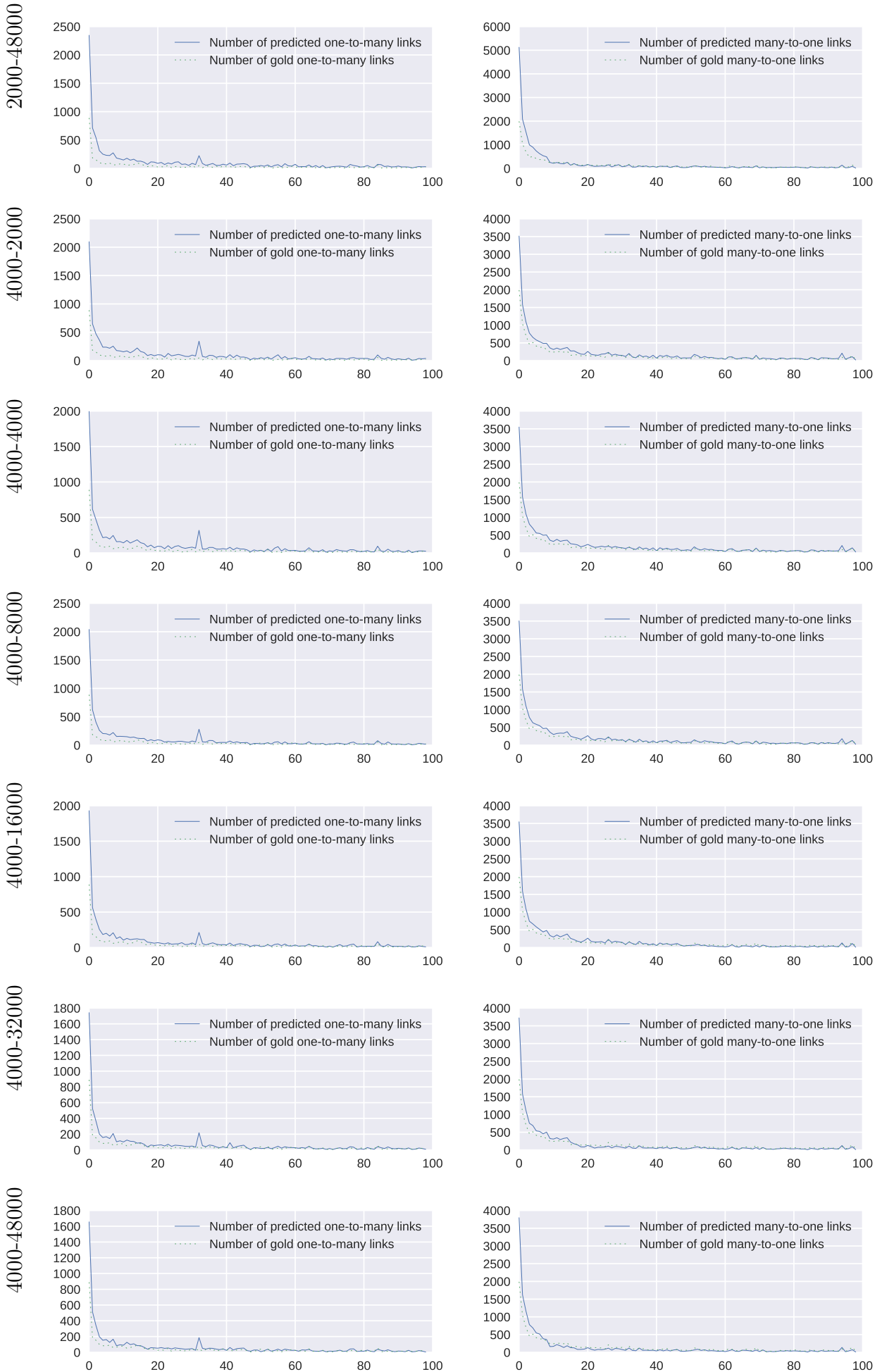
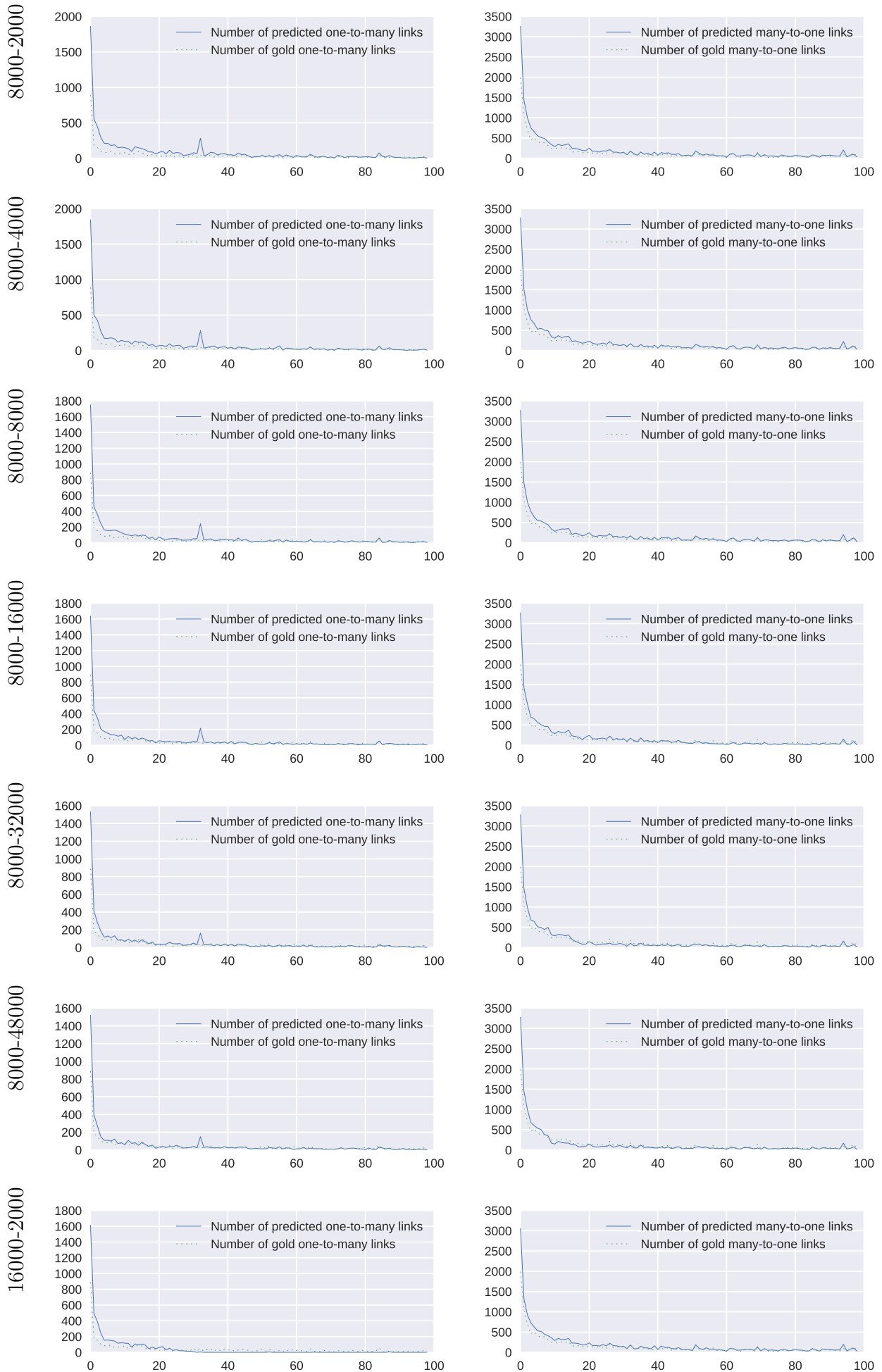


Figure E.88: The direction German-English: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.

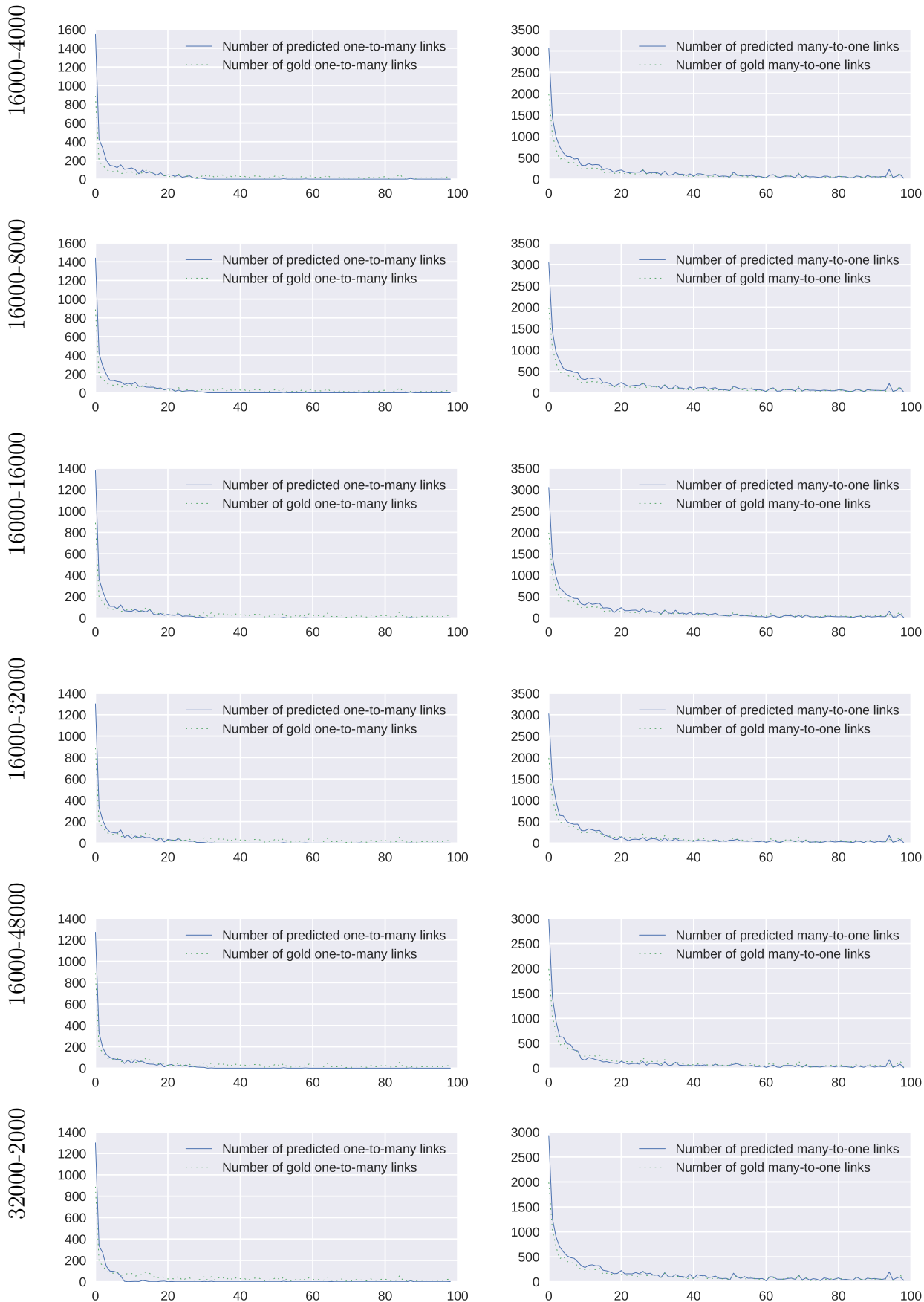


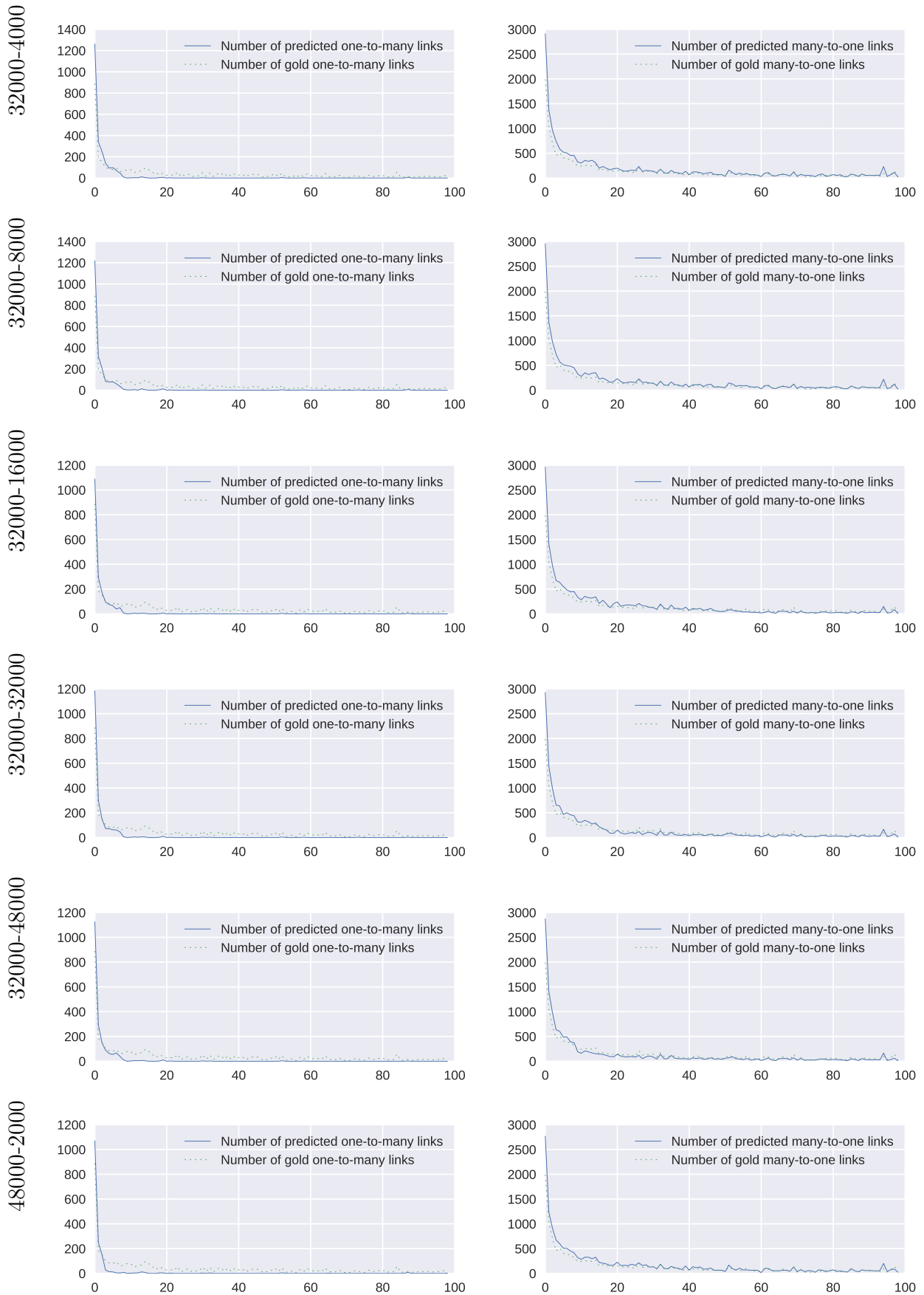
E.7. NUMBER OF ONE-TO-MANY/MANY-TO-ONE AS A FUNCTION OF WORD OCCURRENCE⁵⁰





E.7. NUMBER OF ONE-TO-MANY/MANY-TO-ONE AS A FUNCTION OF WORD OCCURRENCE51





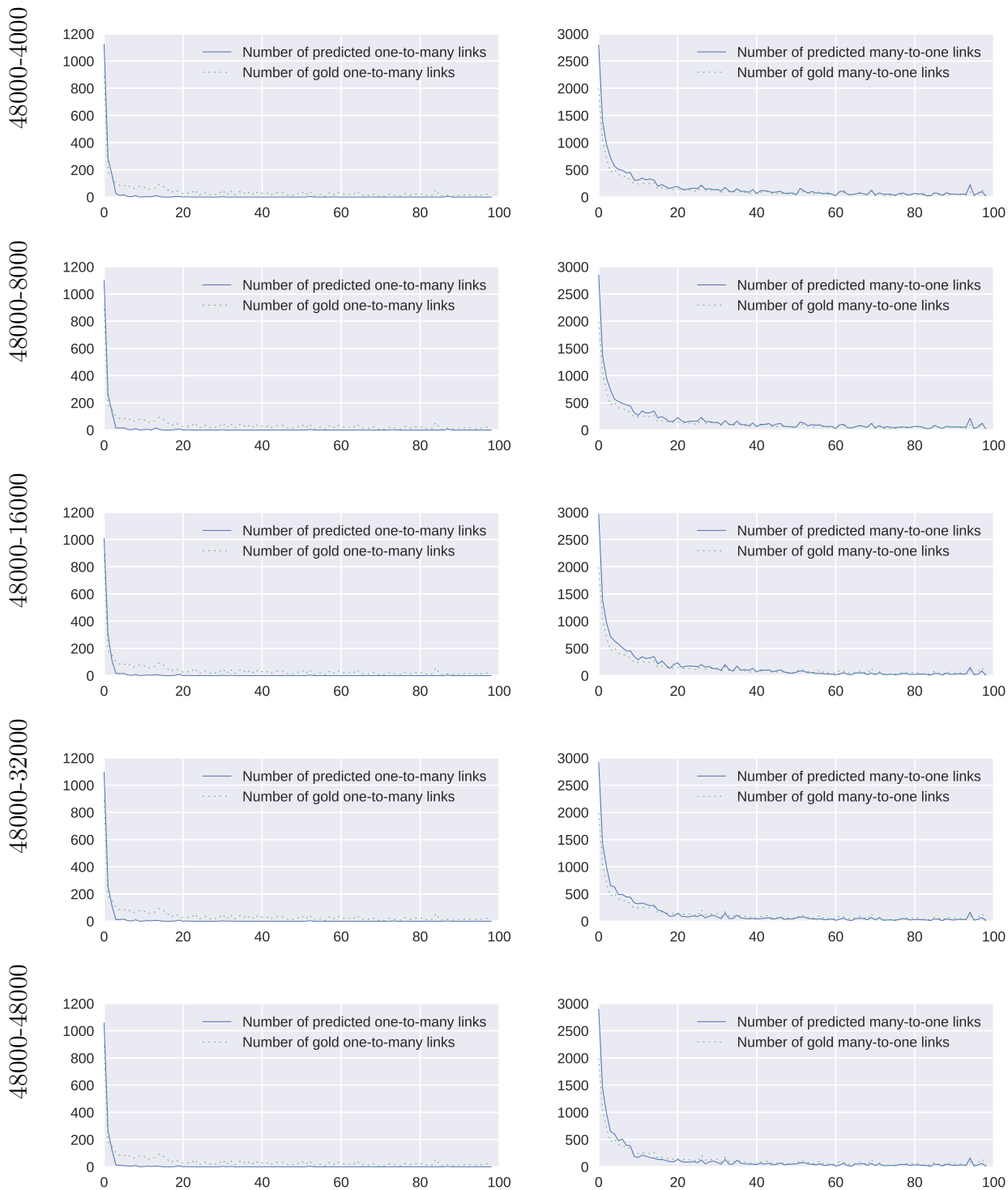
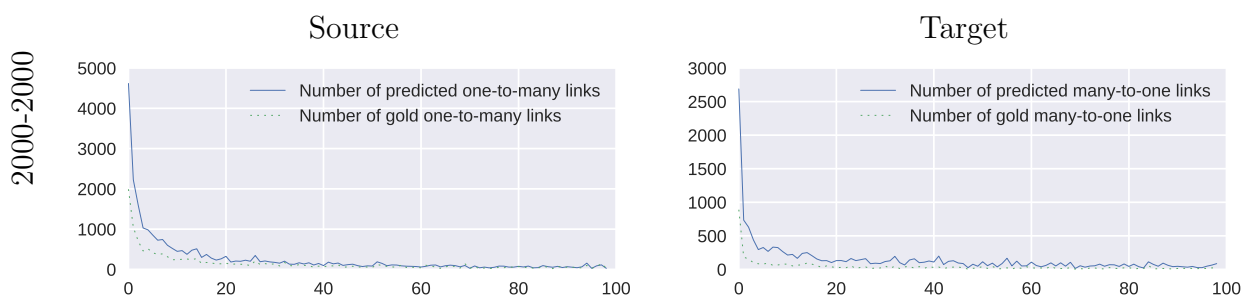
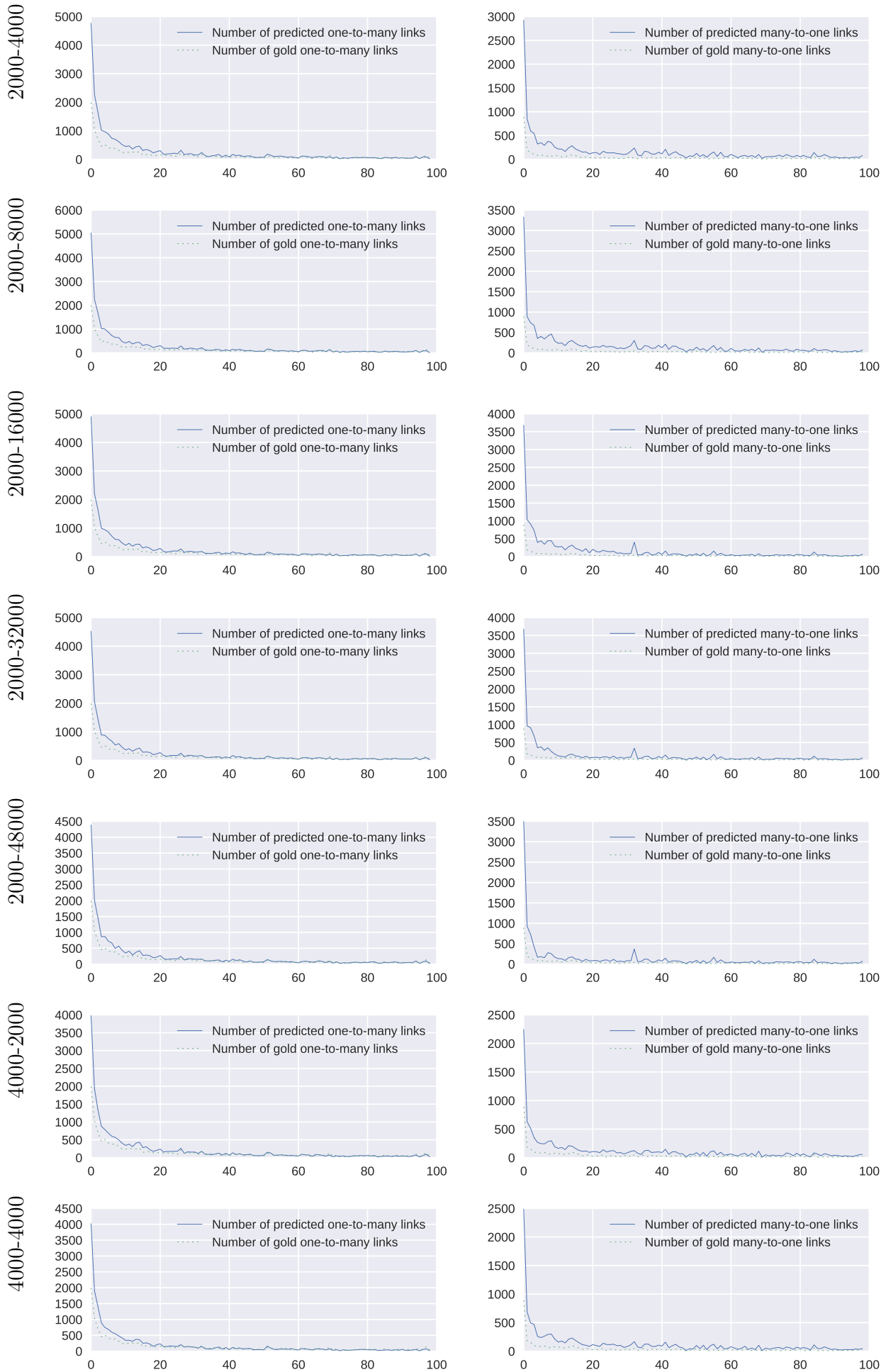
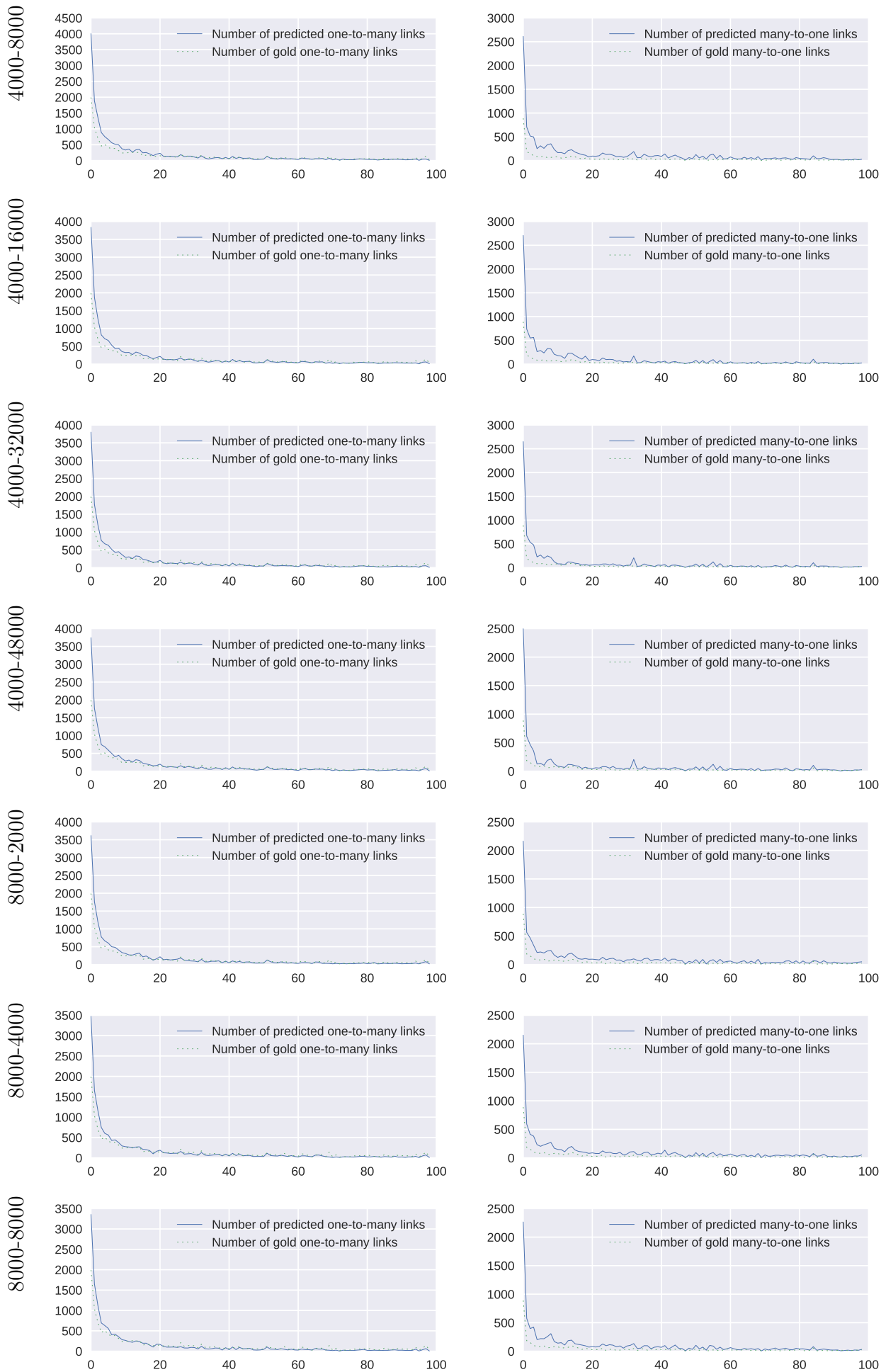


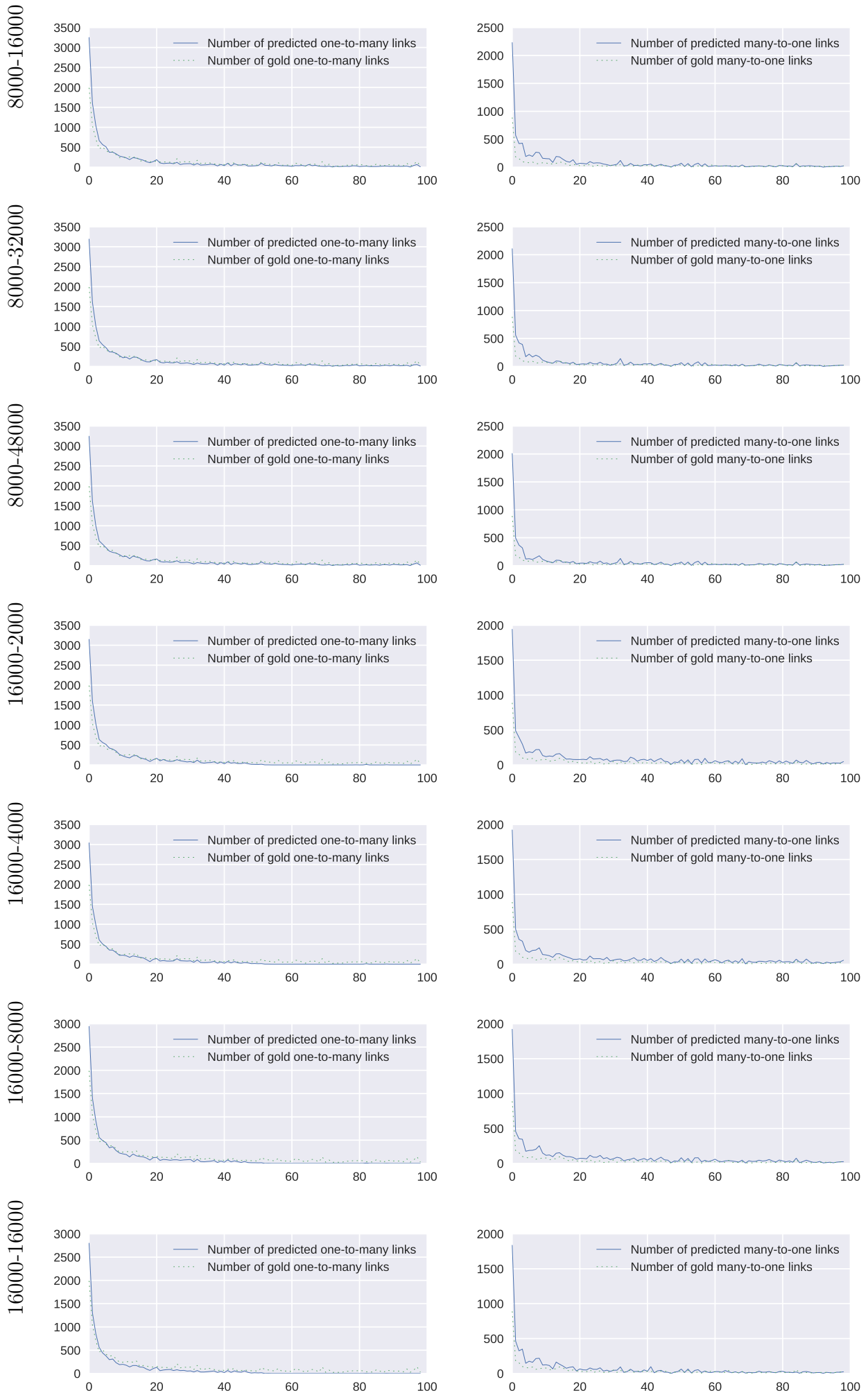
Figure E.89: The direction English-Czech: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.



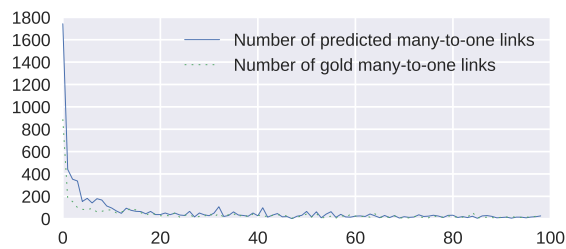
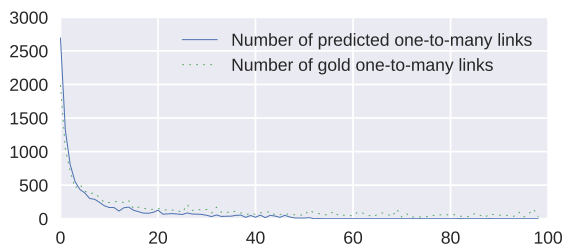


E.7. NUMBER OF ONE-TO-MANY/MANY-TO-ONE AS A FUNCTION OF WORD OCCURRENCE51

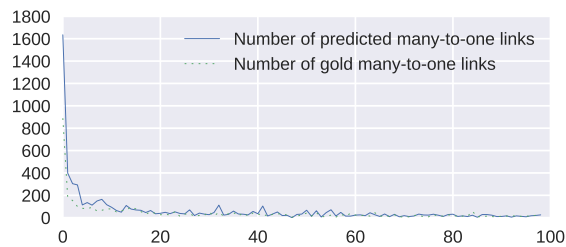
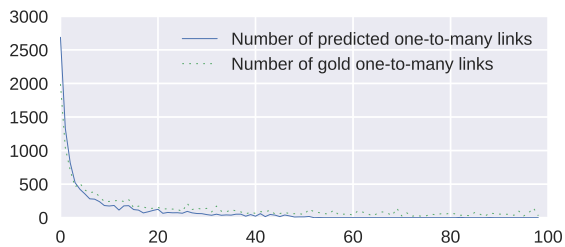




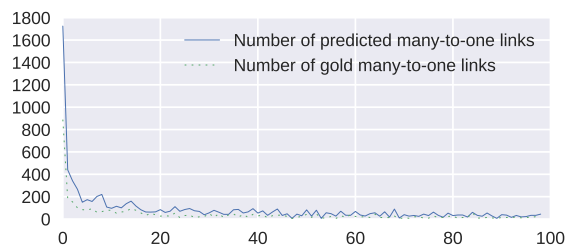
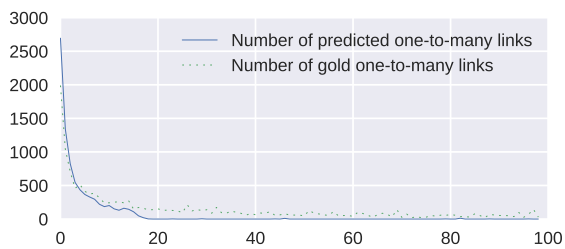
16000-32000



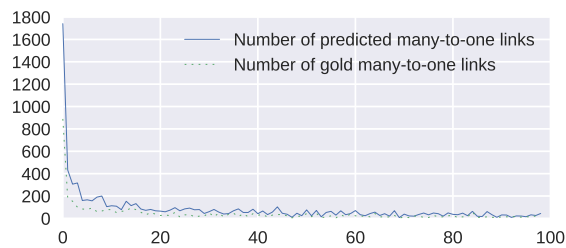
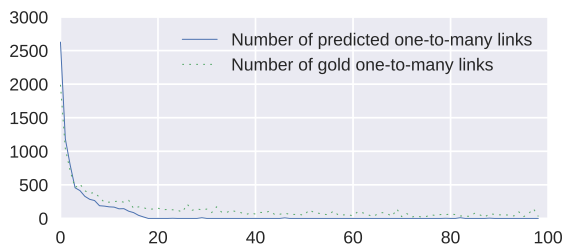
16000-48000



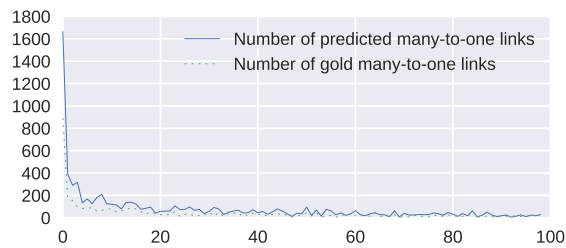
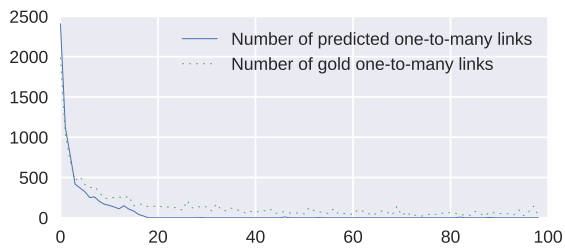
32000-20000



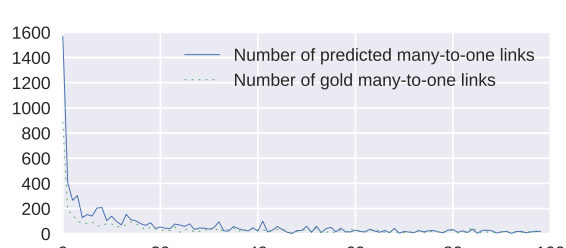
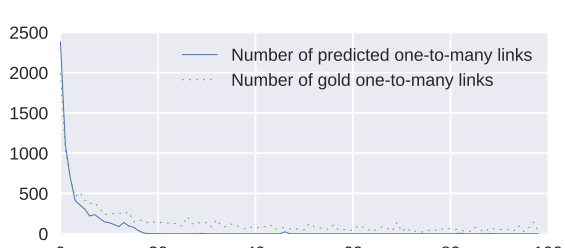
32000-40000



32000-80000



32000-160000



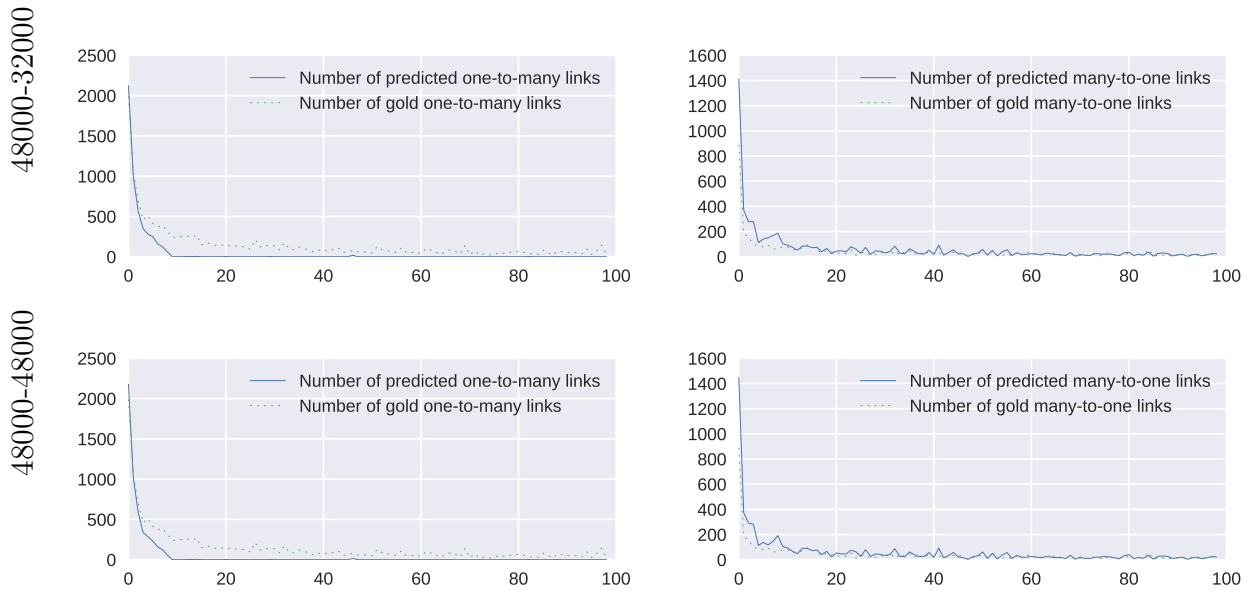
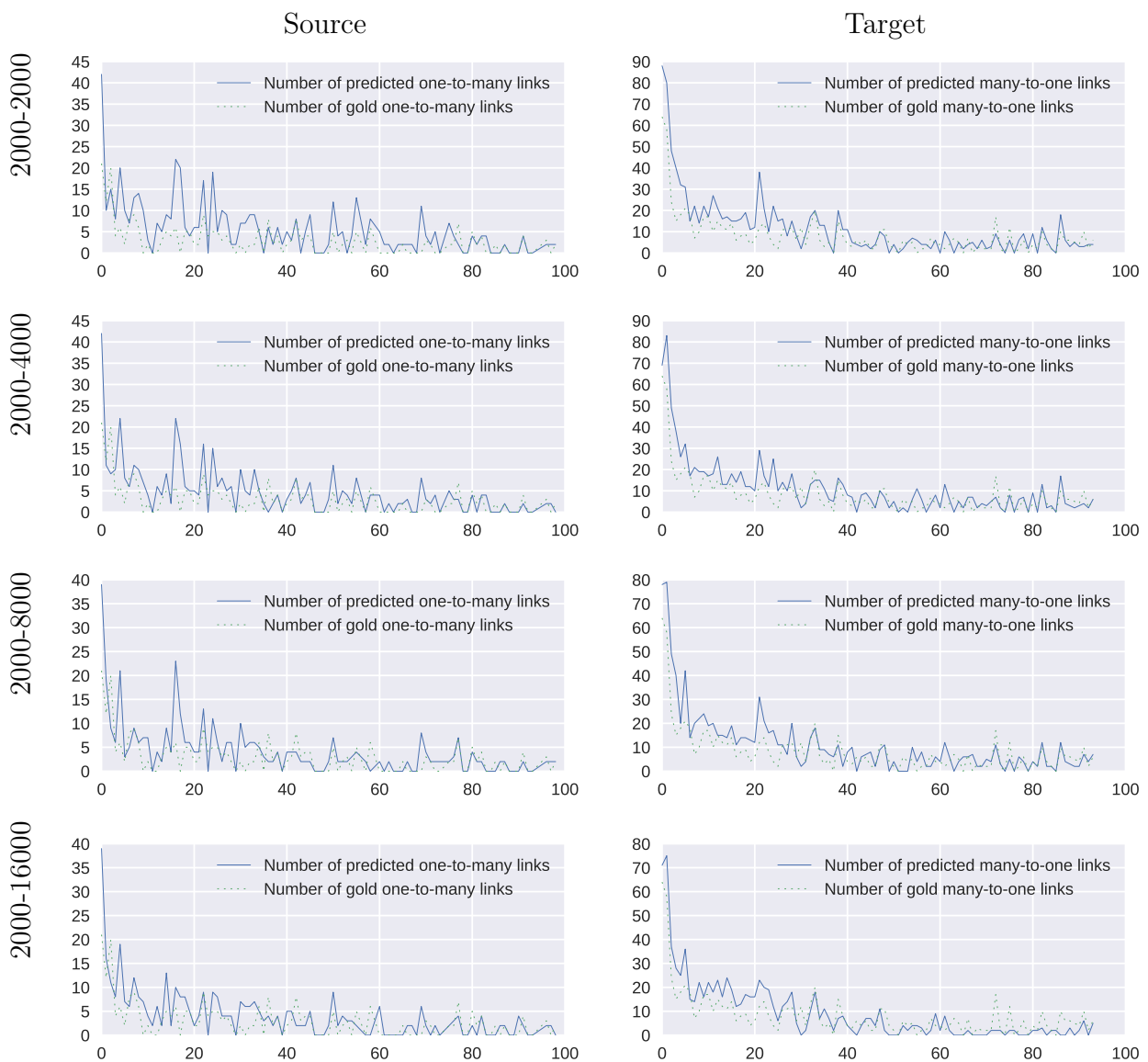
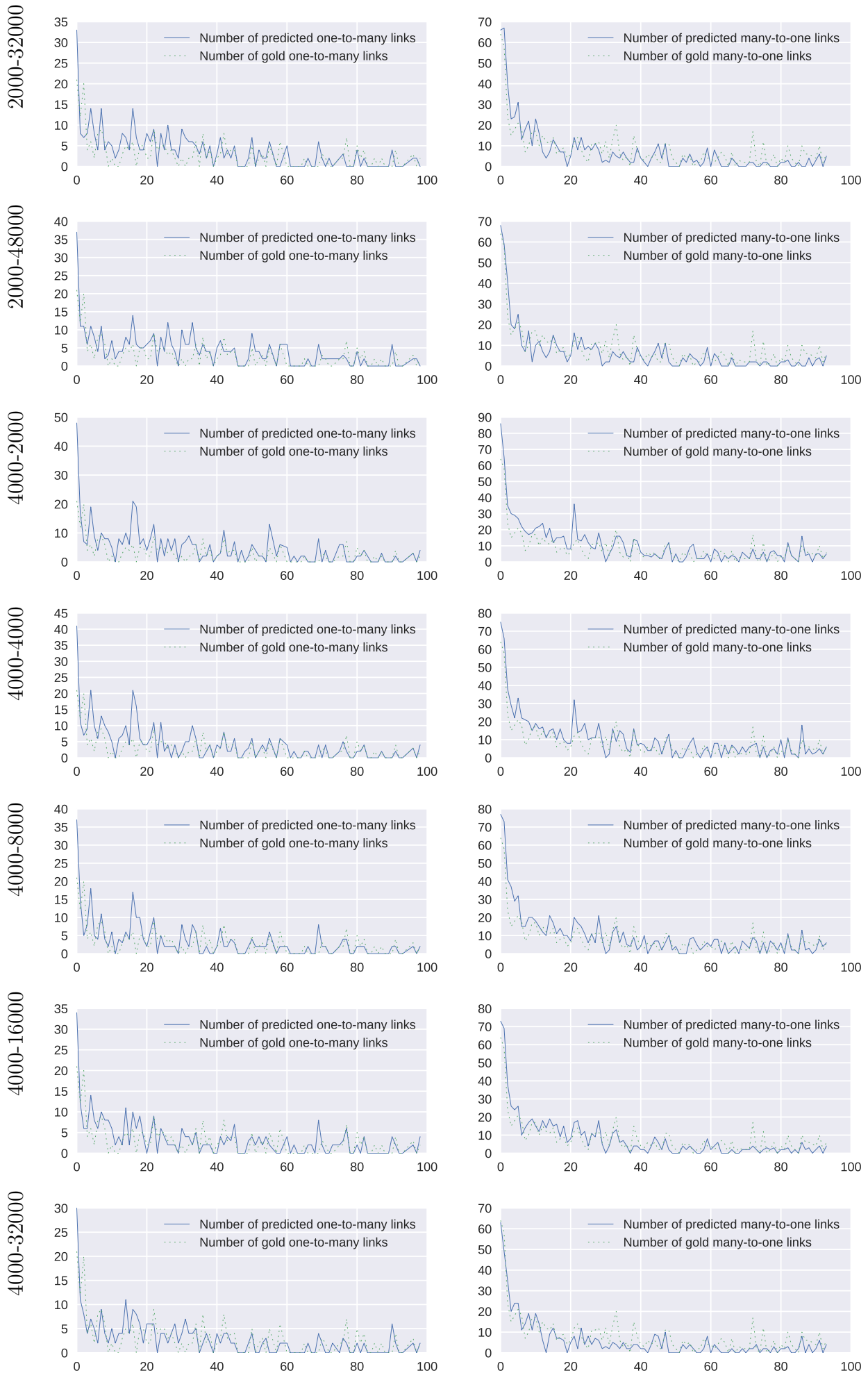


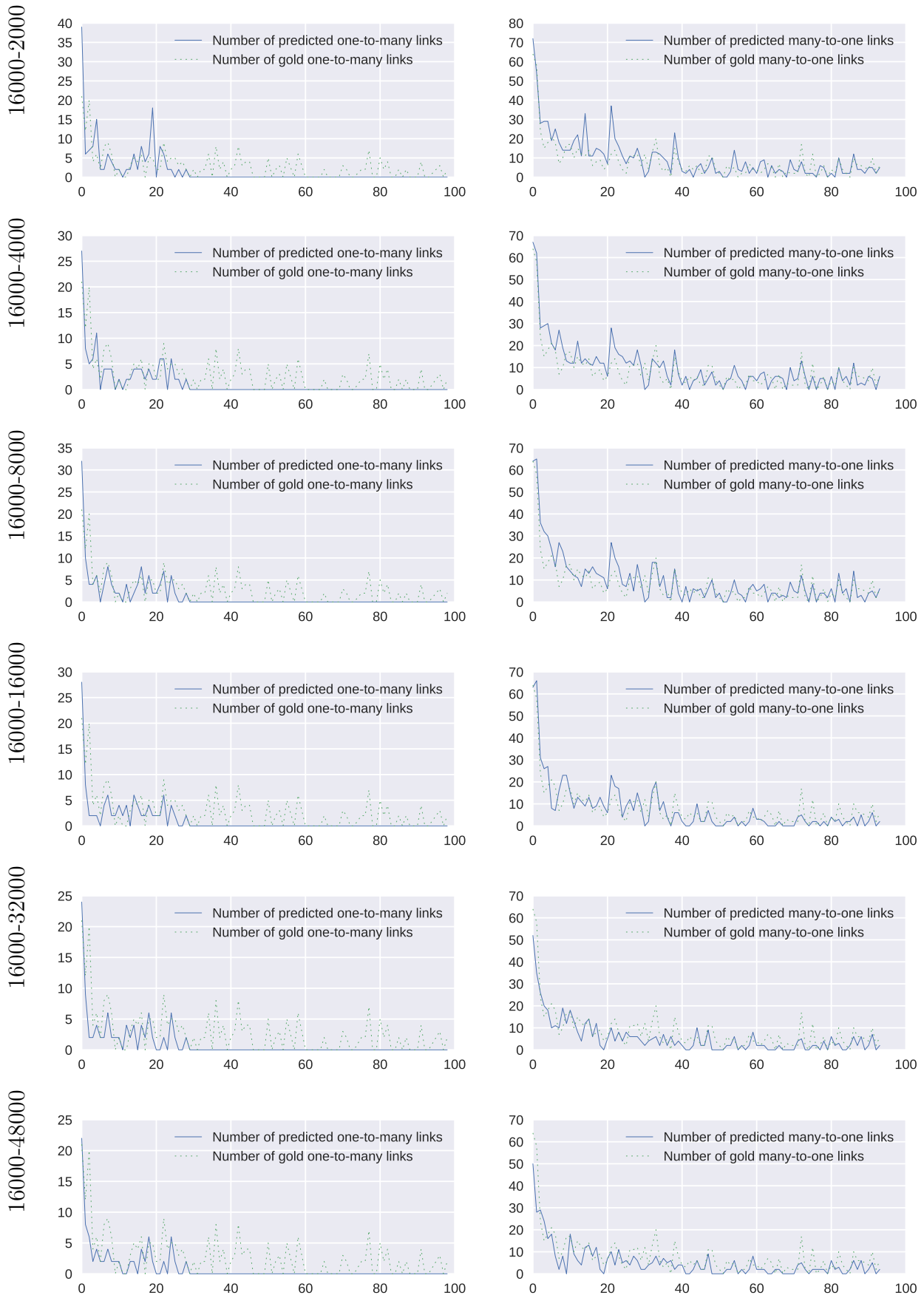
Figure E.90: The direction Czech-English: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.

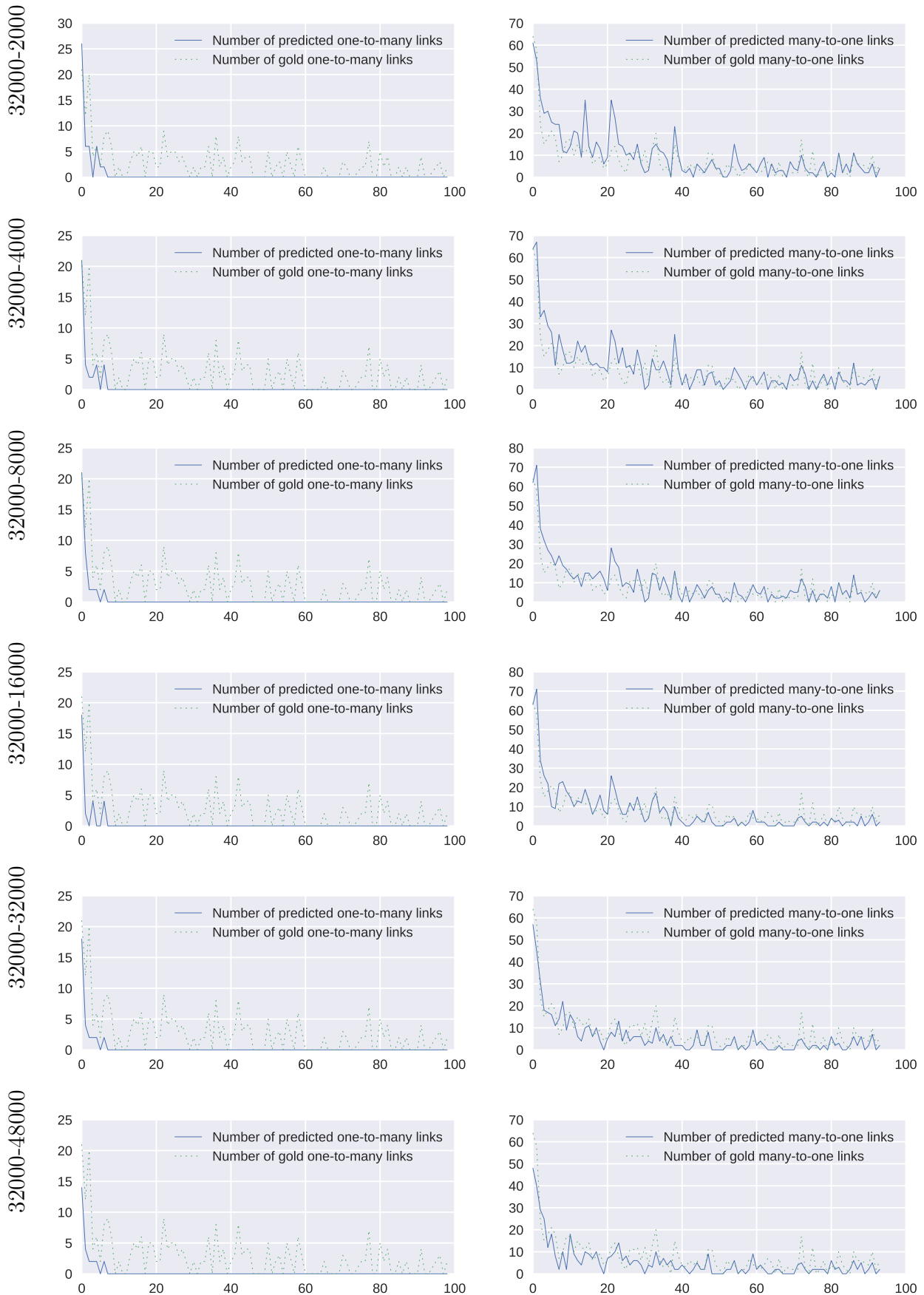




E.7. NUMBER OF ONE-TO-MANY/MANY-TO-ONE AS A FUNCTION OF WORD OCCURRENCE52







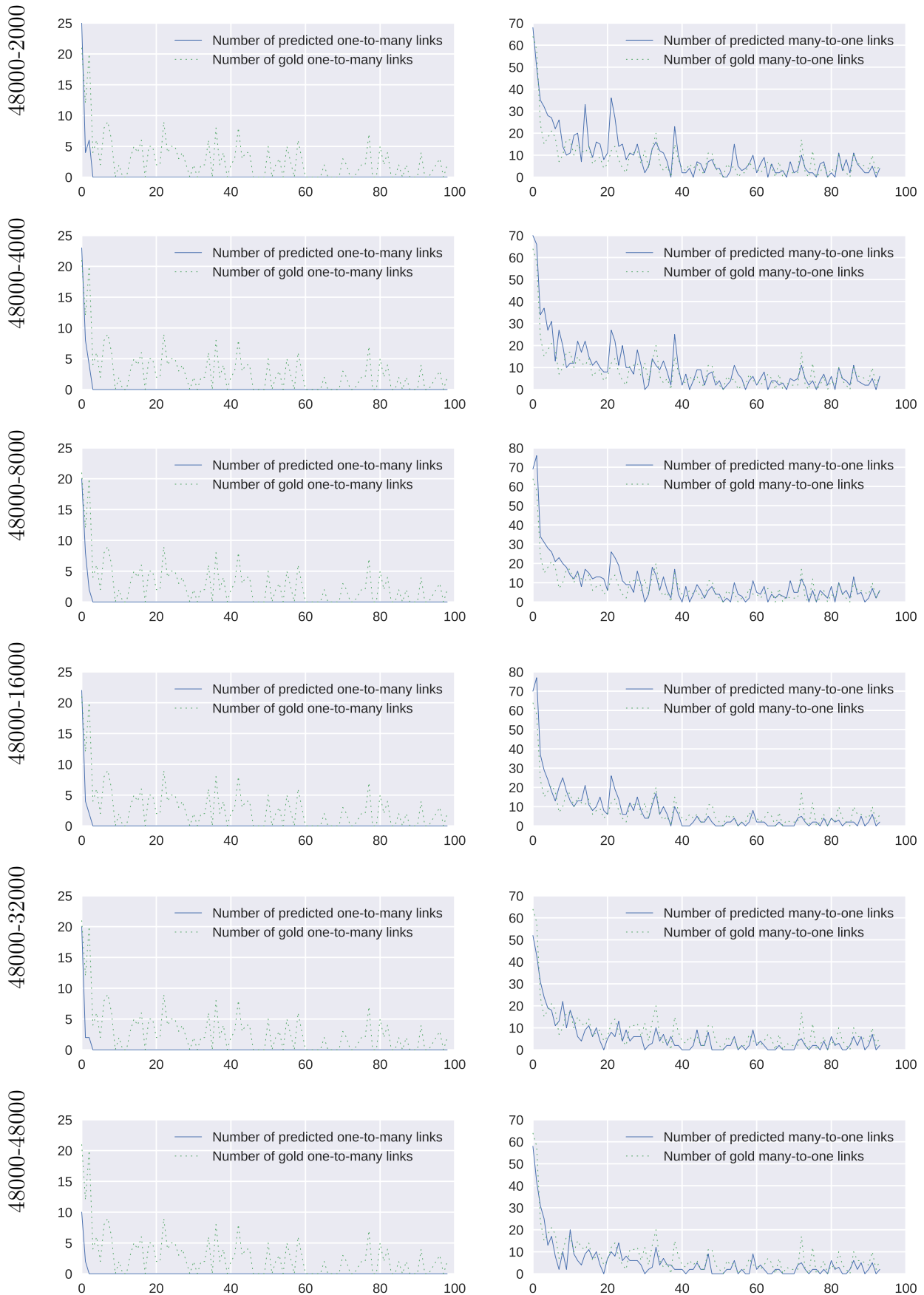
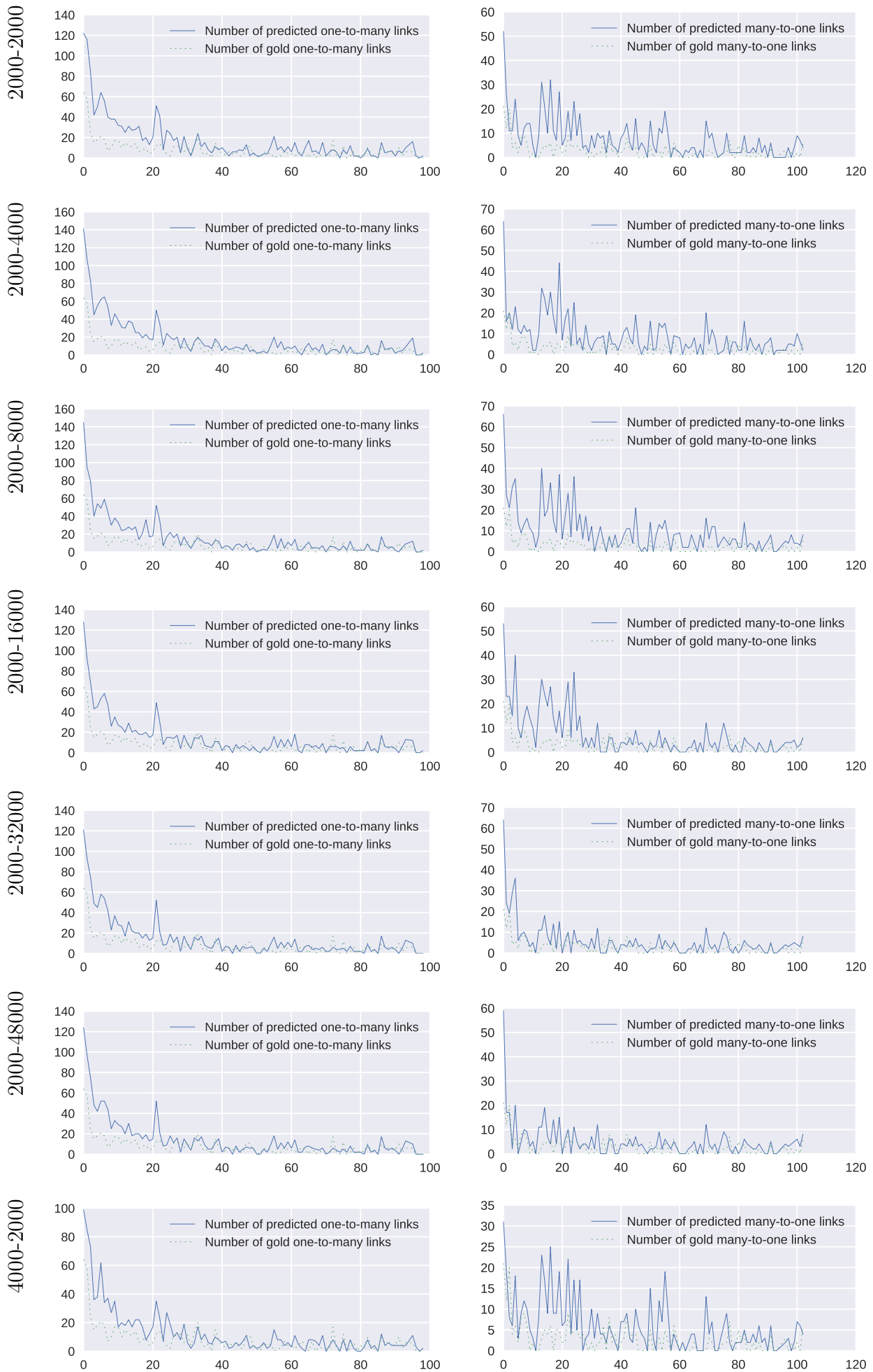
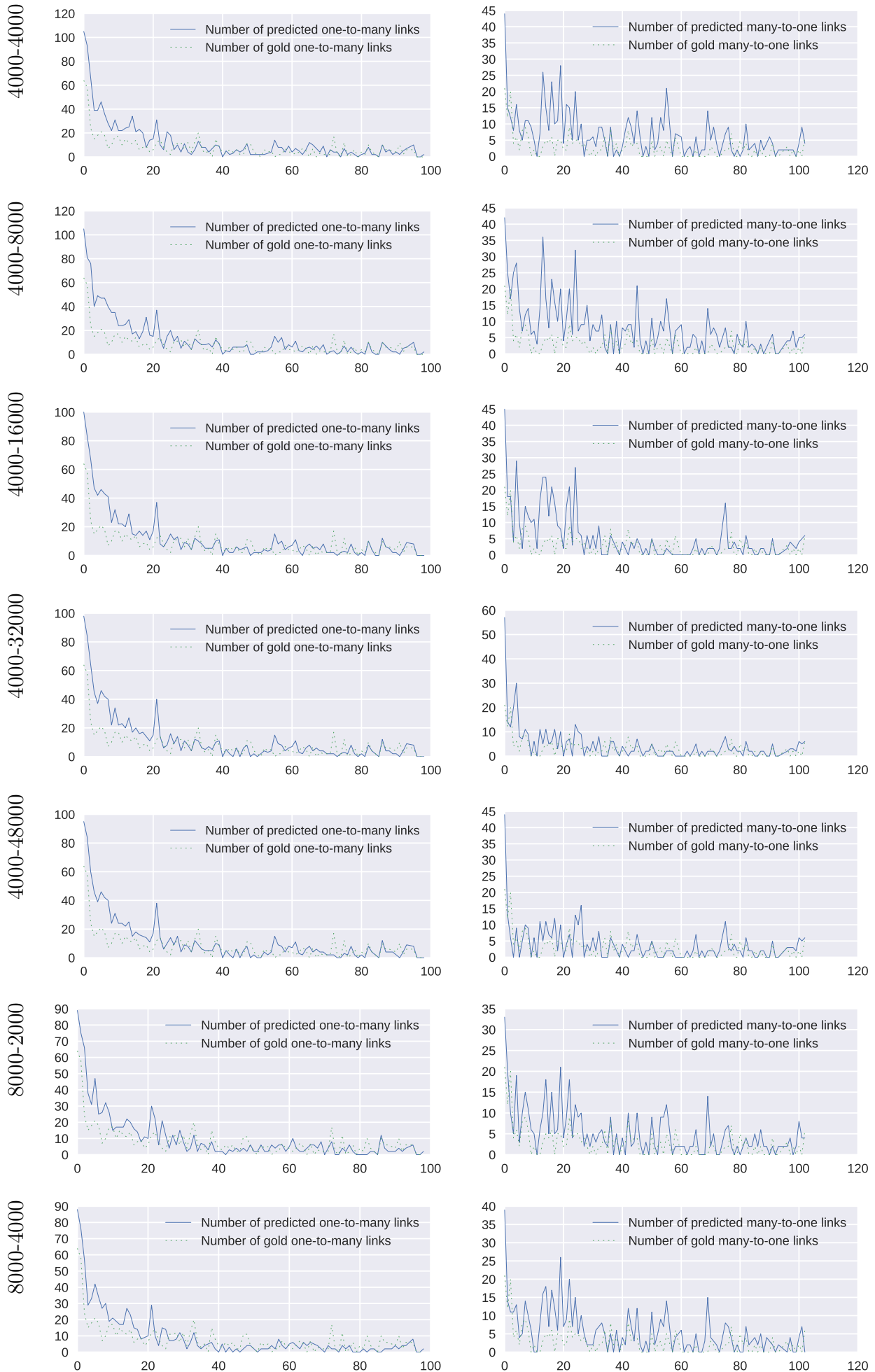


Figure E.91: The direction English-Romanian: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.

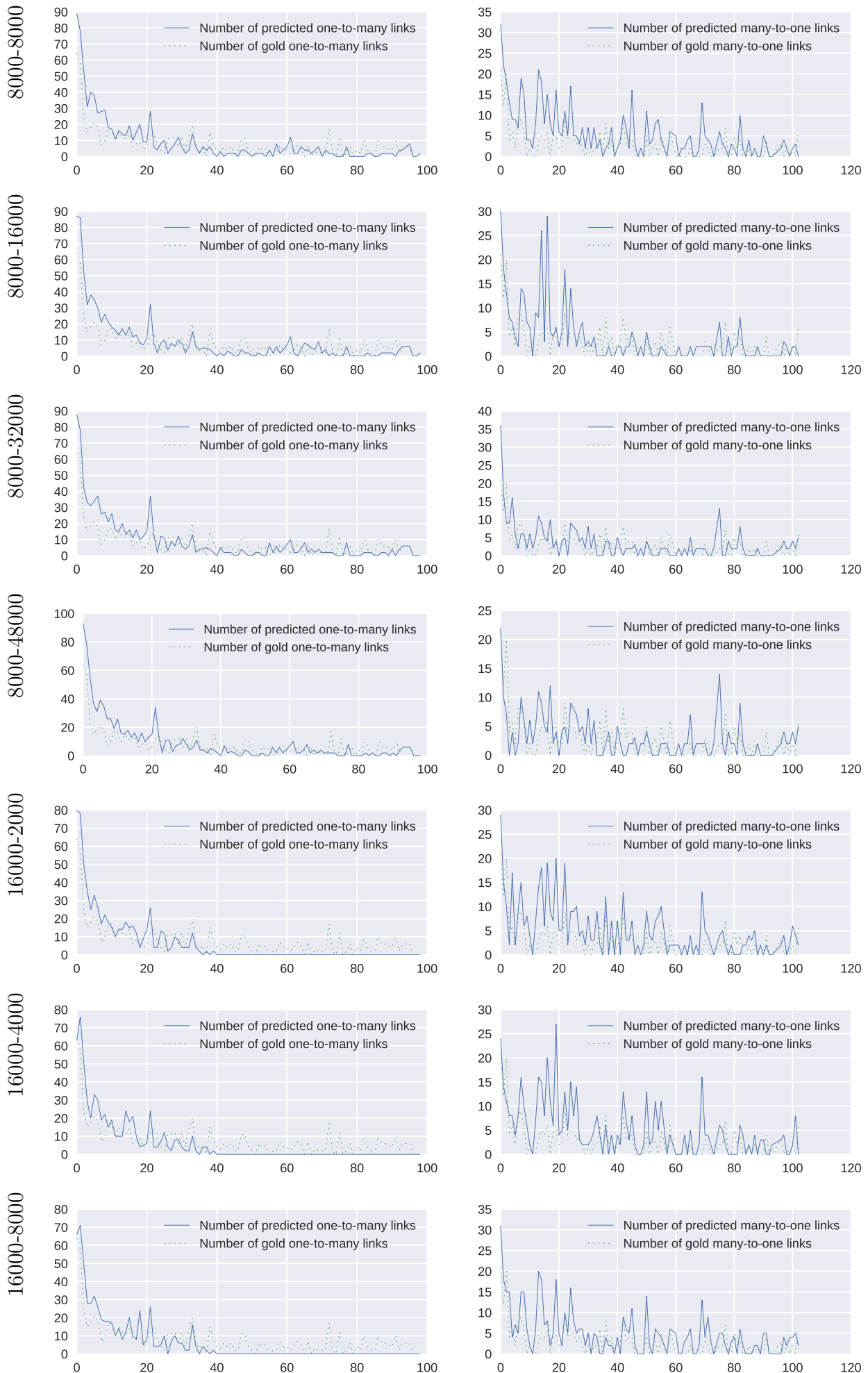
Source

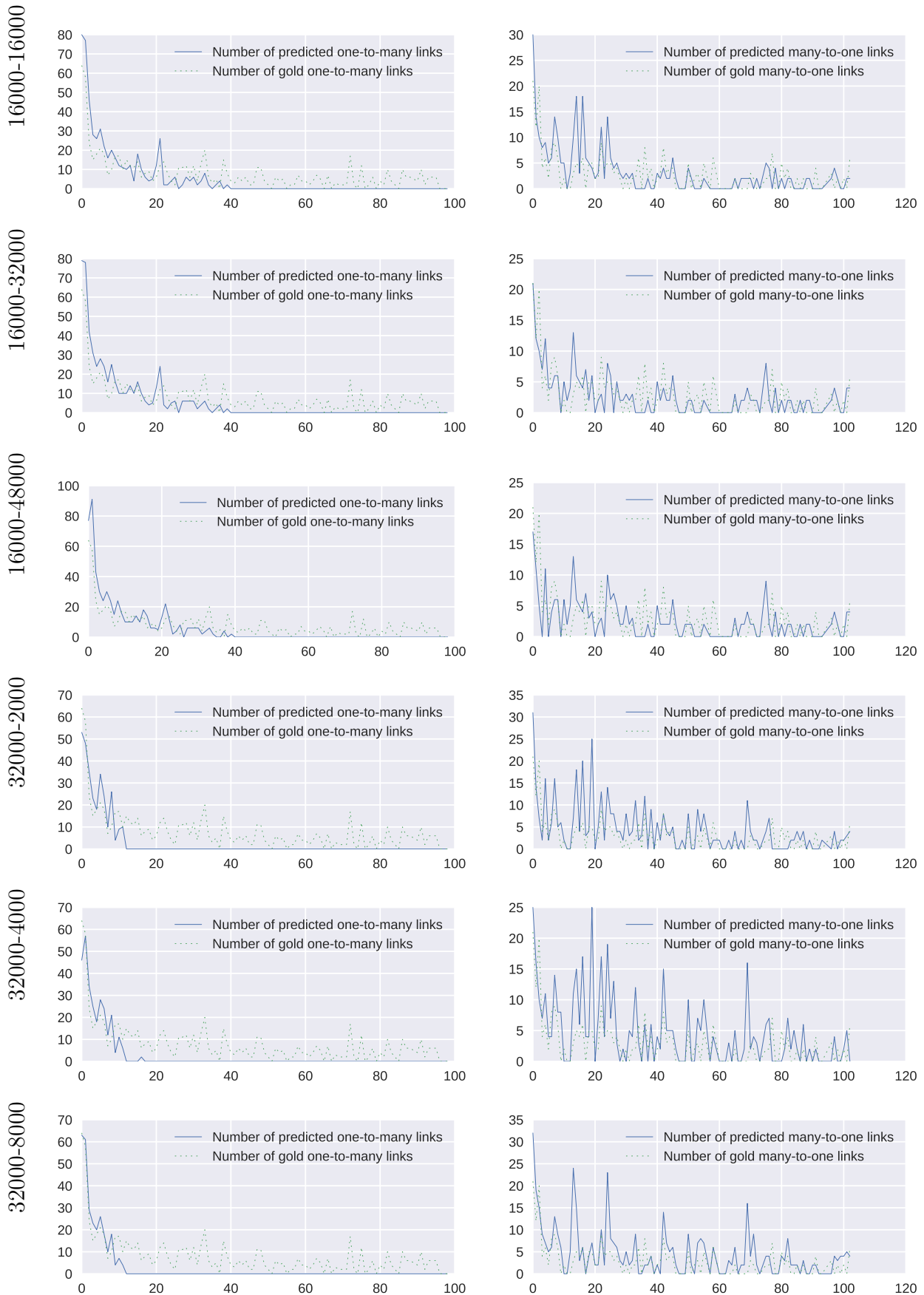
Target

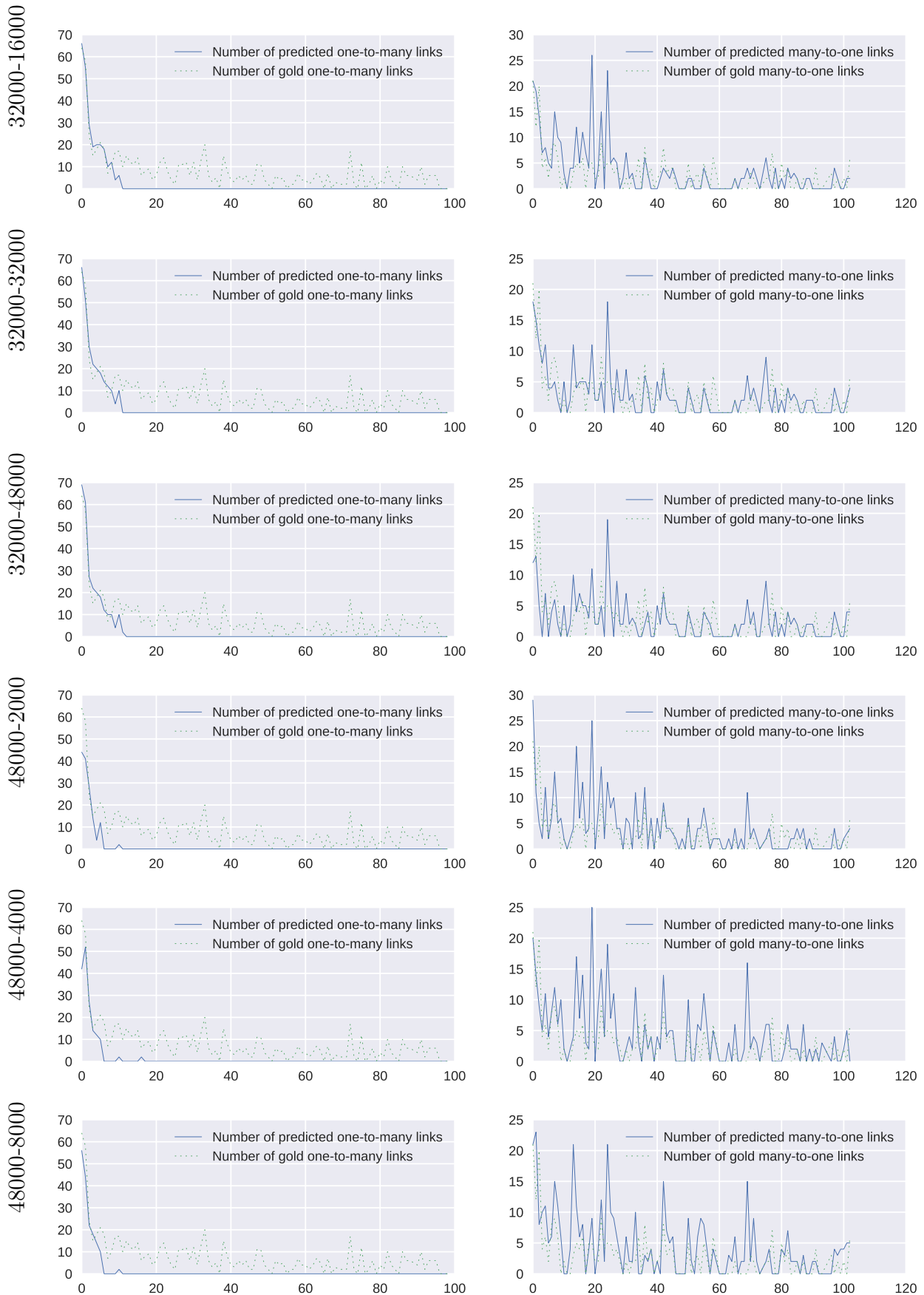




E.7. NUMBER OF ONE-TO-MANY/MANY-TO-ONE AS A FUNCTION OF WORD OCCURRENCE52







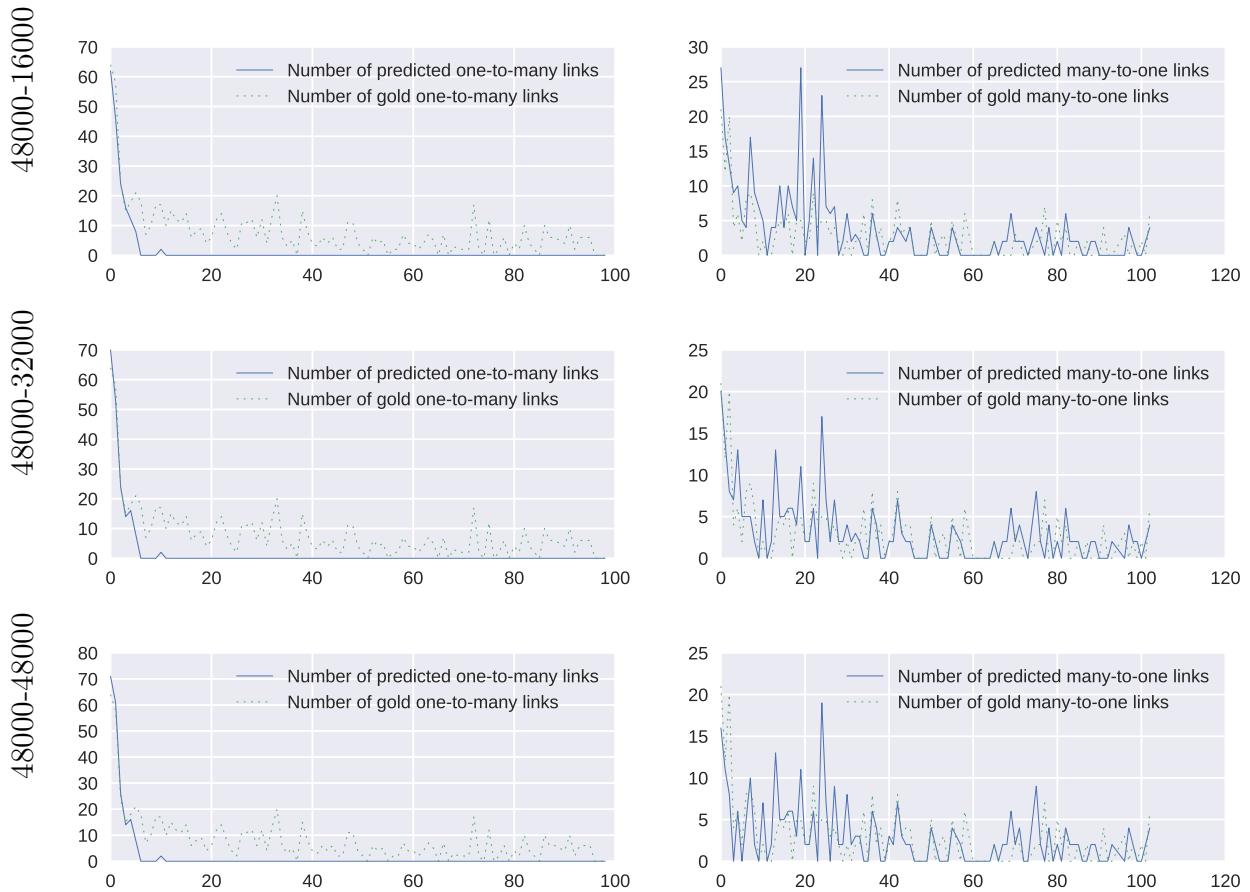
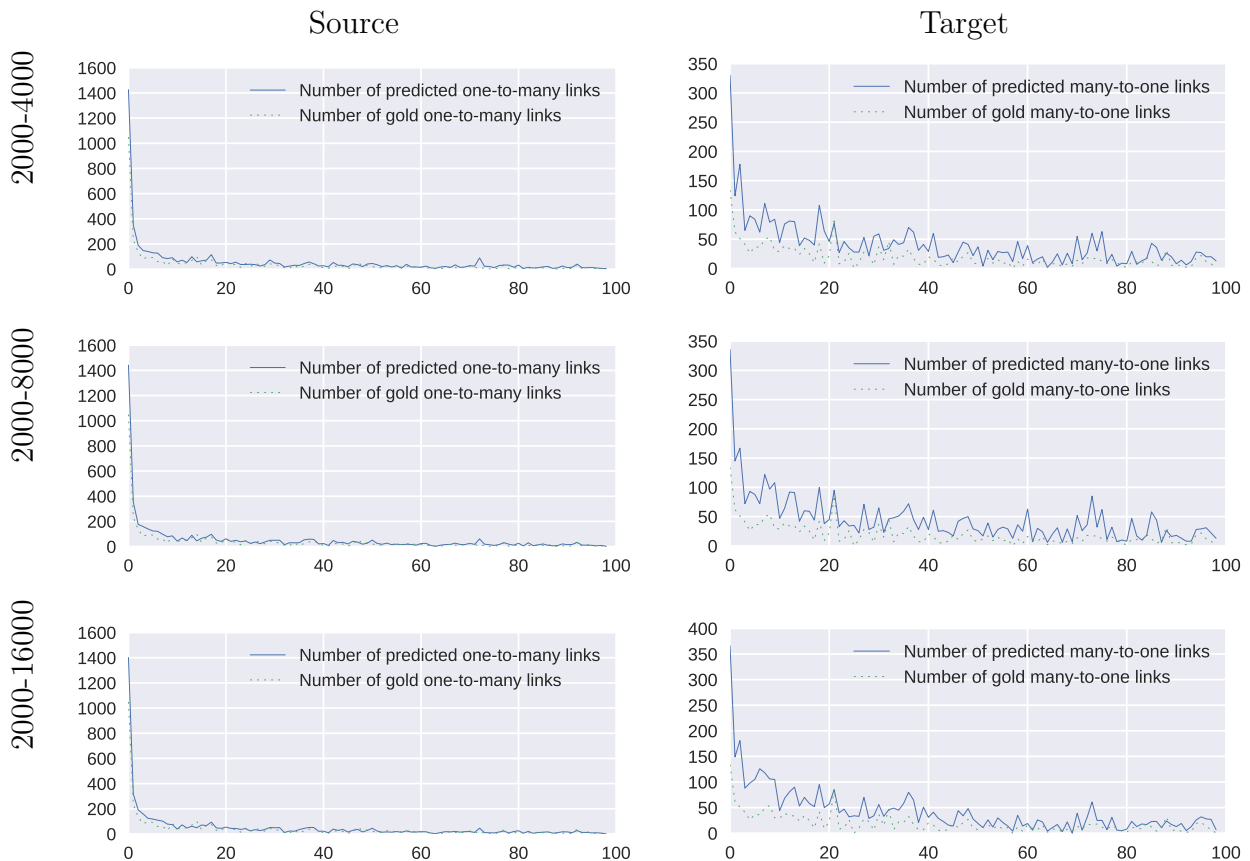
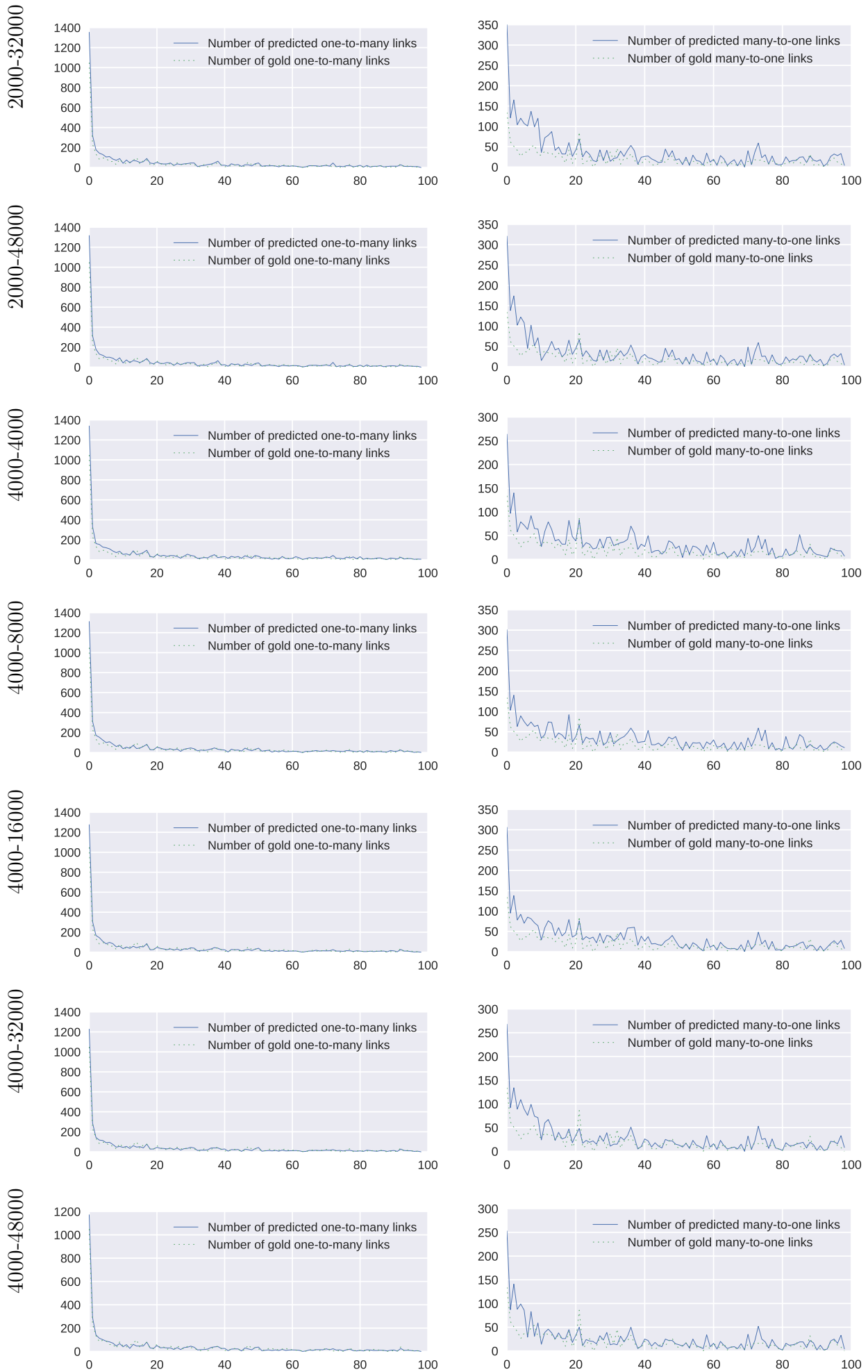
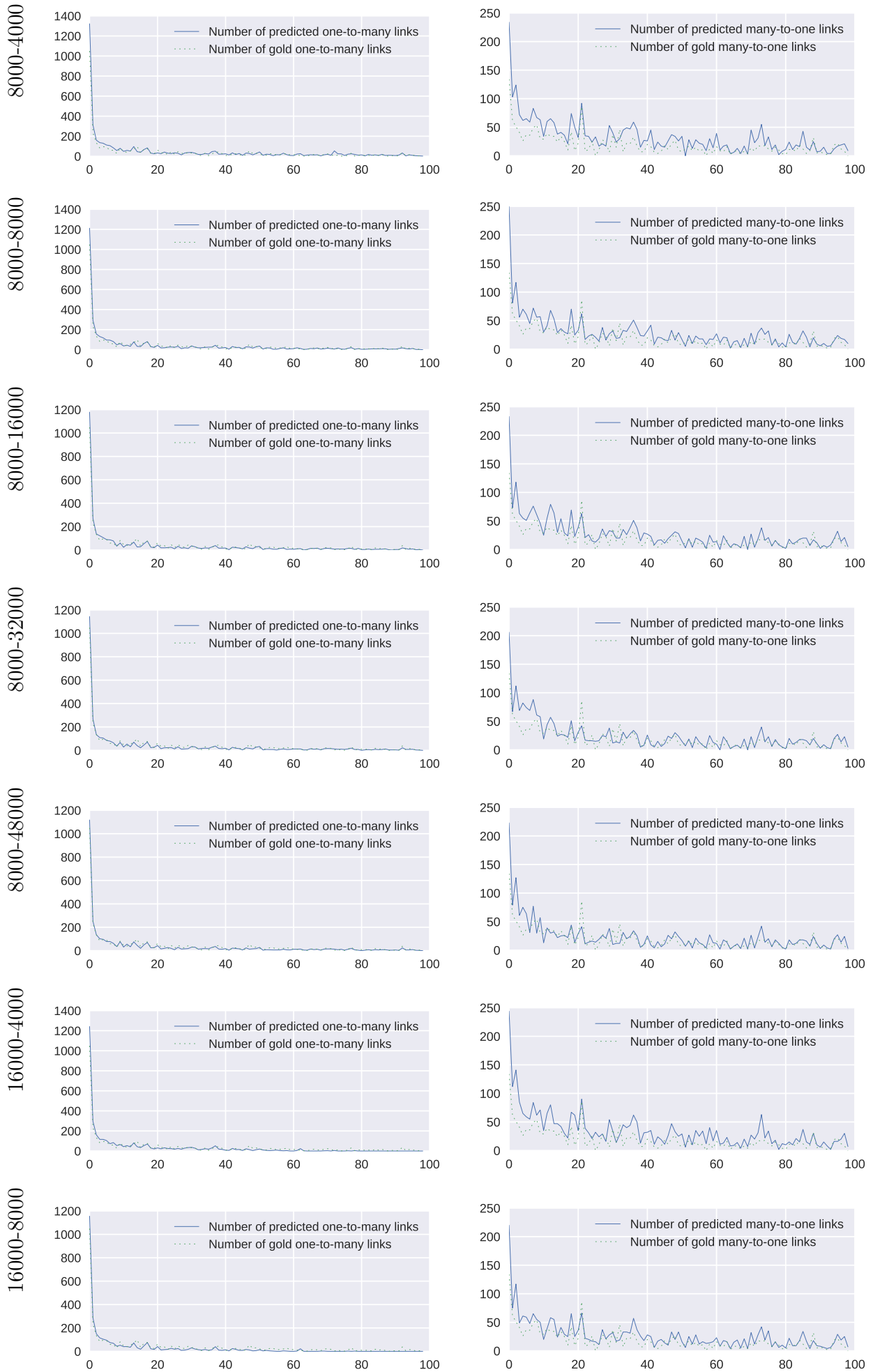


Figure E.92: The direction Romanian-English: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.



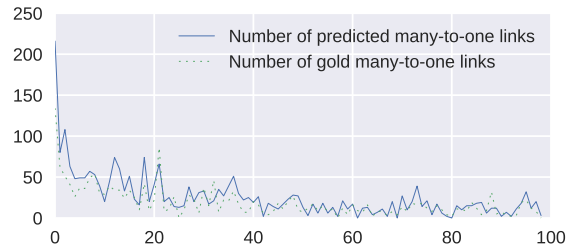
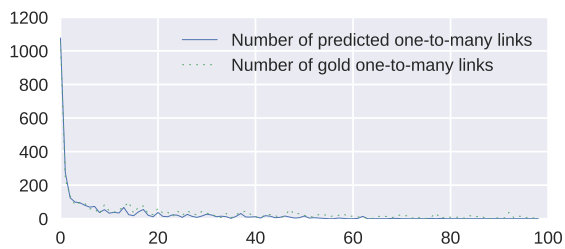
E.7. NUMBER OF ONE-TO-MANY/MANY-TO-ONE AS A FUNCTION OF WORD OCCURRENCE



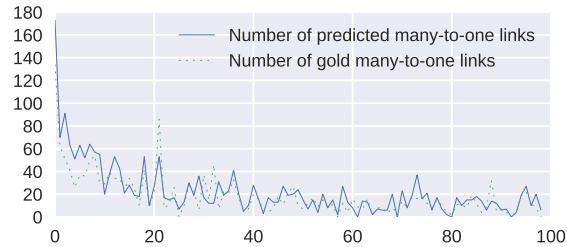
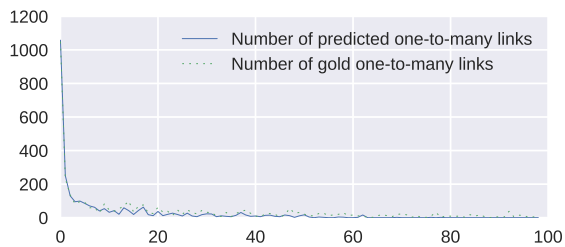


E.7. NUMBER OF ONE-TO-MANY/MANY-TO-ONE AS A FUNCTION OF WORD OCCURRENCE53

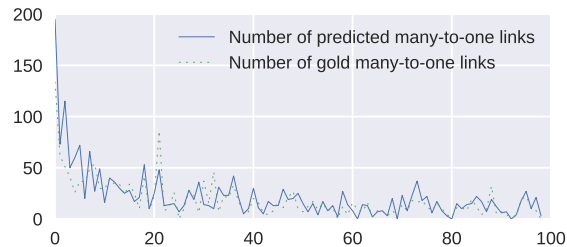
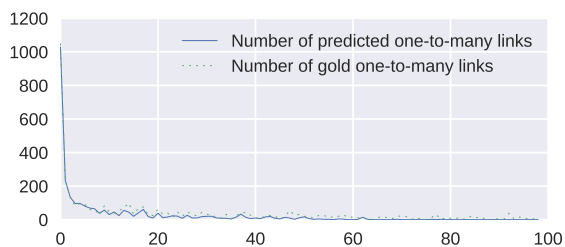
16000-16000



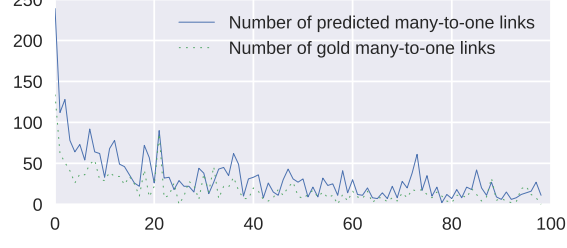
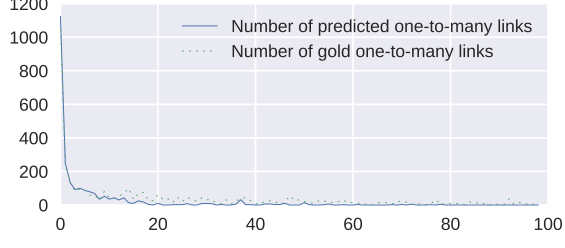
16000-32000



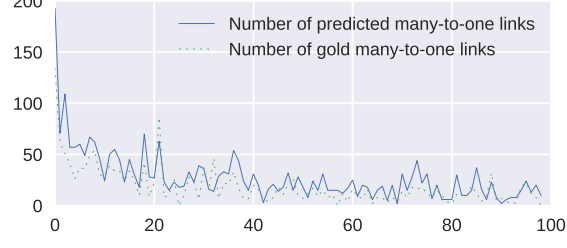
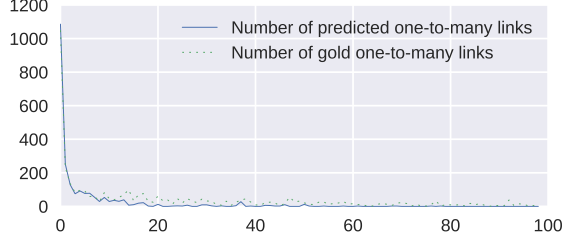
16000-48000



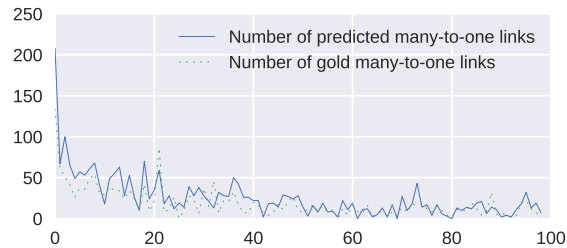
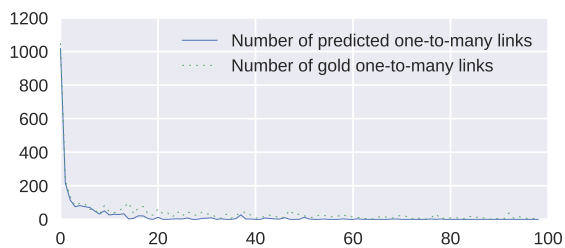
32000-4000

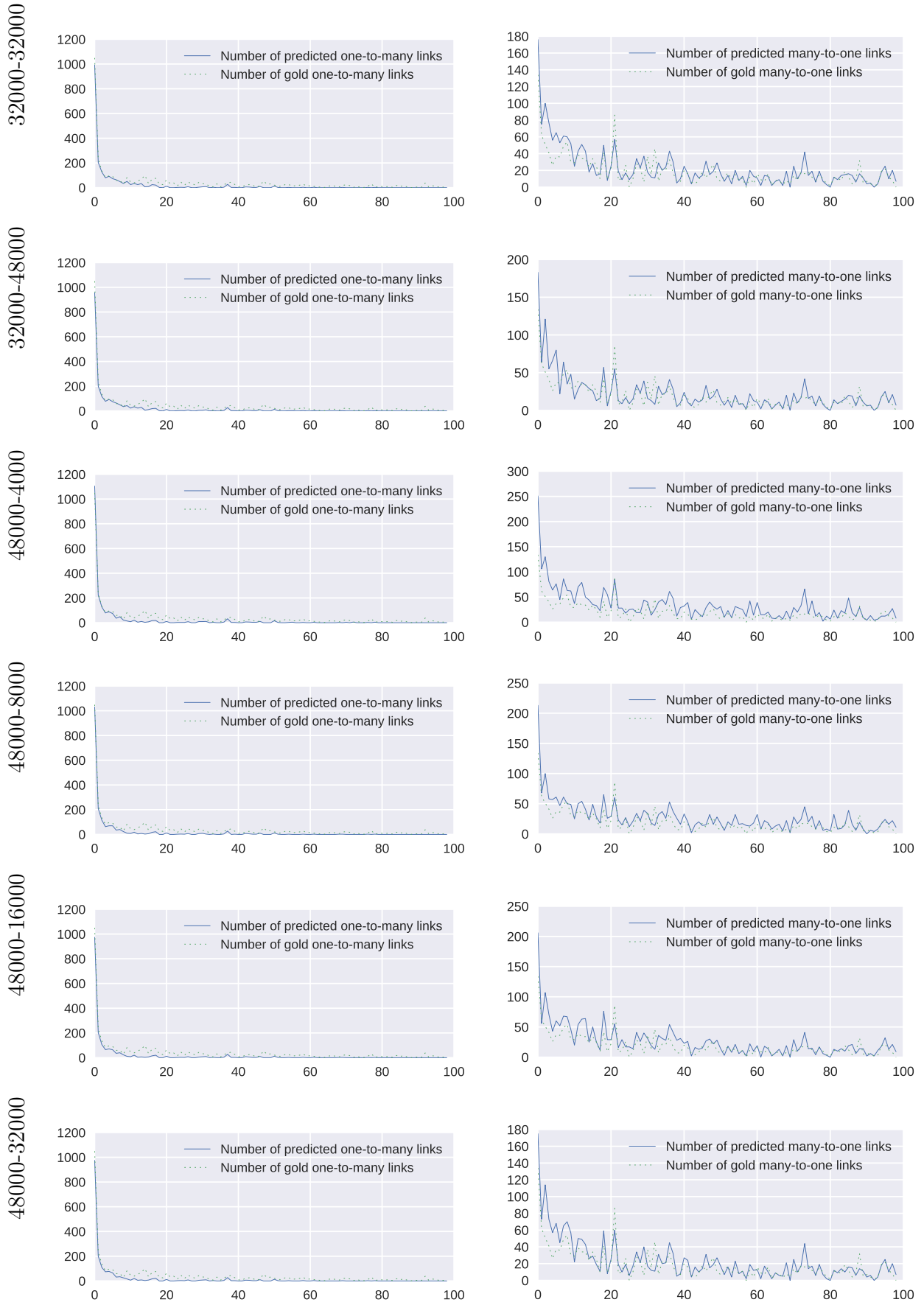


32000-8000



32000-16000





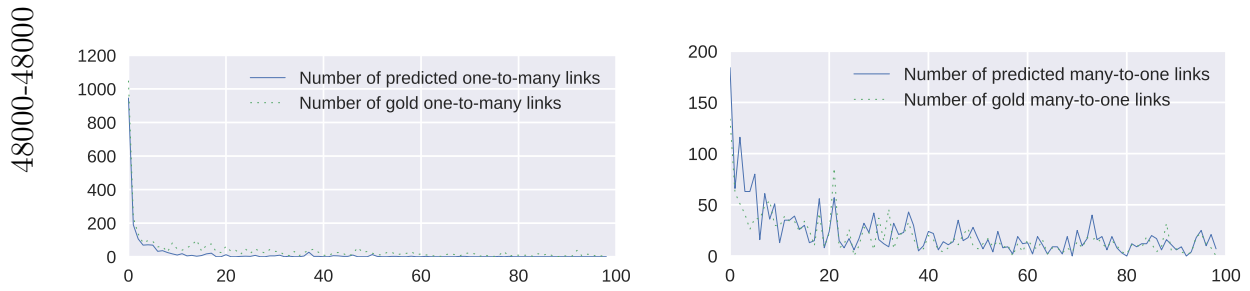
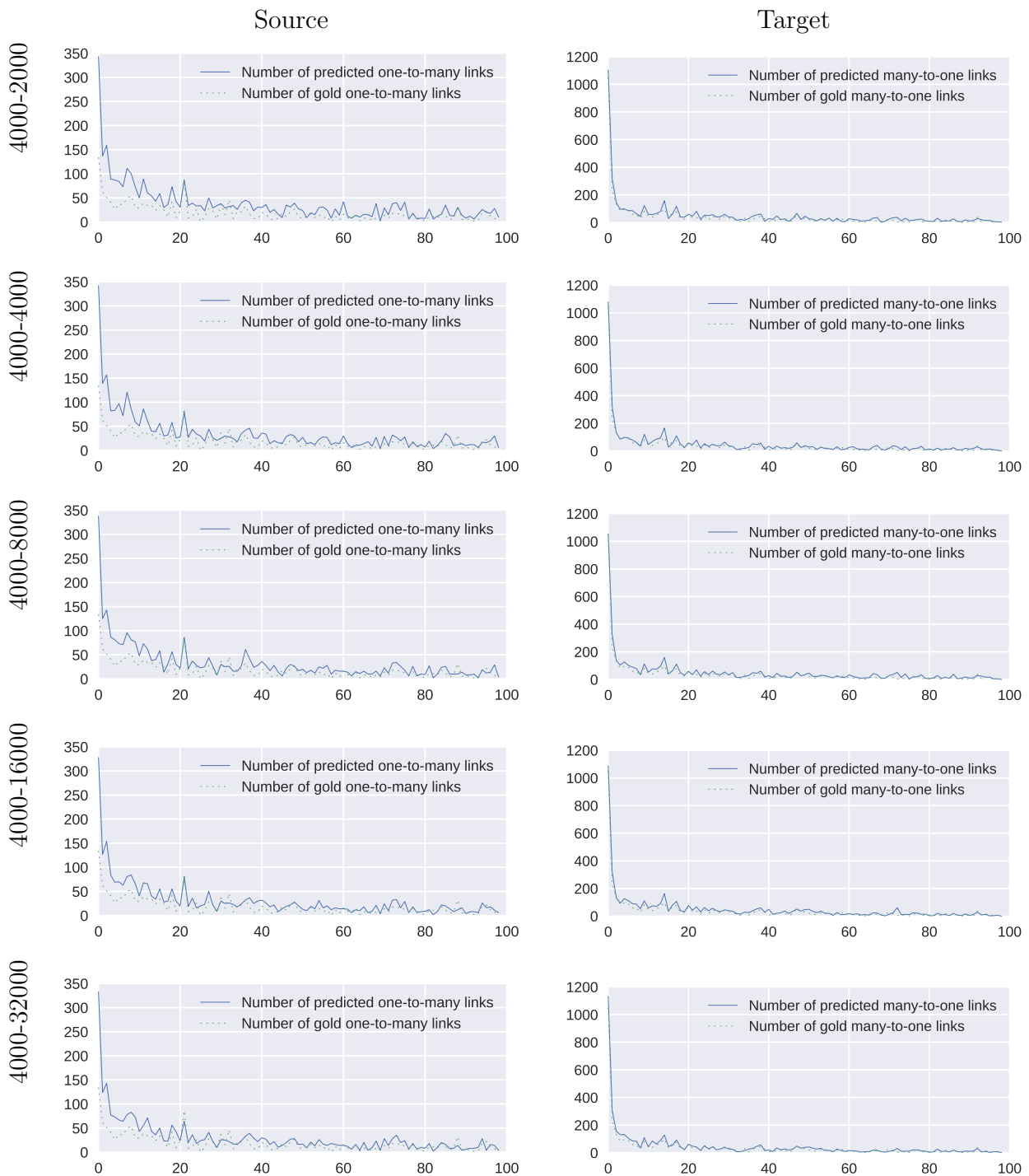
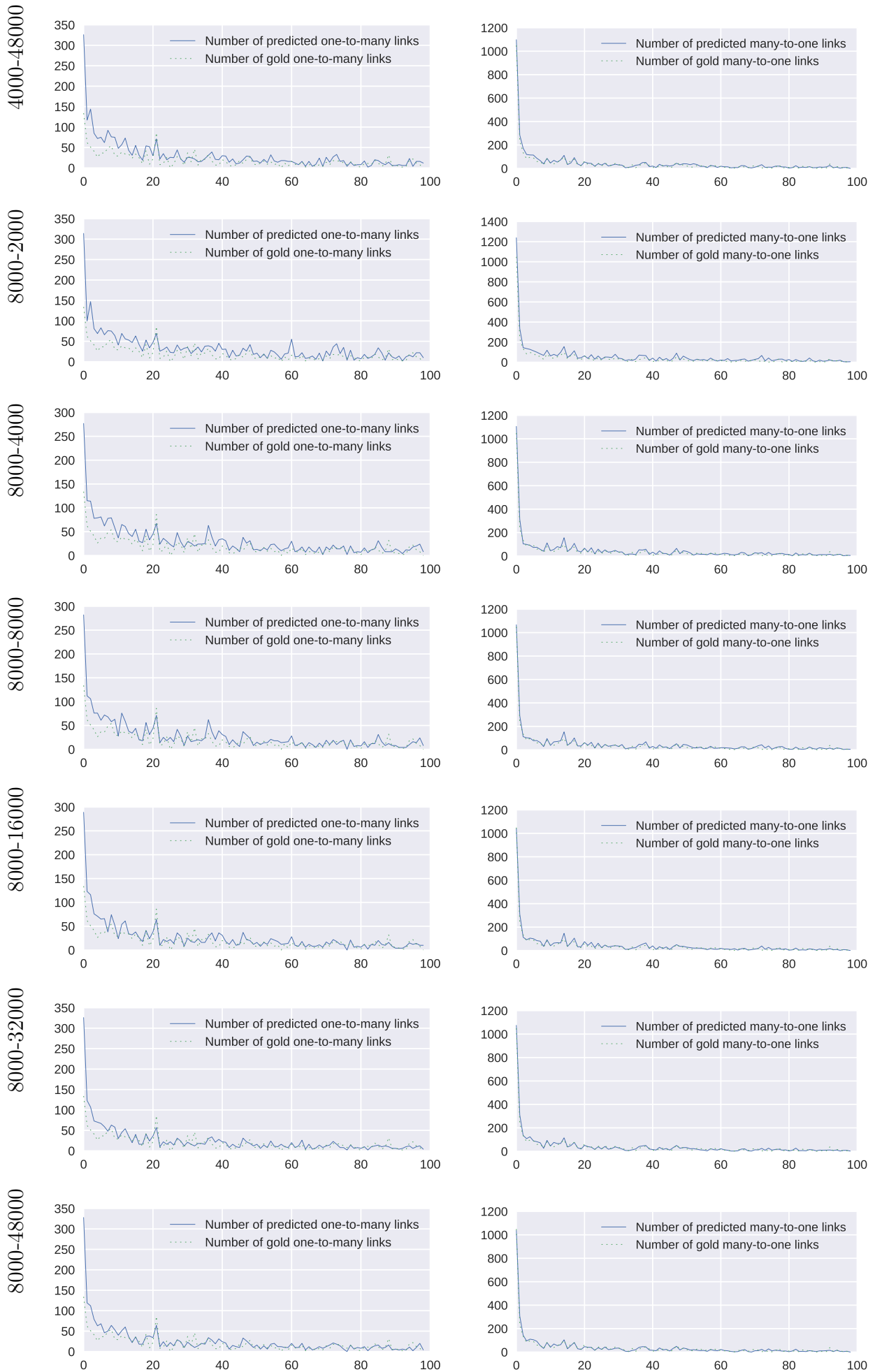
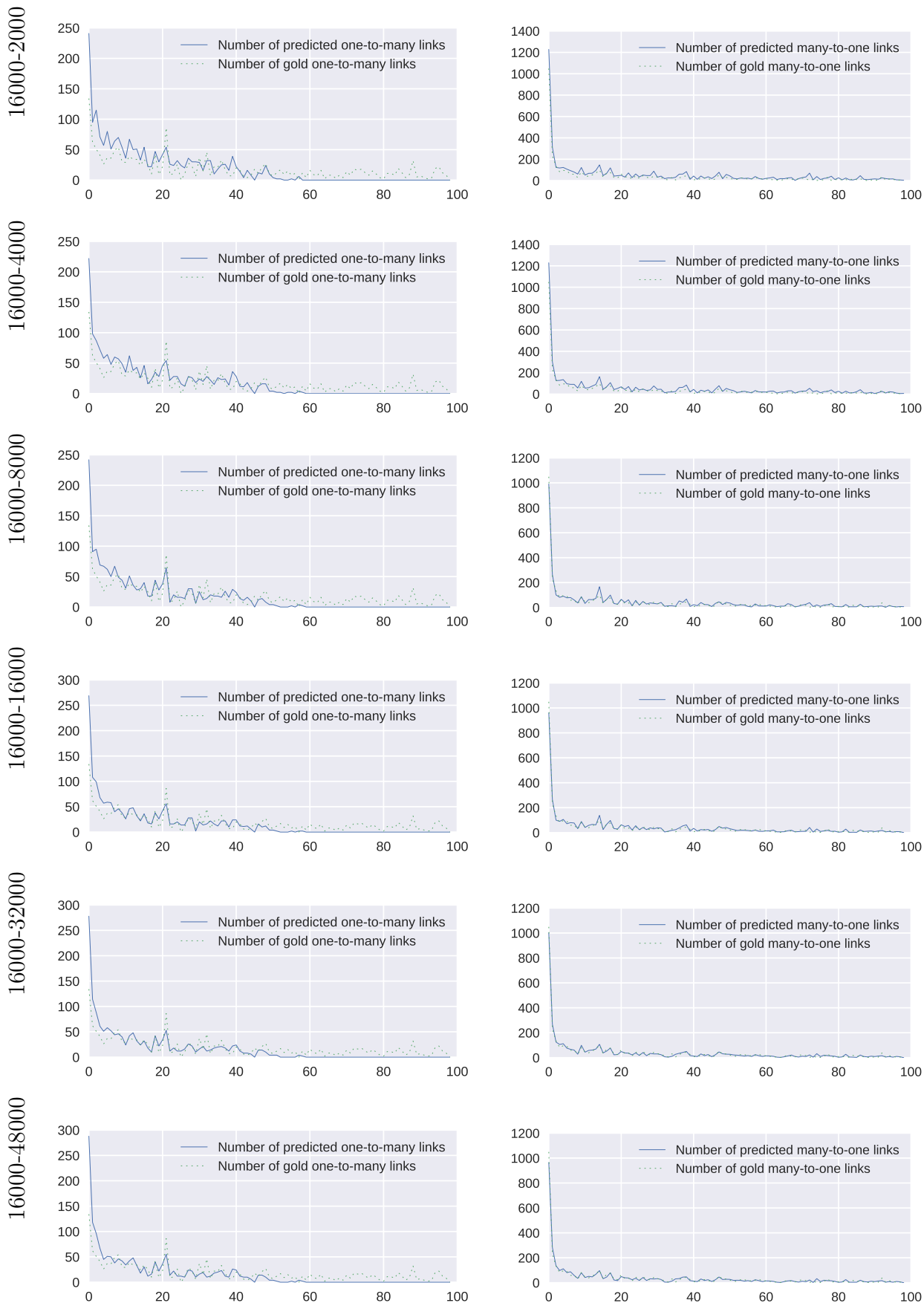


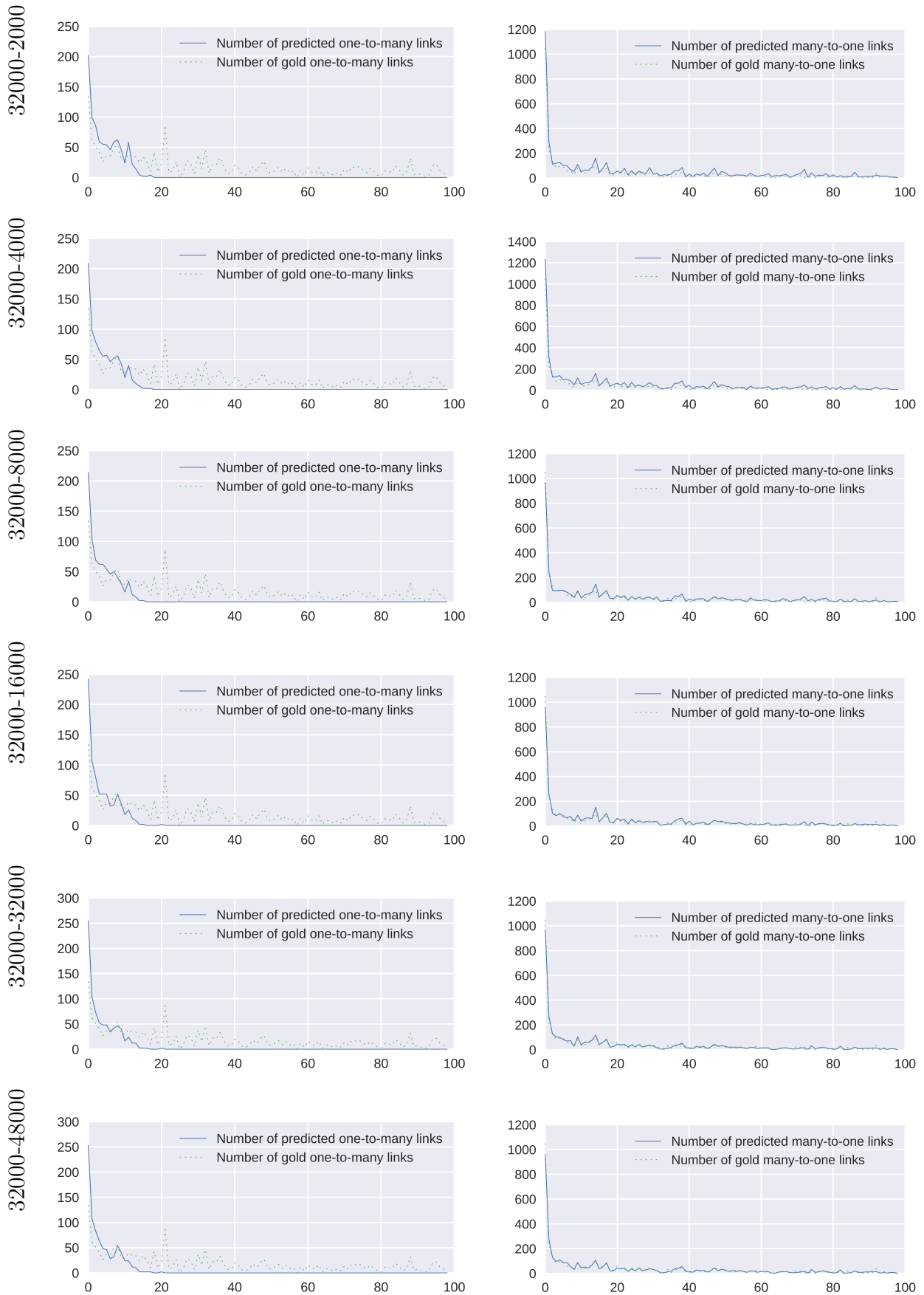
Figure E.93: The direction English-Japanese: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.





E.7. NUMBER OF ONE-TO-MANY/MANY-TO-ONE AS A FUNCTION OF WORD OCCURRENCE





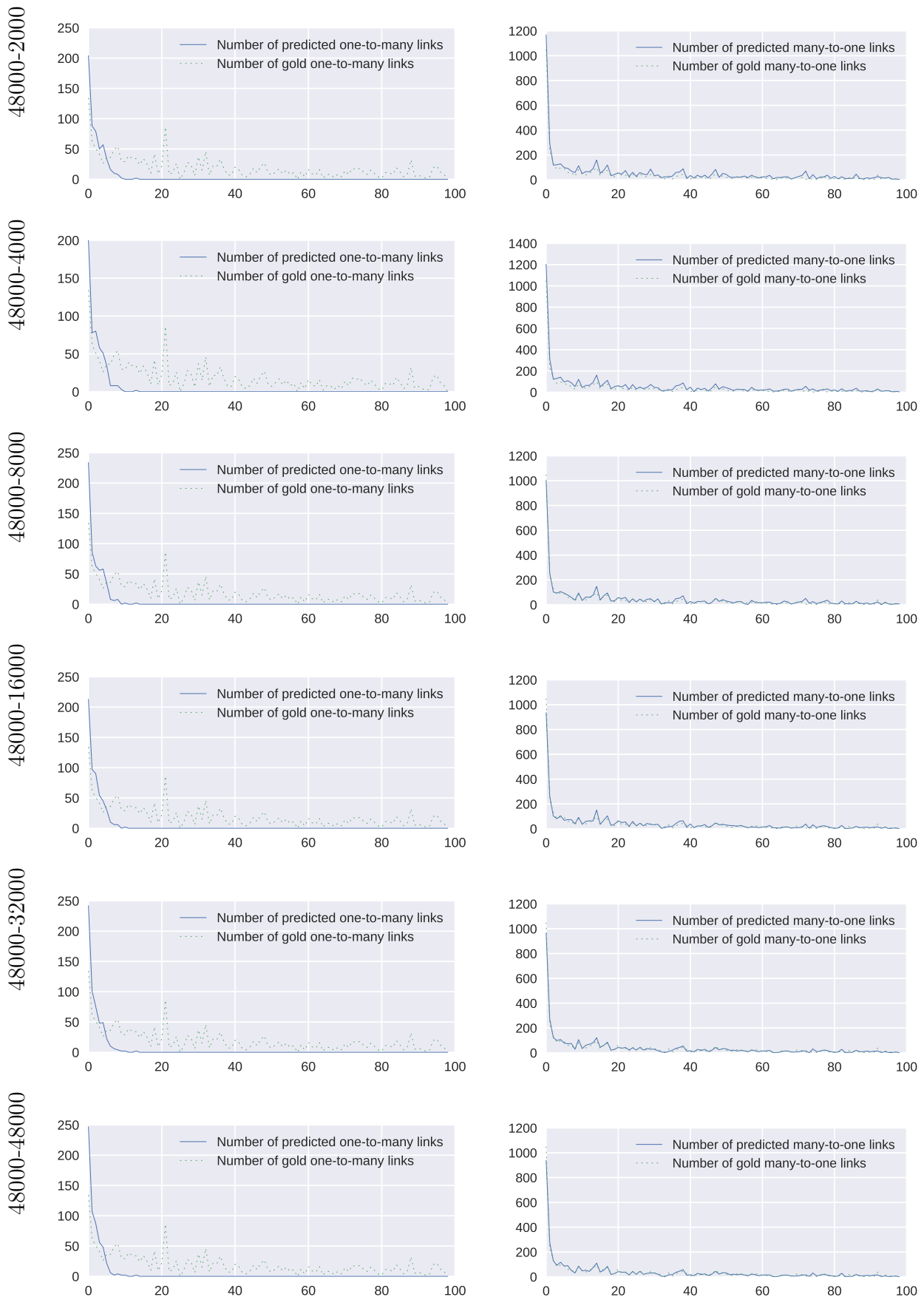
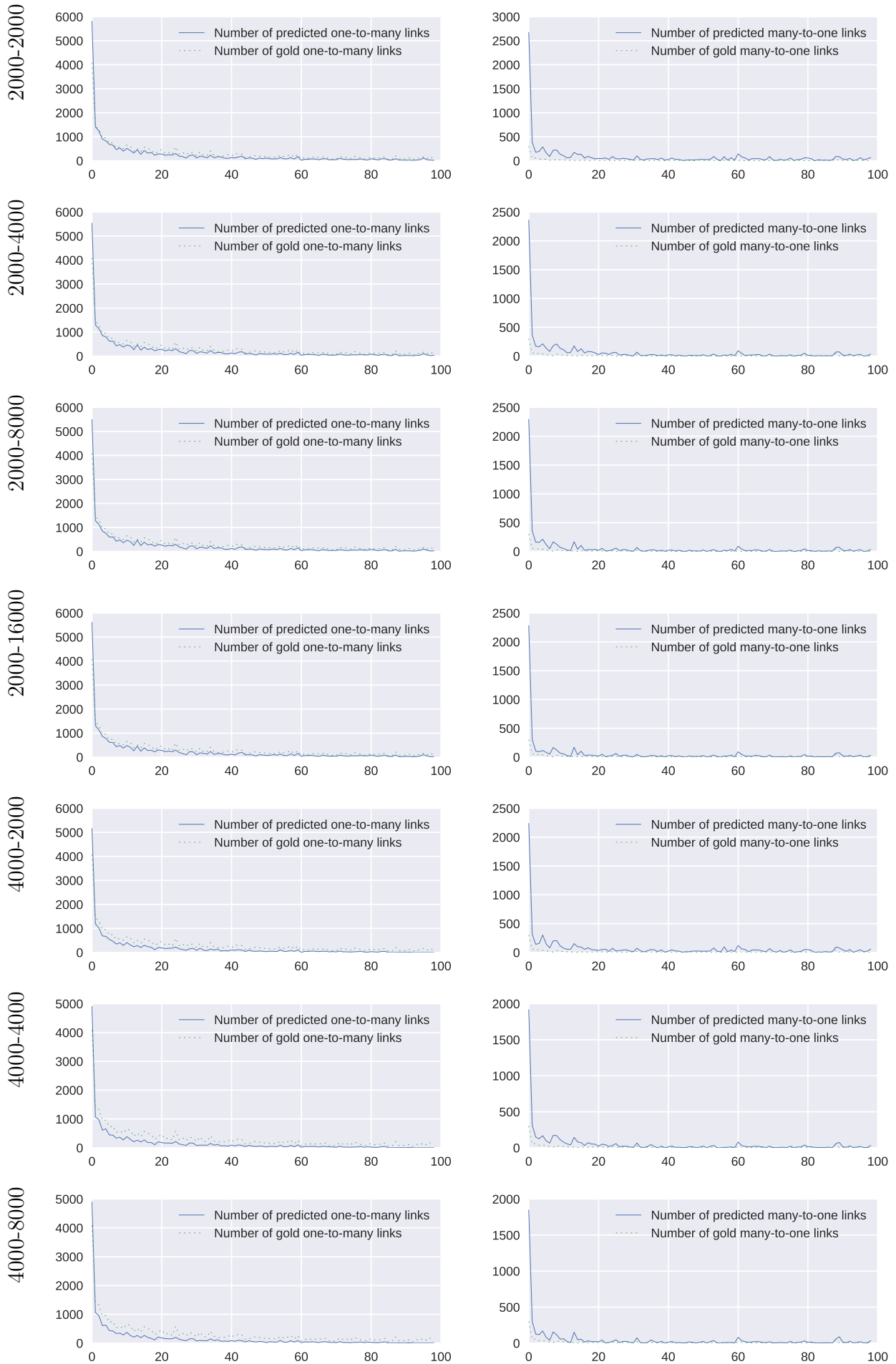


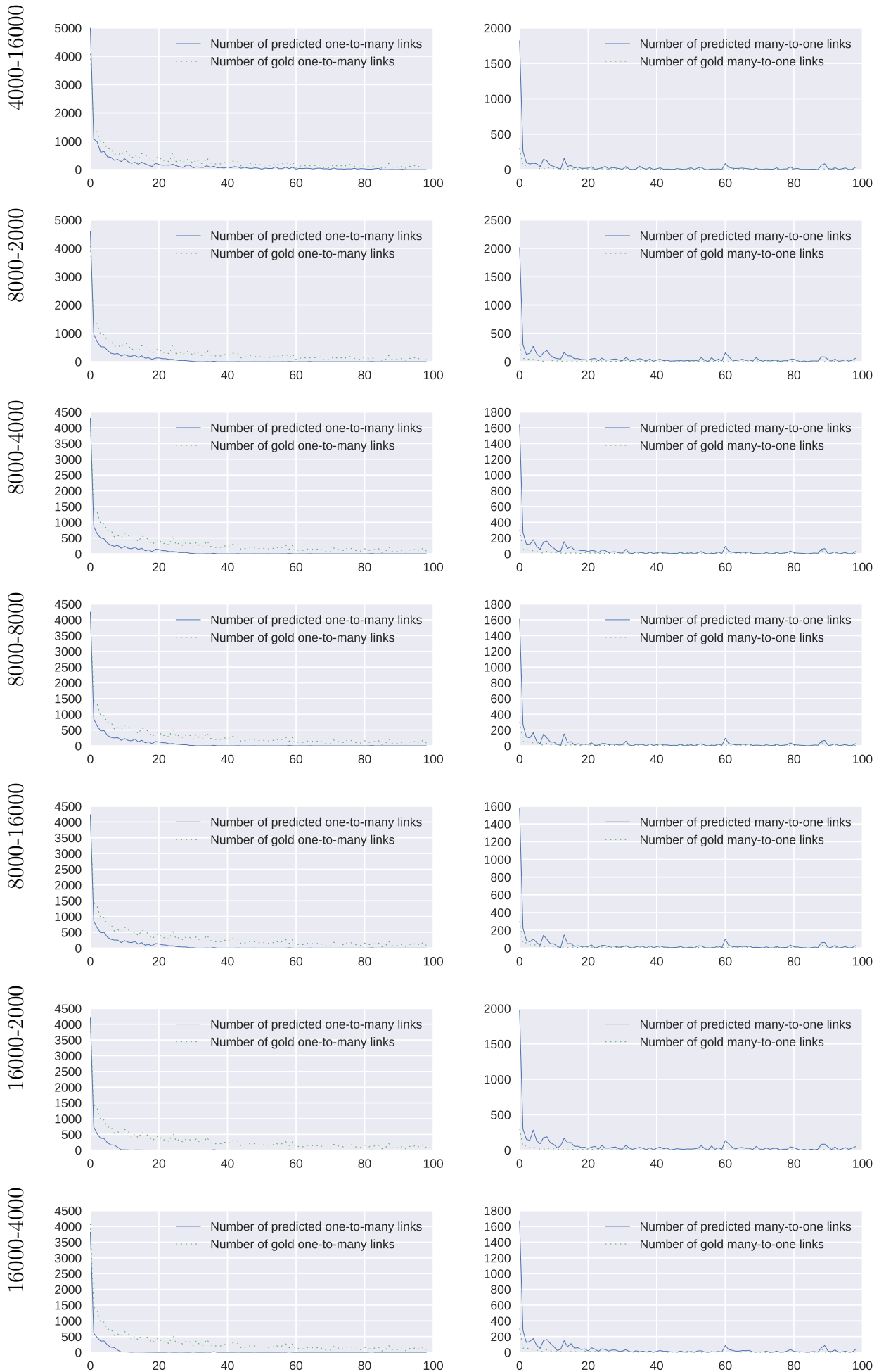
Figure E.94: The direction Japanese-English: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.

Source

Target



E.7. NUMBER OF ONE-TO-MANY/MANY-TO-ONE AS A FUNCTION OF WORD OCCURRENCE⁵⁴



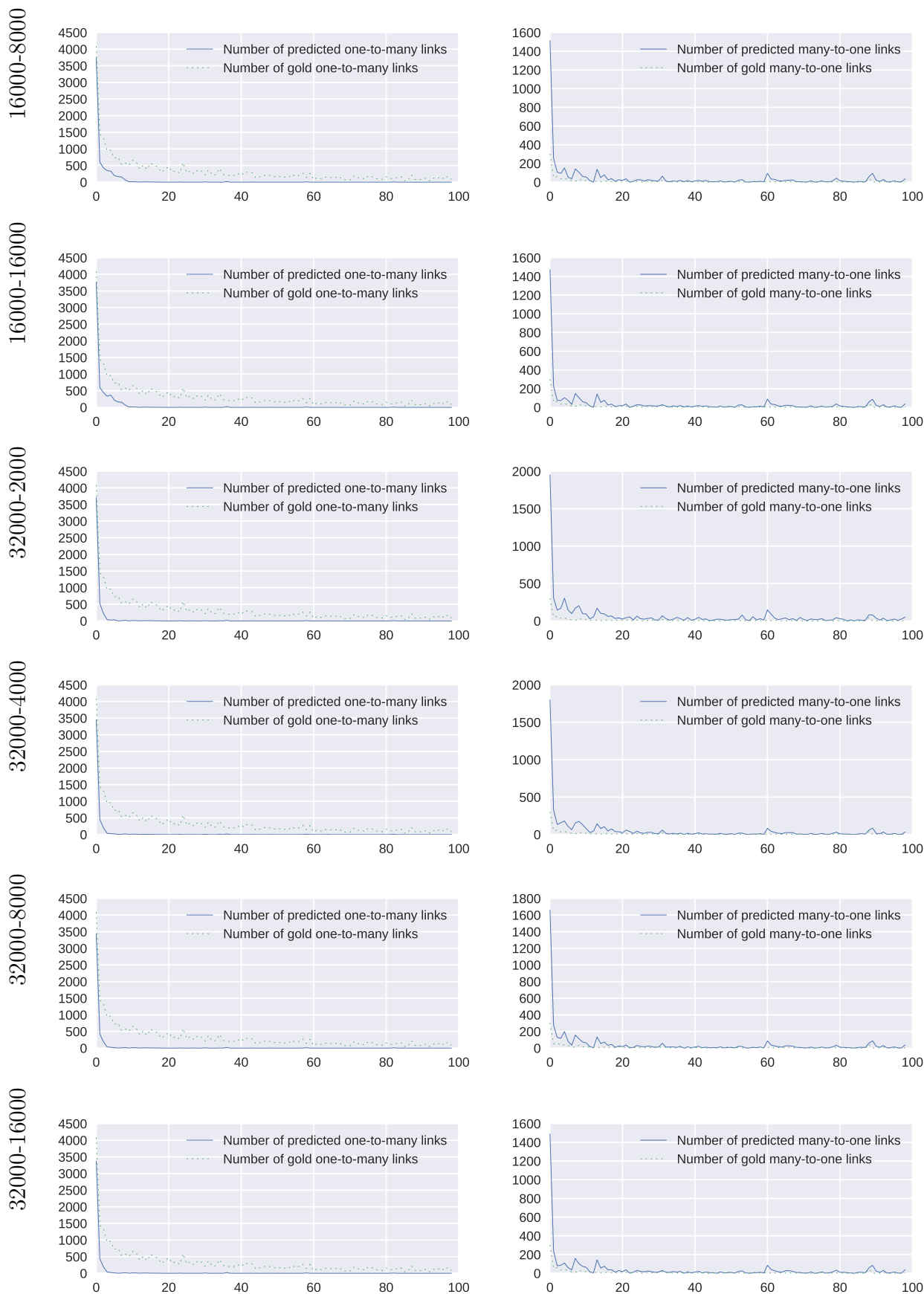
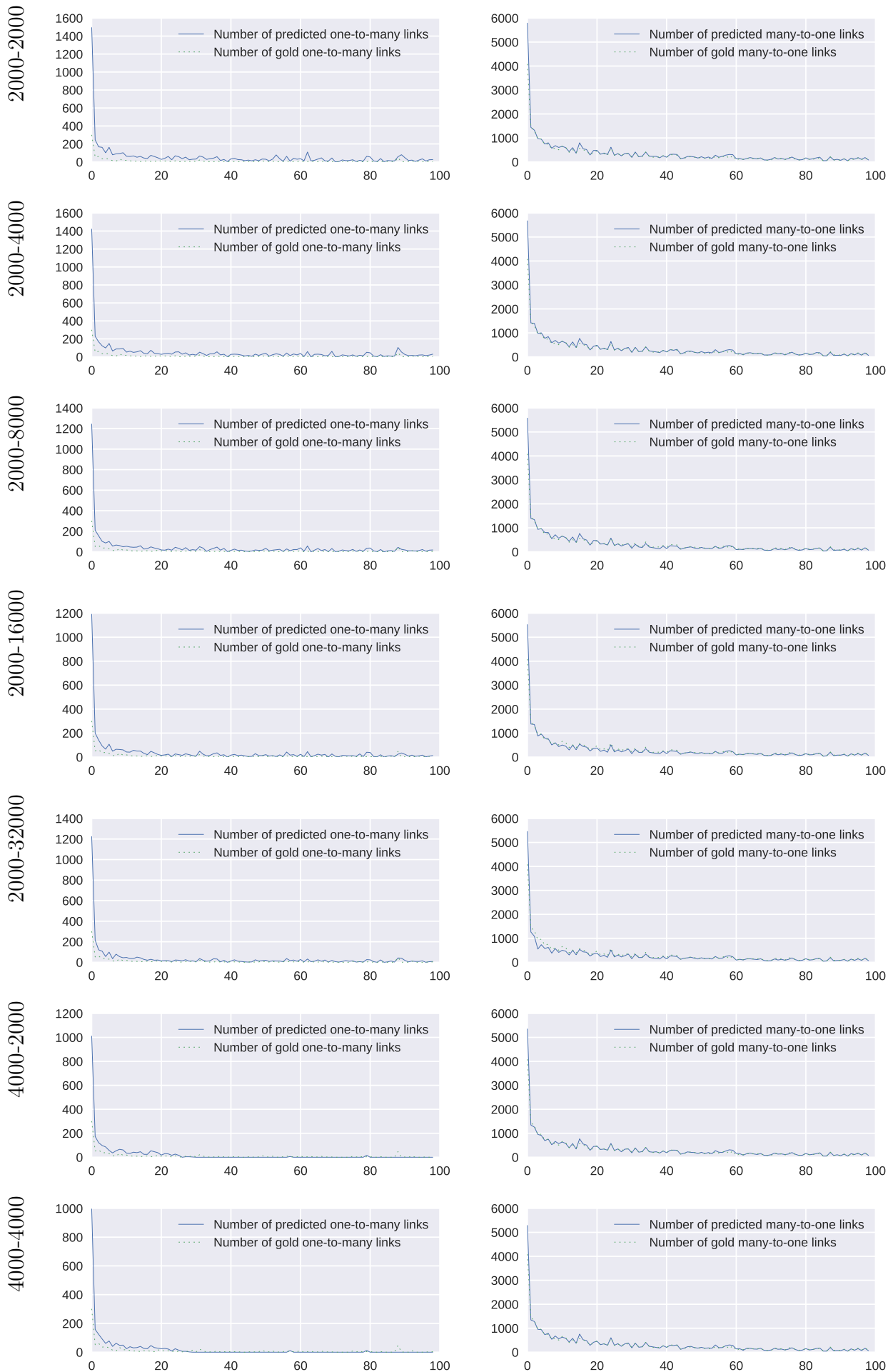


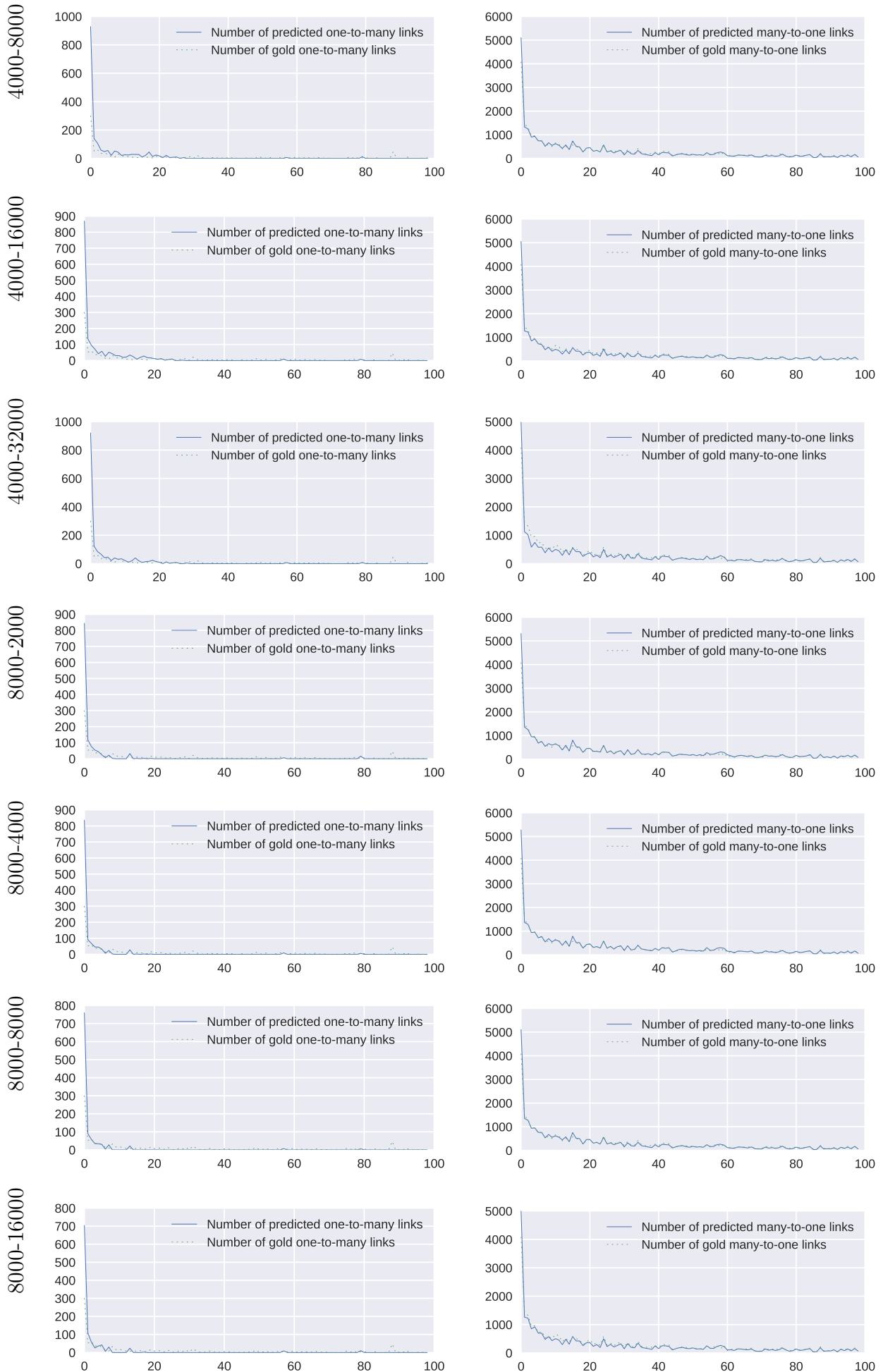
Figure E.95: The direction English-Vietnamese: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.

Source

Target

E.7. NUMBER OF ONE-TO-MANY/MANY-TO-ONE AS A FUNCTION OF WORD OCCURRENCE⁵⁴





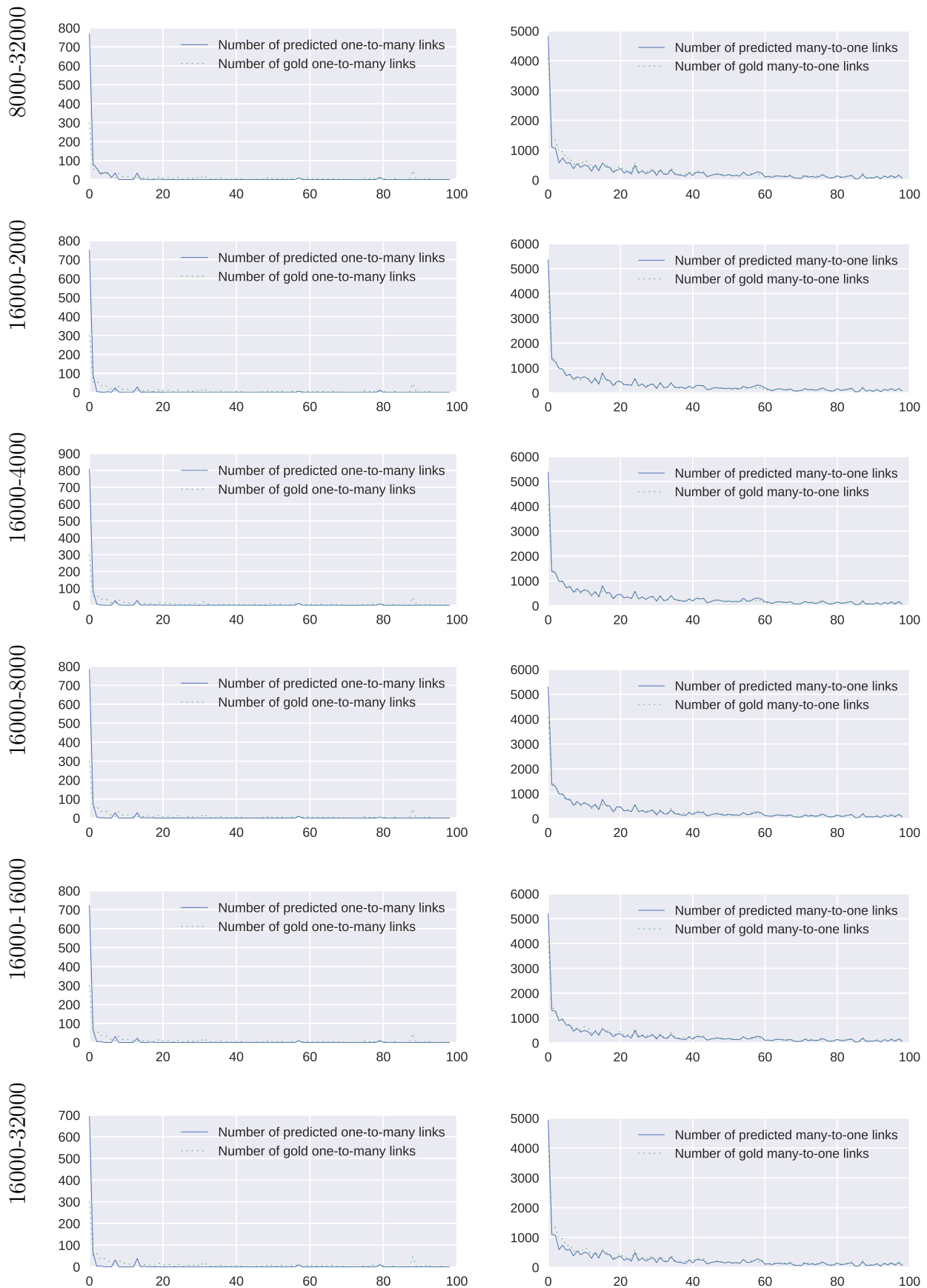


Figure E.96: The direction Vietnamese-English: Number of one-to-many (left graphs) and many-to-one (right graphs) links as a function of word occurrence.

E.8 Part-of-Speech

E.8.1 Alignment links for content and function words



Figure E.97: BPE-based `Fastalign` for English-French: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).

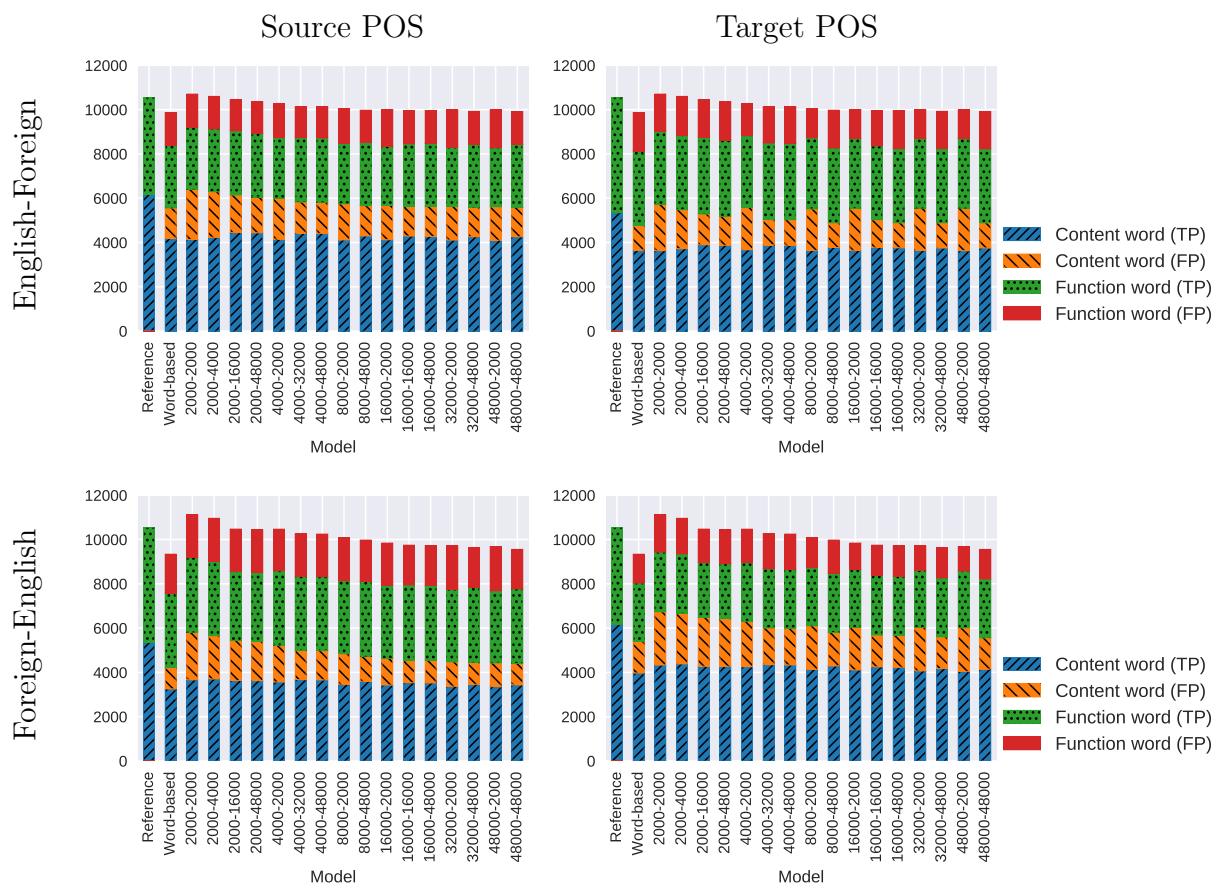


Figure E.98: BPE-based Fastalign for English-German: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).

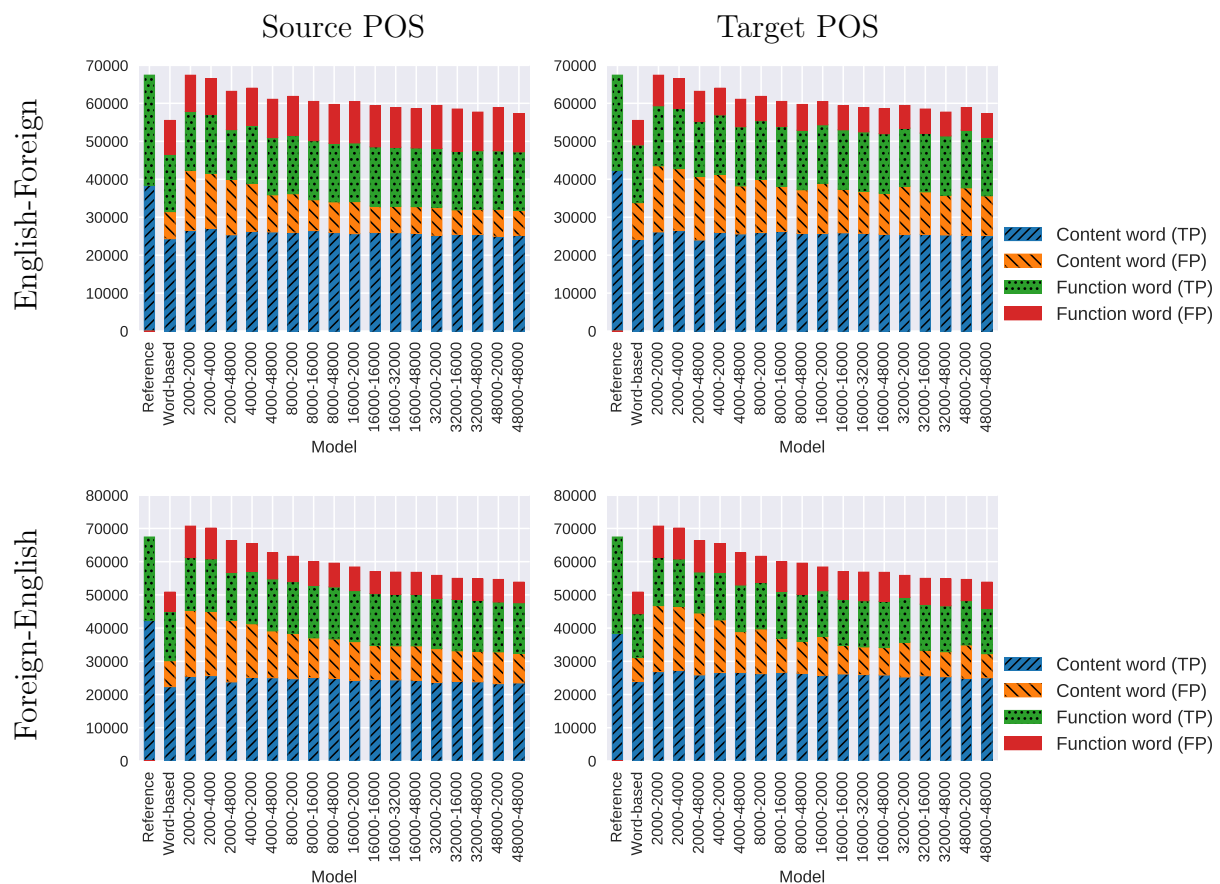


Figure E.99: BPE-based *Fastalign* for English-Czech: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).

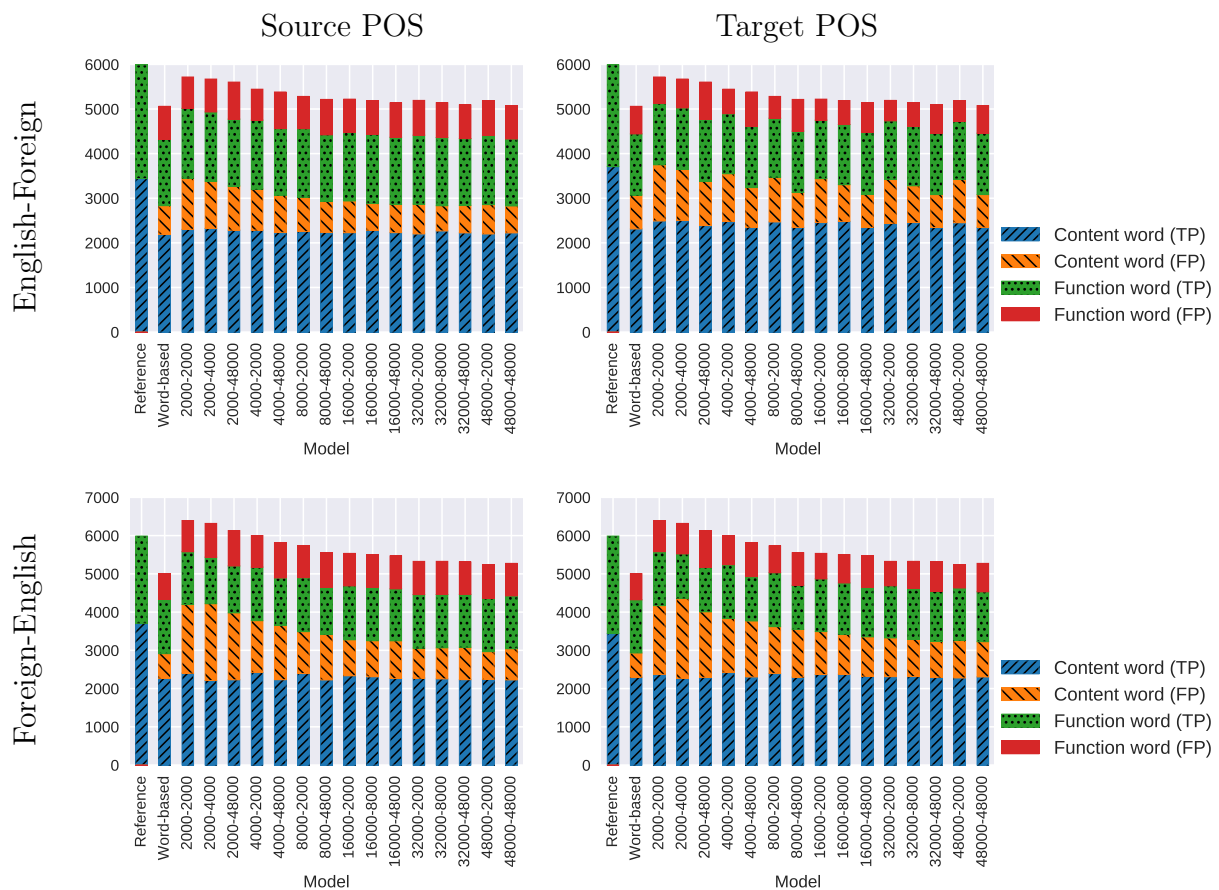


Figure E.100: BPE-based **Fastalign** for English-Romanian: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).

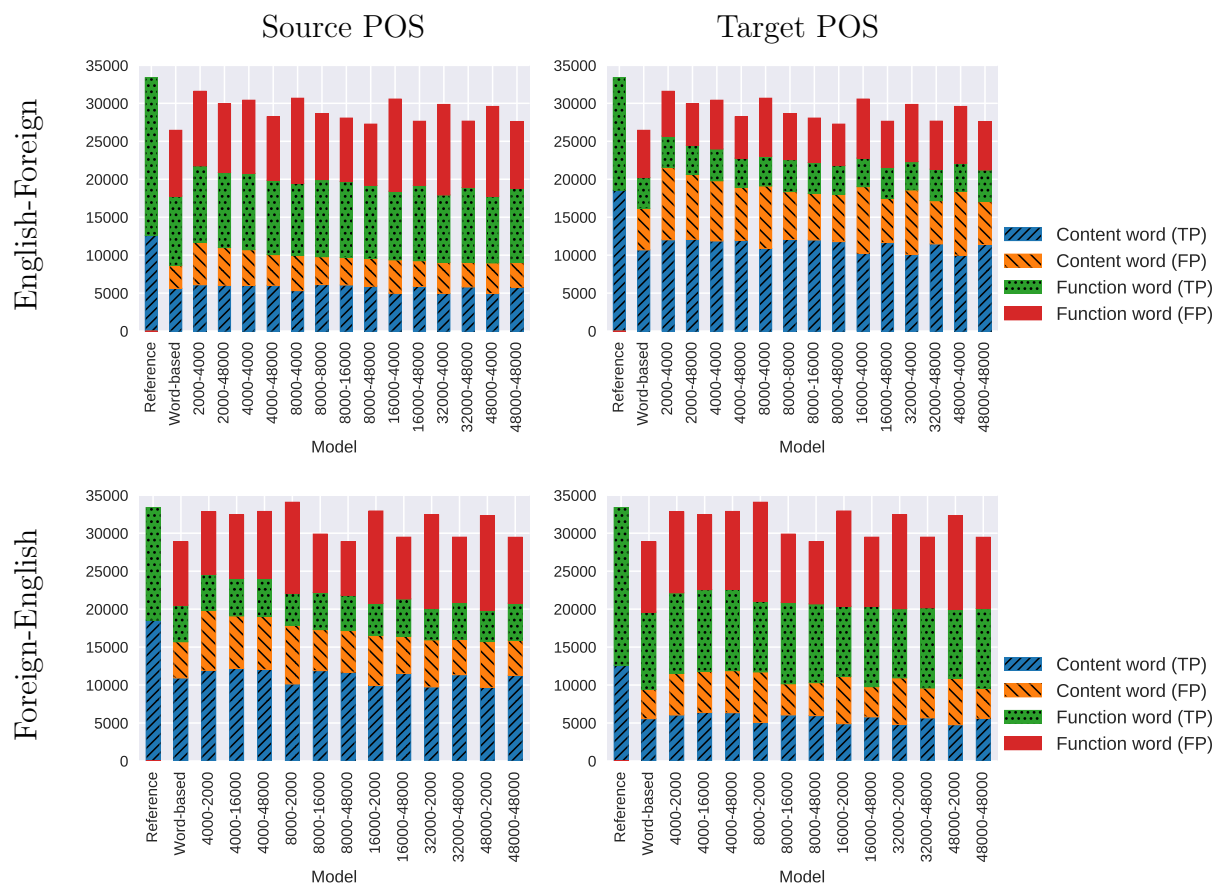


Figure E.101: BPE-based `Fastalign` for English-Japanese: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).

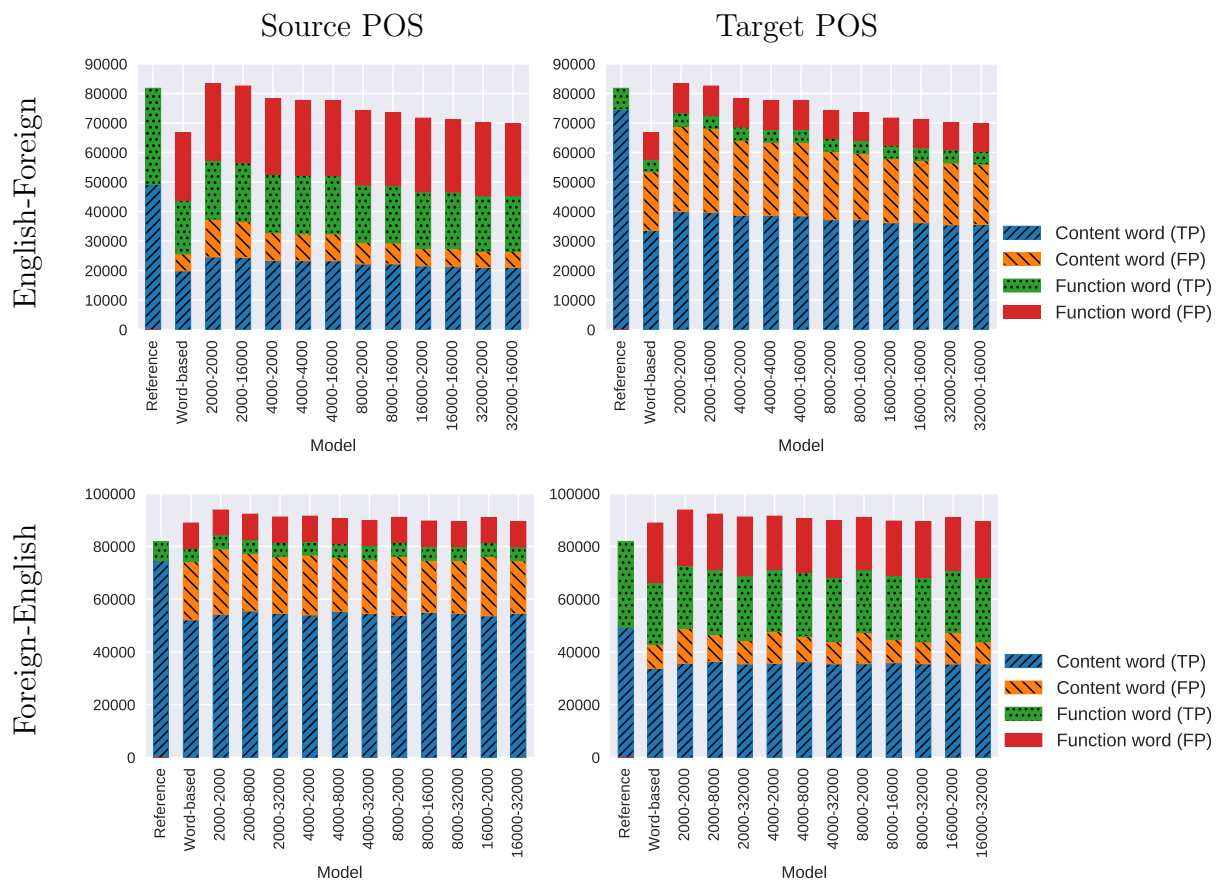


Figure E.102: BPE-based **Fastalign** for English-Vietnamese: The number of target words that align with a content/function source word (Left graph). The number of source words that align with a content/function target words (Right graph).

E.8.2 Unaligned words for content and function words

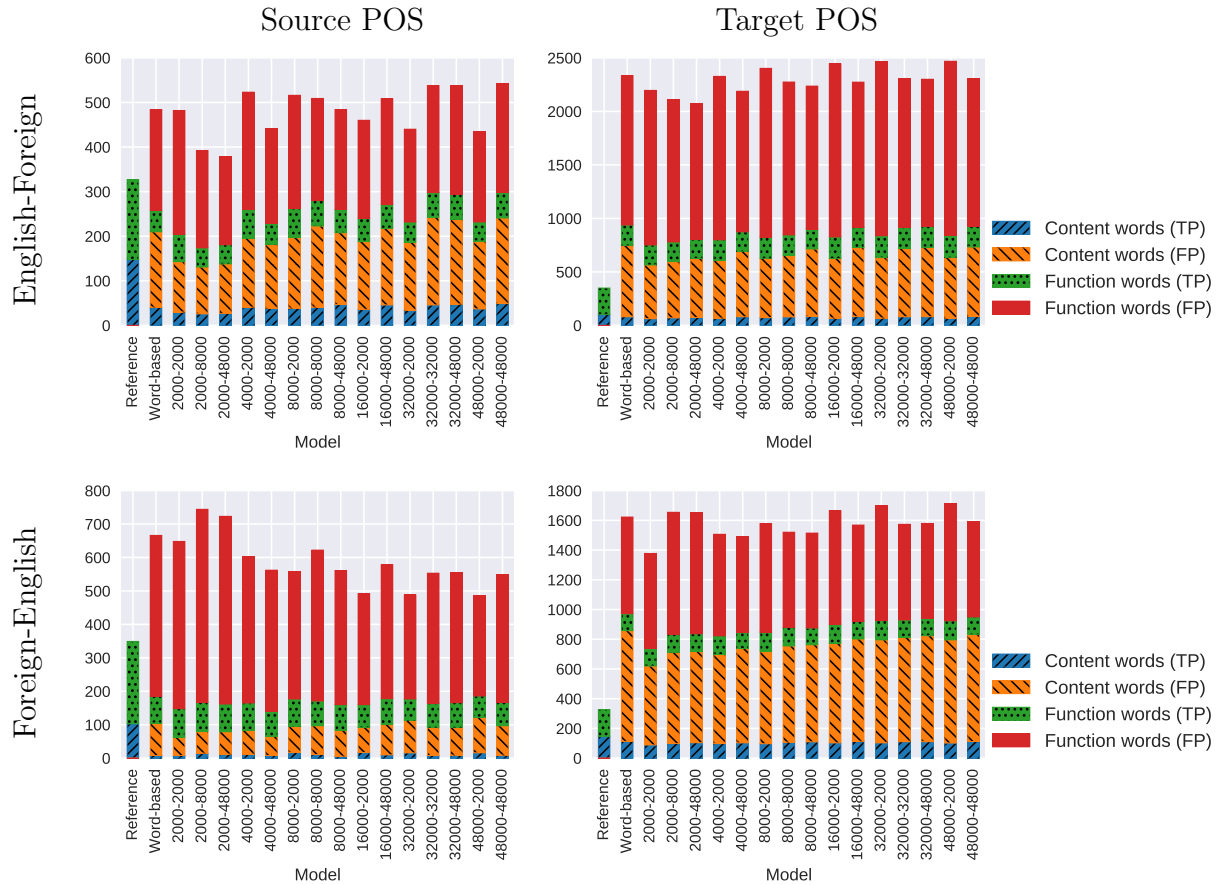


Figure E.103: BPE-based *Fastalign* for English-French: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).



Figure E.104: BPE-based `Fastalign` for English-German: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).



Figure E.105: BPE-based **Fastalign** for English-Czech: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).

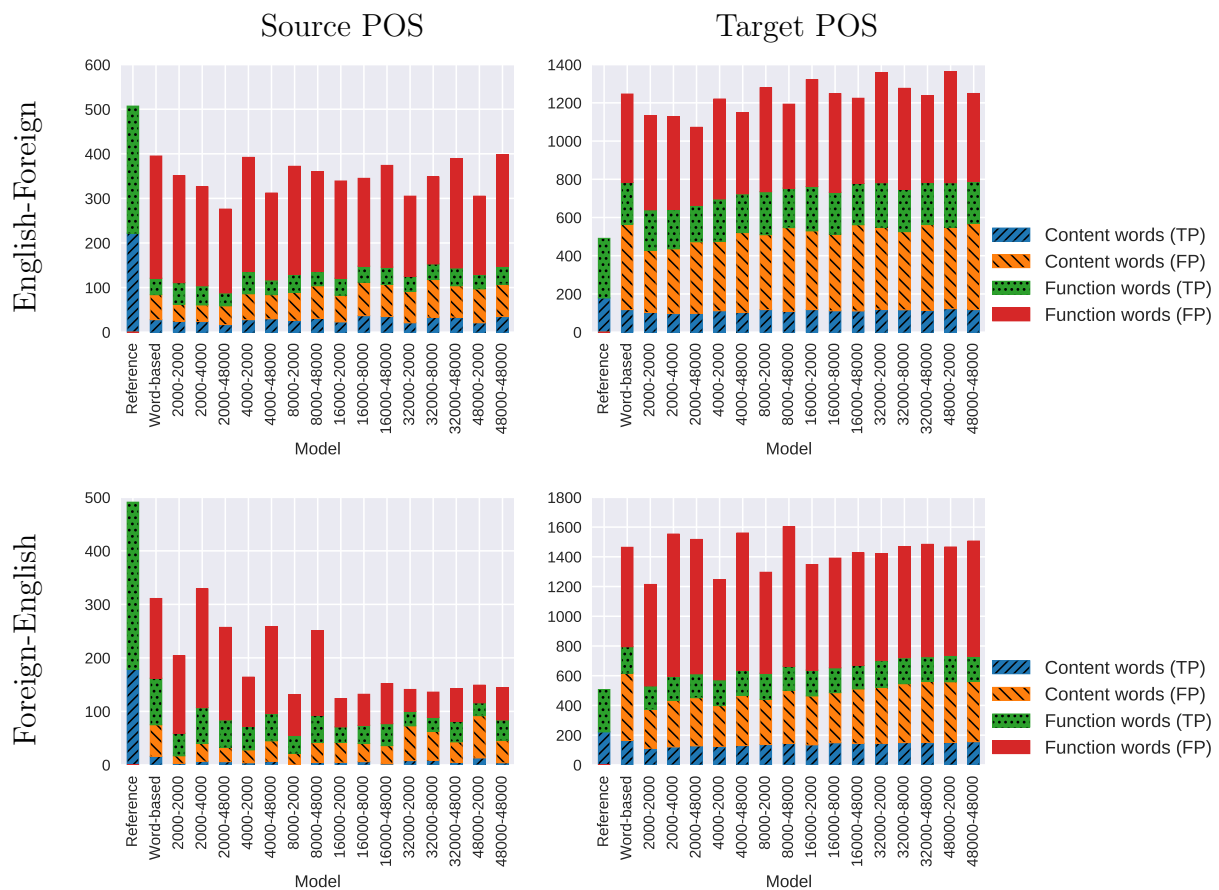


Figure E.106: BPE-based **Fastalign** for English-Romanian: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).

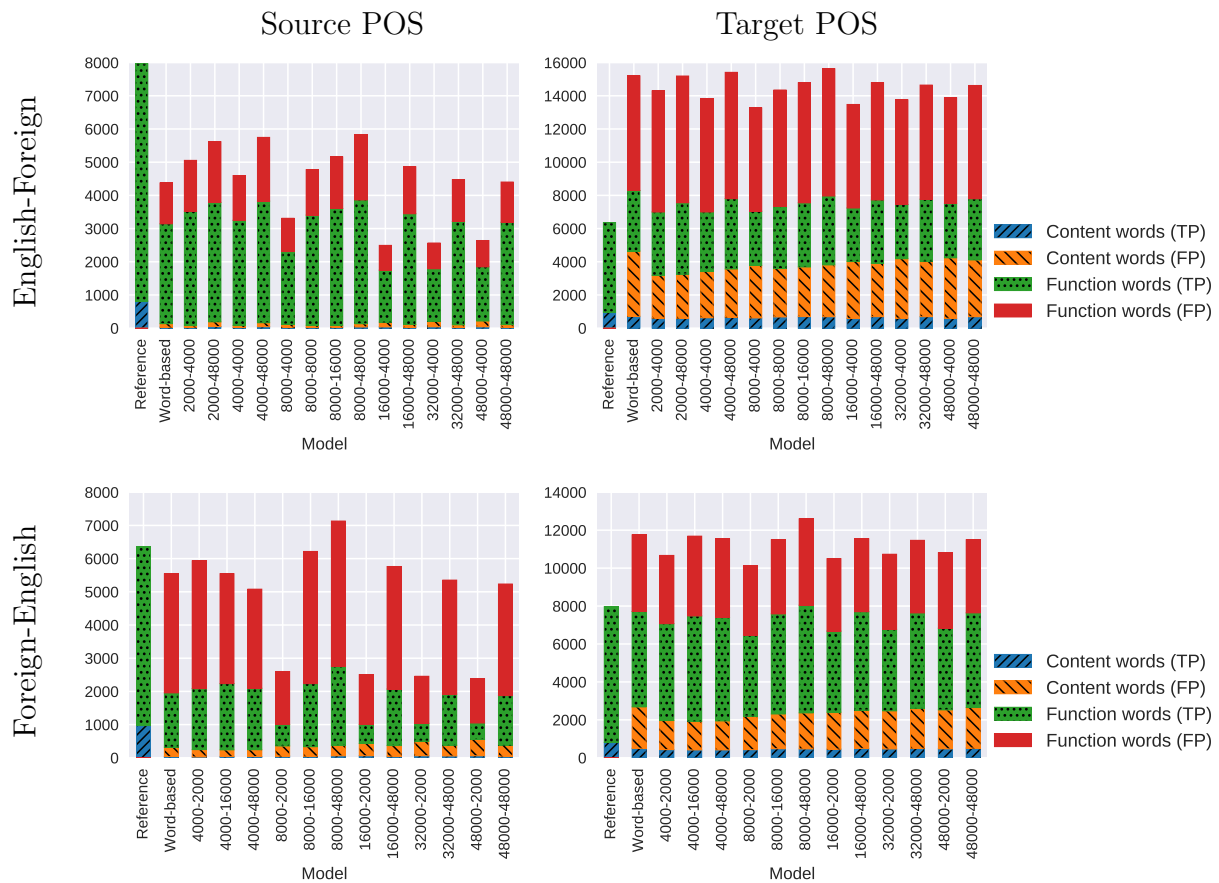


Figure E.107: BPE-based **Fastalign** for English-Japanese: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).

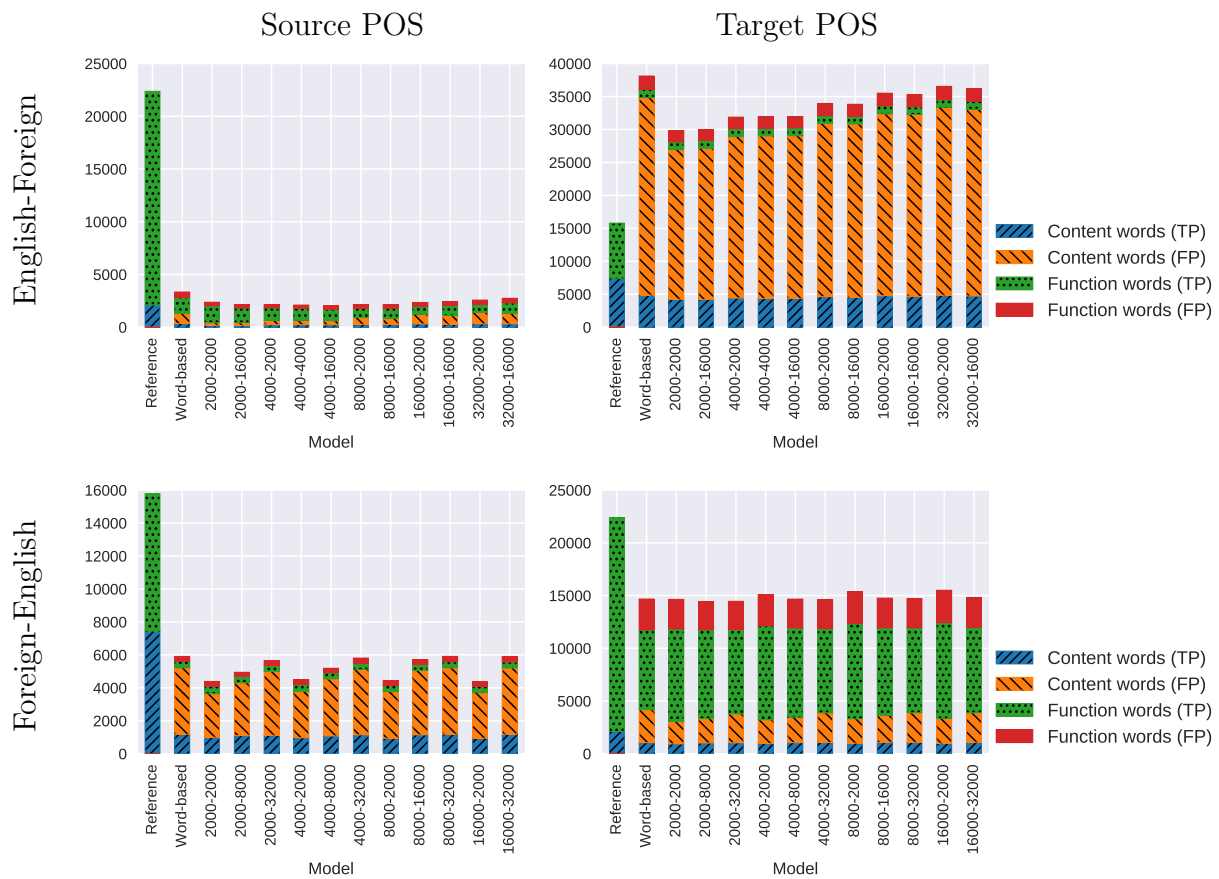


Figure E.108: BPE-based Fastalign for English-Vietnamese: The number of unaligned content/function source words (Left graph). The number of unaligned content/function target words (Right graph).

E.9 Symmetrization

E.9.1 GDF before the transformation step

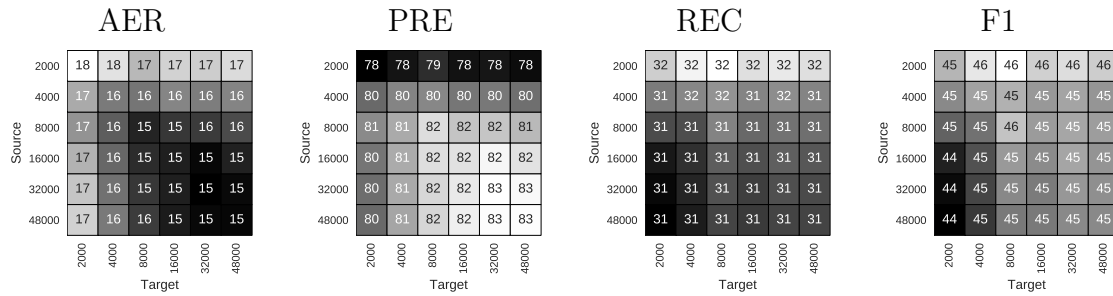


Figure E.109: GDF-before: BPE-based Fastalign for English-French. For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) of symmetrical alignment. Note that we first apply GDF to asymmetrical BPE-based alignments of both directions for obtaining a symmetrical alignment, and then transform it into the word-based alignment.

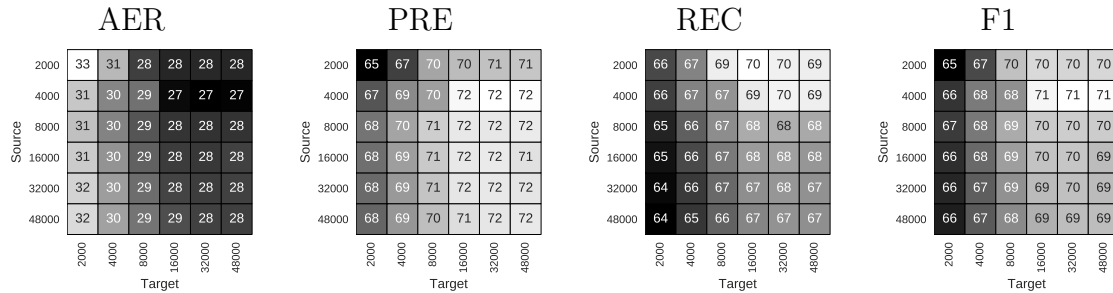


Figure E.110: GDF-before: BPE-based Fastalign for English-German. For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) of symmetrical alignment. Note that we first apply GDF to asymmetrical BPE-based alignments of both directions for obtaining a symmetrical alignment, and then transform it into the word-based alignment.

		AER						PRE						REC						F1					
Source	2000	30	28	27	28	31	32	62	64	65	64	61	61	62	63	63	62	58	57	62	63	64	63	59	59
	4000	28	27	26	25	26	26	65	66	68	68	68	67	62	63	63	62	62	61	63	64	65	65	65	64
	8000	27	26	26	25	25	25	67	68	69	69	70	69	61	62	62	62	62	61	64	65	65	66	65	65
	16000	27	26	25	25	25	25	68	69	69	70	70	70	61	62	62	62	62	61	64	65	65	66	66	65
	32000	27	26	25	25	25	25	69	69	70	70	71	71	60	61	61	61	61	61	64	65	65	65	65	65
	48000	27	26	26	26	25	25	69	69	70	70	71	71	60	60	61	60	61	60	64	65	65	65	65	65
		2000	4000	8000	16000	32000	48000	2000	4000	8000	16000	32000	48000	2000	4000	8000	16000	32000	48000	2000	4000	8000	16000	32000	48000
		Target						Target						Target						Target					

Figure E.111: GDF-before: BPE-based `Fastalign` for English-Czech. For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) of symmetrical alignment. Note that we first apply GDF to asymmetrical BPE-based alignments of both directions for obtaining a symmetrical alignment, and then transform it into the word-based alignment.

		AER						PRE						REC						F1					
Source	2000	34	34	33	34	34	35	68	68	69	69	68	68	65	65	65	64	64	63	66	66	67	66	66	65
	4000	33	32	32	33	34	34	70	72	72	71	70	69	64	64	65	64	63	62	67	68	68	67	66	66
	8000	33	32	32	32	33	34	72	73	73	73	72	71	63	64	64	63	63	62	67	68	68	68	67	66
	16000	33	32	31	32	32	33	72	74	74	73	73	73	63	64	64	63	63	62	67	68	69	68	68	67
	32000	33	32	32	32	33	33	72	74	74	74	73	73	63	63	63	63	62	62	67	68	68	68	67	67
	48000	33	32	32	32	32	33	72	74	74	74	74	73	63	63	63	63	62	62	67	68	68	68	68	67
		2000	4000	8000	16000	32000	48000	2000	4000	8000	16000	32000	48000	2000	4000	8000	16000	32000	48000	2000	4000	8000	16000	32000	48000
		Target						Target						Target						Target					

Figure E.112: GDF-before: BPE-based `Fastalign` for English-Romanian. For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) of symmetrical alignment. Note that we first apply GDF to asymmetrical BPE-based alignments of both directions for obtaining a symmetrical alignment, and then transform it into the word-based alignment.

		AER						PRE						REC						F1					
Source	2000	50	50	50	50	50	51	53	53	53	53	48	48	48	48	48	50	50	50	50	50				
	4000	50	48	48	48	49	53	56	56	56	56	48	48	48	48	47	50	52	52	52	51				
	8000	54	48	48	49	49	48	57	57	57	57	44	49	48	47	47	46	52	52	51	51				
	16000	56	48	48	48	48	46	57	57	57	57	42	48	48	48	47	44	52	52	52	52				
	32000	56	49	48	48	49	46	56	57	57	57	41	47	47	47	47	44	51	52	52	51				
	48000	56	49	49	49	49	46	56	56	56	56	41	47	47	47	47	44	51	51	51	51				
		4000	8000	16000	32000	48000	4000	8000	16000	32000	48000	4000	8000	16000	32000	48000	4000	8000	16000	32000	48000				
		Target						Target						Target						Target					

Figure E.113: GDF-before: BPE-based `Fastalign` for English-Japanese. For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) of symmetrical alignment. Note that we first apply GDF to asymmetrical BPE-based alignments of both directions for obtaining a symmetrical alignment, and then transform it into the word-based alignment.

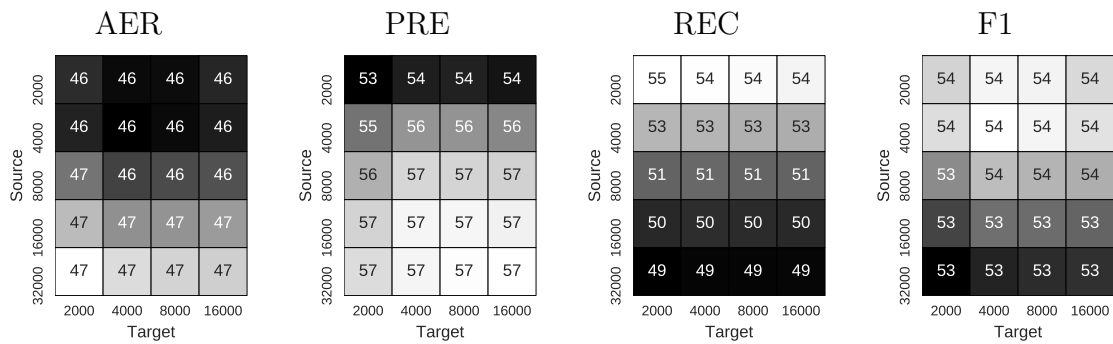


Figure E.114: GDF-before: BPE-based Fastalign for English-Vietnamese. For each pair (vocabulary size of source and target), we display Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) of symmetrical alignment. Note that we first apply GDF to asymmetrical BPE-based alignments of both directions for obtaining a symmetrical alignment, and then transform it into the word-based alignment.

E.9.2 GDF after the transformation step

2000-2000	19	19	21	21	21	21	19	18	19	19	19	19	19	19	18	18	18	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18		
2000-4000	19	19	20	20	20	20	19	18	19	19	19	19	18	18	18	18	18	18	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	
2000-8000	19	19	20	20	20	20	19	18	19	19	19	19	18	18	18	18	18	18	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	
2000-16000	19	19	21	21	21	21	19	18	19	19	19	19	18	18	18	18	18	18	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	
2000-32000	19	19	21	21	21	21	19	18	19	19	19	19	18	18	18	18	18	18	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	
2000-48000	19	19	21	21	21	21	19	18	19	19	19	19	18	18	18	18	18	18	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	
4000-2000	19	19	20	21	20	21	19	18	18	19	19	19	18	18	18	18	18	18	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	
4000-4000	18	19	20	20	20	20	18	18	18	18	19	19	18	18	18	18	18	18	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	
4000-8000	18	19	20	20	20	20	18	18	18	19	19	19	18	18	18	18	18	18	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	
4000-16000	19	19	20	20	20	20	19	18	18	19	19	19	18	18	18	18	18	18	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	
4000-32000	19	19	20	20	20	20	19	18	18	18	19	19	19	18	18	18	18	18	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	
4000-48000	19	19	20	20	20	20	19	18	18	19	19	19	19	18	18	18	18	18	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	
8000-2000	19	19	21	21	21	21	19	18	19	19	19	19	18	18	18	18	18	18	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	
8000-4000	19	19	20	20	20	20	18	18	18	19	19	19	18	18	18	18	18	18	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	
8000-8000	18	18	20	20	20	20	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	
8000-16000	18	18	20	20	20	20	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	
8000-32000	18	18	20	20	20	20	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	
8000-48000	18	18	20	20	20	20	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	
16000-2000	19	19	21	21	21	22	19	19	19	19	19	19	19	18	18	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	
16000-4000	19	19	21	21	21	21	18	18	19	19	19	19	18	18	18	18	19	19	19	18	18	18	18	19	19	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	
16000-8000	18	18	20	20	20	20	18	18	18	18	18	18	18	17	17	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	
16000-16000	18	18	20	20	20	20	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	
16000-32000	18	18	20	20	20	20	18	17	18	18	18	18	18	17	17	17	18	18	18	18	18	18	18	18	18	18	18	18	18	17	17	17	17	18	18	18	18	18	18		
16000-48000	18	18	20	20	20	20	18	18	18	18	18	18	18	18	17	17	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	
32000-2000	19	20	22	22	22	22	19	19	19	19	20	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	
32000-4000	19	19	21	21	21	21	18	18	19	19	19	19	18	18	18	18	19	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18
32000-8000	18	18	20	20	20	20	18	18	18	18	19	19	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18
32000-16000	18	18	20	20	20	20	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18
32000-32000	18	18	20	20	20	20	18	18	18	18	18	18	18	18	17	17	17	18	18	18	18	18	18	18	18	18	18	18	18	17	17	17	18	18	18	18	18	18	18	18	
32000-48000	18	18	20	20	20	20	18	18	18	18	18	18	18	17	17	18	18	18	18	18	18	18	18	18	18	18	18	18	18	17	17	18	18	18	18	18	18	18	18	18	
48000-2000	19	20	22	22	22	22	19	19	19	19	20	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	
48000-4000	19	19	21	21	21	21	19	18	19	19	19	19	18	18	18	18	19	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18	19	19	18	18	18	18
48000-8000	18	19	20	20	20	20	18	18	18	18	19	19	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18
48000-16000	18	18	20	20	20	20	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18
48000-32000	18	18	20	20	20	20	18	18	18	18	18	18	18	17	17	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	
48000-48000	18	18	20	20	20	20	18	18	18	18	18	18	18	17	17	17	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18

Figure E.115: GDF-after: BPE-based Fastalign for English-French. we display Alignment error rate (AER) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.

	4000-2000	49	47	47	47	47	47	53	48	47	47	47	47	53	52	48	47	47	47	53	52	49	48	48	47	53	52	49	48	48	47	
	4000-4000	48	47	47	47	47	47	52	48	47	47	46	47	52	51	47	47	47	47	52	52	48	47	47	47	52	52	48	48	47	47	
	4000-8000	48	47	47	47	47	47	52	47	47	46	46	46	52	51	48	47	47	47	52	52	48	48	47	47	52	52	49	48	47	47	
	4000-16000	48	47	47	47	47	47	52	47	47	46	46	47	52	51	48	47	47	47	52	52	48	48	47	47	52	52	49	48	47	47	
	4000-32000	48	47	47	47	47	47	52	48	47	47	47	47	52	51	48	47	47	47	52	52	48	48	47	47	52	52	49	48	48	47	
	4000-48000	48	47	47	47	47	47	53	48	47	47	47	47	53	52	48	47	47	47	53	52	48	48	47	47	53	52	49	48	48	47	
	8000-2000	47	46	46	46	46	46	51	47	46	46	46	46	51	51	47	46	46	46	52	51	48	47	47	46	52	51	48	47	47	47	
	8000-4000	47	46	46	46	47	47	51	47	46	46	46	46	52	51	47	47	47	47	52	51	48	47	47	47	52	51	48	47	47	47	
	8000-8000	47	46	46	46	46	46	51	47	46	46	46	46	52	51	47	47	47	47	52	51	48	47	47	47	52	51	48	47	47	47	
	8000-16000	47	46	46	46	47	47	51	47	47	46	46	46	52	51	47	47	47	47	52	51	48	47	47	47	52	52	48	48	47	47	
	8000-32000	51	50	50	50	50	50	56	51	50	50	49	50	56	55	50	50	50	50	56	55	51	50	50	50	56	55	51	51	50	50	
	8000-48000	47	46	46	46	46	46	51	47	46	46	46	46	51	51	47	46	46	46	51	51	48	47	47	46	51	51	48	47	47	47	
	16000-2000	47	46	46	46	46	46	51	47	46	46	46	46	51	51	47	47	47	46	51	51	48	47	47	47	52	51	48	47	47	47	
	16000-4000	47	46	46	46	47	47	51	47	47	46	46	46	52	51	47	47	47	47	52	51	48	47	47	47	52	51	48	48	47	47	
	16000-8000	47	46	46	47	47	47	51	47	47	46	46	46	52	51	47	47	47	47	52	51	48	47	47	47	52	52	49	48	47	47	
	16000-16000	52	51	51	51	52	52	57	52	51	51	51	51	57	56	52	51	51	51	57	56	52	52	51	51	57	56	53	52	52	51	
	16000-32000	47	46	46	46	47	47	51	47	47	46	46	46	52	51	47	47	47	47	52	51	48	47	47	47	52	51	48	47	47	47	
	16000-48000	47	46	46	46	47	47	51	47	46	46	46	46	52	51	47	47	47	47	52	51	48	47	47	47	52	51	48	47	47	47	
	32000-2000	47	46	46	46	47	47	51	47	47	46	46	46	52	51	47	47	47	47	52	51	48	47	47	47	52	51	48	48	47	47	
	32000-4000	47	46	46	46	47	47	51	47	47	46	46	46	52	51	47	47	47	47	52	51	48	48	47	47	52	51	48	48	47	47	
	32000-8000	52	51	51	51	52	52	57	52	51	51	51	51	57	56	52	51	51	51	57	56	52	52	52	51	57	56	53	52	52	52	
	32000-16000	48	47	47	47	47	47	52	47	47	47	47	47	52	51	47	47	47	47	52	51	48	47	47	47	52	52	48	48	47	47	
	32000-32000	47	46	46	46	47	47	52	47	47	46	46	46	52	51	47	47	47	47	52	51	48	47	47	47	52	51	48	48	47	47	
	32000-48000	47	47	46	47	47	47	52	47	47	46	47	46	52	51	47	47	47	47	52	51	48	48	48	47	52	52	48	48	48	47	
	48000-2000	48	47	47	47	47	47	52	47	47	47	47	47	52	51	48	47	47	47	52	52	48	48	48	48	52	52	49	48	48	48	
	48000-4000	52	51	51	51	52	52	57	52	51	51	51	51	57	56	52	51	51	51	57	56	52	52	52	52	57	56	53	52	52	52	
	48000-8000	48	47	47	47	47	47	52	48	47	47	47	47	52	51	48	47	47	47	52	52	48	48	47	47	52	52	49	48	48	47	
	48000-16000	48	47	47	47	47	47	52	47	47	47	47	47	52	51	48	47	47	47	52	52	48	48	47	47	52	52	48	48	48	47	
	48000-32000	48	47	47	47	47	47	52	47	47	47	47	47	52	51	48	47	47	47	52	52	48	48	48	48	52	52	49	48	48	48	
	48000-48000	48	47	47	47	47	47	52	47	47	47	47	47	52	51	48	47	47	47	52	52	48	48	48	48	52	52	49	48	48	48	
Source																																
	4000-2000																															
	4000-4000																															
	4000-8000																															
	4000-16000																															
	4000-32000																															
	4000-48000																															
	8000-2000																															
	8000-4000																															
	8000-8000																															
	8000-16000																															
	8000-32000																															
	8000-48000																															
	16000-2000																															
	16000-4000																															
	16000-8000																															
	16000-16000																															
	16000-32000																															
	16000-48000																															
	32000-2000																															
	32000-4000																															
	32000-8000																															
	32000-16000																															
	32000-32000																															
	32000-48000																															
	48000-2000																															
	48000-4000																															
	48000-8000																															
	48000-16000																															
	48000-32000																															
	48000-48000																															
	Target																															

Figure E.131: GDF-after: BPE-based `Fastalign` for English-Japanese. we display Alignment error rate (AER) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.

	2000-2000	35	34	33	33	33	34	33	33	33	33	35	34	33	33	33	35	34	33	33	33
	2000-4000	34	33	33	33	33	34	33	33	33	33	35	34	33	33	33	35	34	33	33	33
	2000-8000	34	33	33	33	33	34	33	33	33	33	35	34	33	33	33	35	34	33	33	33
	2000-16000	34	33	33	33	33	34	34	33	33	33	35	34	33	33	33	35	34	33	33	33
	2000-32000	35	34	33	33	34	34	34	33	33	33	35	34	33	33	33	35	34	33	33	33
	4000-2000	34	34	33	33	33	34	34	33	33	33	34	34	33	33	33	35	34	33	33	33
	4000-4000	34	34	33	33	33	34	34	33	33	33	34	34	33	33	33	35	34	33	33	33
	4000-8000	34	34	33	33	33	34	34	33	33	33	34	34	33	33	33	35	34	33	33	33
	4000-16000	35	34	34	34	34	35	34	34	34	34	35	34	34	34	34	35	34	34	34	34
	4000-32000	34	34	33	34	34	34	34	33	33	34	34	34	34	33	34	35	34	34	34	34
	8000-2000	34	34	33	34	34	34	34	33	33	34	34	34	34	33	34	35	34	34	34	34
	8000-4000	34	34	33	33	34	34	34	33	33	34	34	34	34	33	34	35	34	34	34	34
	8000-8000	35	34	34	34	34	35	34	34	34	34	35	34	34	34	34	35	34	34	34	34
	8000-16000	34	34	33	34	34	34	34	33	34	34	35	34	34	34	34	35	34	34	34	34
	8000-32000	34	34	33	34	34	34	34	33	34	34	35	34	33	34	34	35	34	34	34	34
	16000-2000	34	34	33	34	34	34	34	33	34	34	35	34	34	34	34	35	34	34	34	34
	16000-4000	35	34	34	34	35	35	34	34	34	34	35	34	34	34	34	35	34	34	34	34
	16000-8000	35	34	33	34	34	35	34	33	34	34	35	34	34	34	34	35	34	34	34	34
	16000-16000	34	34	33	34	34	34	34	33	34	34	35	34	34	34	34	35	34	34	34	34
	16000-32000	34	34	33	34	34	34	34	33	34	34	35	34	33	34	34	35	34	34	34	34
Source		2000-2000	2000-4000	2000-8000	2000-16000	2000-32000	4000-2000	4000-4000	4000-8000	4000-16000	4000-32000	8000-2000	8000-4000	8000-8000	8000-16000	8000-32000	16000-2000	16000-4000	16000-8000	16000-16000	16000-32000
		Target																			

Figure E.135: GDF-after: BPE-based `Fastalign` for English-Vietnamese. we display Alignment error rate (AER) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.

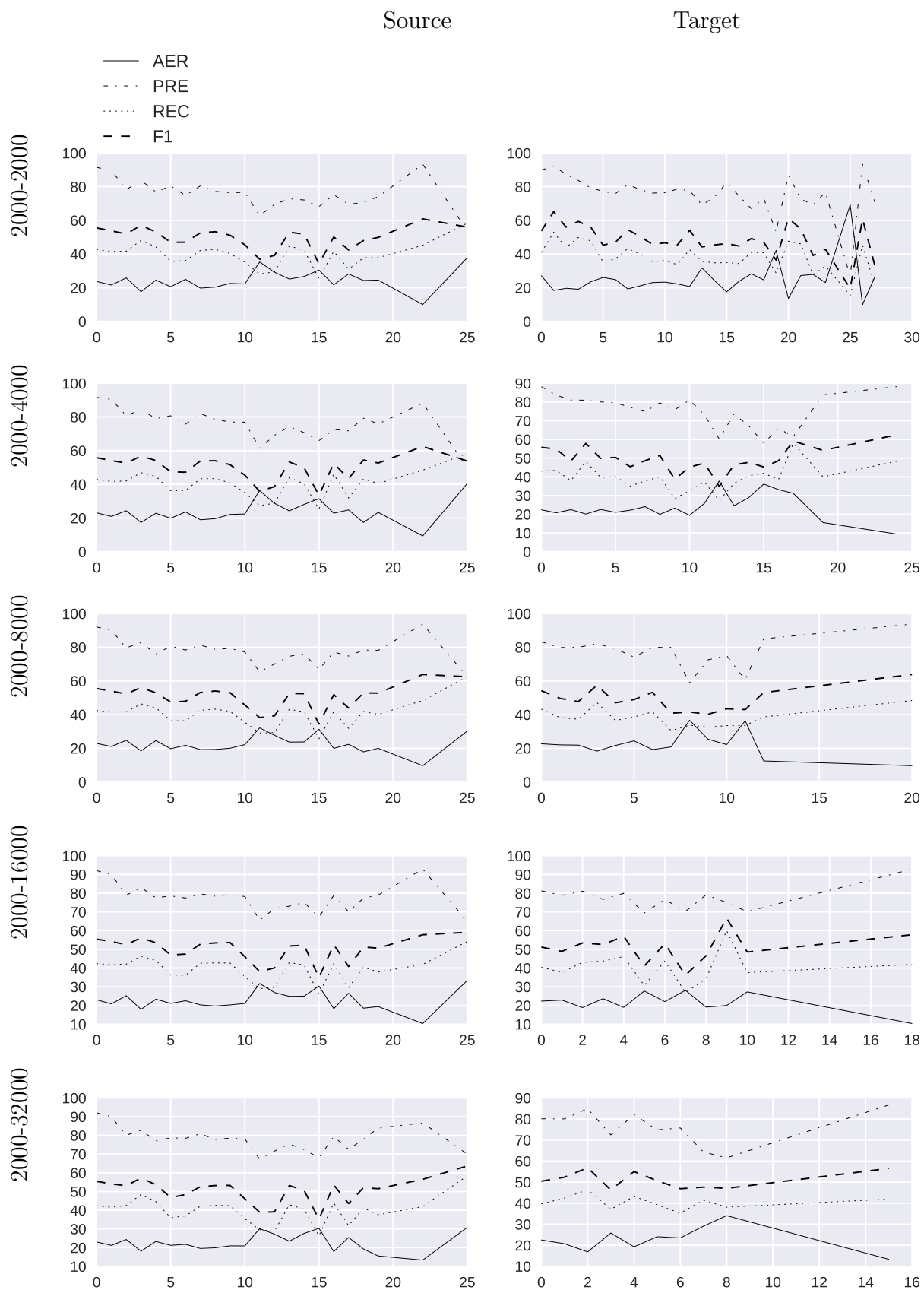
2000-2000	65	66	67	67	67	66	67	67	67	67	65	66	67	67	67	65	66	67	67	67
2000-4000	66	67	67	67	67	66	67	67	67	67	65	66	67	67	67	65	66	67	67	67
2000-8000	66	67	67	67	67	66	67	67	67	67	65	66	67	67	67	65	66	67	67	67
2000-16000	66	67	67	67	67	66	67	67	67	67	65	66	67	67	67	65	66	67	67	67
2000-32000	65	66	67	67	66	66	66	67	67	67	66	66	67	67	67	65	66	67	67	67
4000-2000	66	66	67	67	67	66	66	67	67	67	66	66	67	67	67	65	66	67	67	67
4000-4000	66	66	67	67	67	66	66	67	67	67	66	66	67	67	67	65	66	67	67	67
4000-8000	66	66	67	67	67	66	66	67	67	67	66	66	67	67	67	65	66	67	67	67
4000-16000	65	66	66	66	66	65	66	66	66	66	65	66	66	66	66	65	66	66	66	66
4000-32000	66	66	67	66	66	66	66	67	67	66	66	66	66	67	66	65	66	66	66	66
8000-2000	66	66	67	66	66	66	66	67	67	66	66	66	66	67	66	65	66	66	66	66
8000-4000	66	66	67	67	66	66	66	67	67	66	66	66	66	67	66	65	66	66	66	66
8000-8000	65	66	66	66	66	65	66	66	66	66	65	66	66	66	66	65	66	66	66	66
8000-16000	66	66	67	66	66	66	66	67	66	66	65	66	66	66	66	65	66	66	66	66
8000-32000	66	66	67	66	66	66	66	67	66	66	65	66	67	66	66	65	66	66	66	66
16000-2000	66	66	67	66	66	66	66	67	66	66	65	66	66	66	66	65	66	66	66	66
16000-4000	65	66	66	66	65	65	66	66	66	66	65	66	66	66	66	65	66	66	66	66
16000-8000	65	66	67	66	66	65	66	67	66	66	65	66	66	66	66	65	66	66	66	66
16000-16000	66	66	67	66	66	66	66	67	66	66	65	66	66	66	66	65	66	66	66	66
16000-32000	66	66	67	66	66	66	66	67	66	66	65	66	67	66	66	65	66	66	66	66
	2000-2000	2000-4000	2000-8000	2000-16000	2000-32000	4000-2000	4000-4000	4000-8000	4000-16000	4000-32000	8000-2000	8000-4000	8000-8000	8000-16000	8000-32000	16000-2000	16000-4000	16000-8000	16000-16000	16000-32000

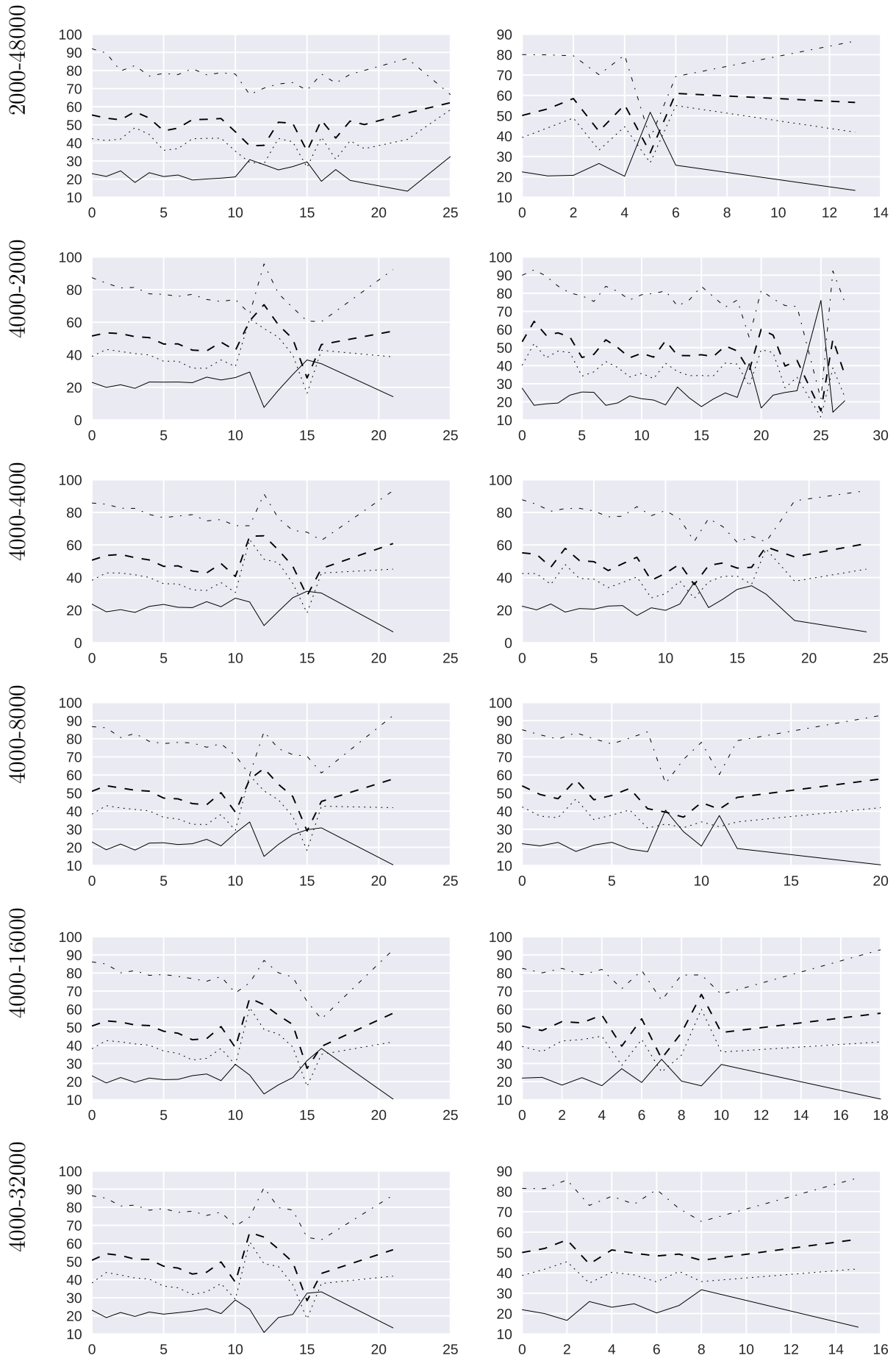
Figure E.136: GDF-after: BPE-based Fastalign for English-Vietnamese. We show F-score (F1) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.

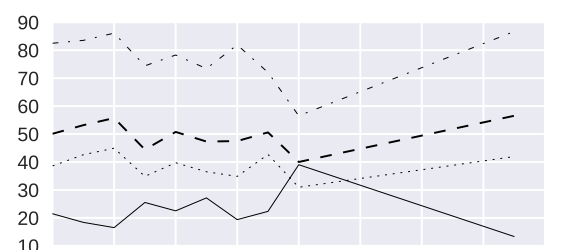
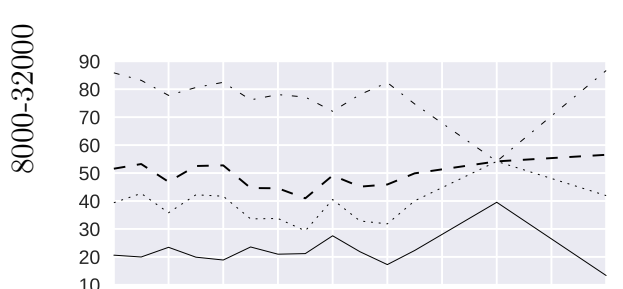
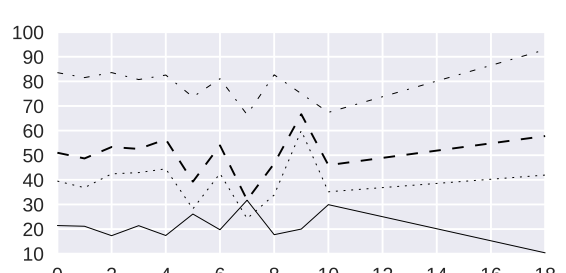
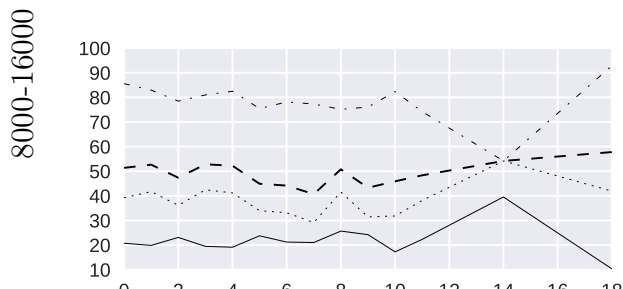
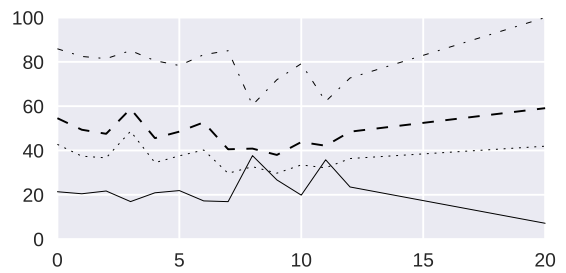
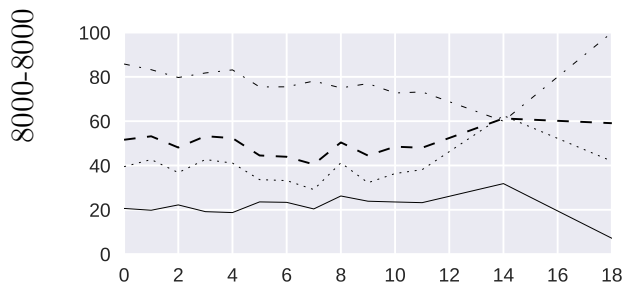
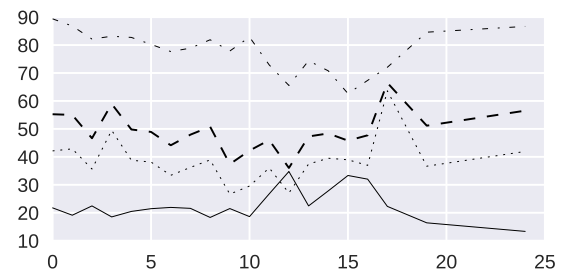
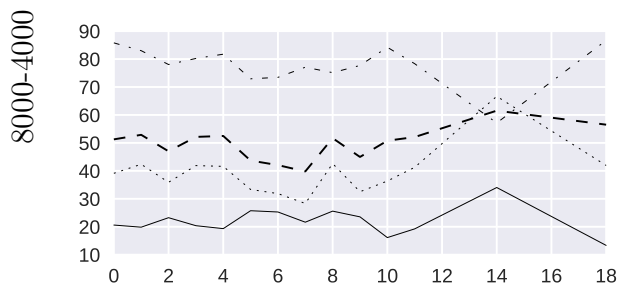
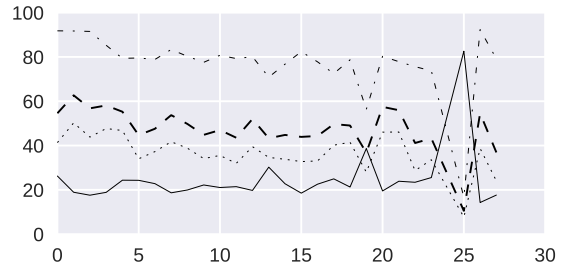
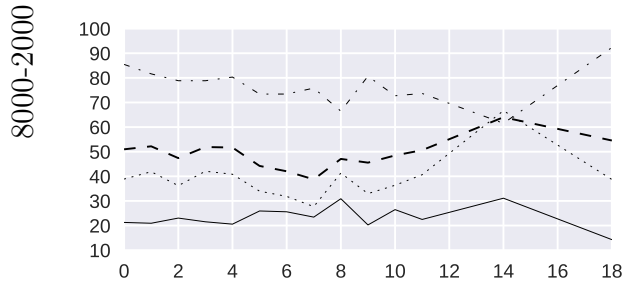
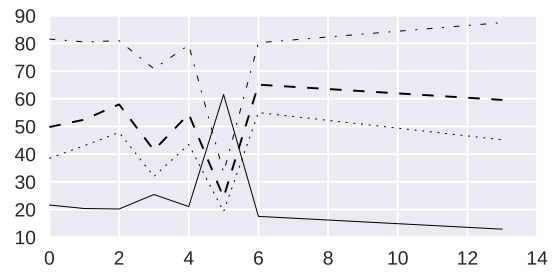
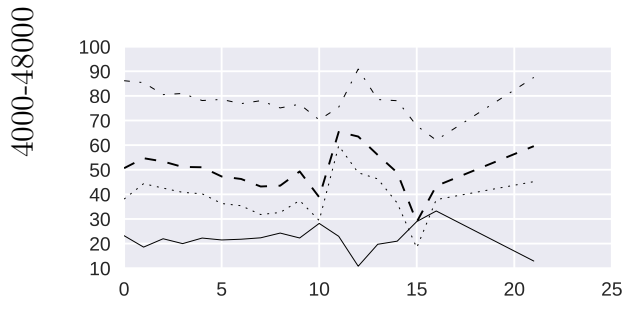
2000-2000	58	59	60	60	60	59	59	60	60	60	58	59	60	60	60	58	59	60	60	60
2000-4000	59	59	60	60	60	59	59	60	60	60	58	59	60	60	60	58	59	60	60	60
2000-8000	59	59	60	60	60	59	59	60	60	60	58	59	60	60	60	58	59	60	60	60
2000-16000	59	59	60	60	60	59	59	60	60	60	58	59	60	60	60	58	59	60	60	60
2000-32000	58	59	60	60	59	59	59	60	60	60	58	59	60	60	60	58	59	60	60	60
4000-2000	59	59	60	60	60	59	59	60	60	60	59	59	60	60	60	58	59	60	60	60
4000-4000	59	59	60	60	60	59	59	60	60	60	59	59	60	60	60	58	59	60	60	60
4000-8000	59	59	60	60	60	59	59	60	60	60	59	59	60	60	60	58	59	60	60	60
4000-16000	58	59	59	59	59	58	59	60	60	59	58	59	60	60	60	58	59	59	60	60
4000-32000	59	60	60	60	60	59	59	60	60	60	59	59	60	60	60	58	59	60	60	60
8000-2000	59	60	60	60	60	59	60	60	60	60	59	59	60	60	60	58	59	60	60	60
8000-4000	59	60	60	60	60	59	60	60	60	60	59	59	60	60	60	58	59	60	60	60
8000-8000	58	59	60	59	59	58	59	60	60	59	58	59	60	60	59	58	59	59	59	59
8000-16000	59	59	60	60	59	59	59	60	60	60	59	59	60	60	60	58	59	60	60	60
8000-32000	59	60	60	60	59	59	60	60	60	60	59	59	60	60	60	58	59	60	60	60
16000-2000	59	60	60	60	59	59	60	60	60	60	59	59	60	60	60	58	59	60	60	60
16000-4000	58	59	59	59	59	58	59	60	59	59	58	59	59	59	59	58	59	59	59	59
16000-8000	59	59	60	60	59	59	59	60	60	59	59	59	60	60	60	58	59	60	60	59
16000-16000	59	60	60	60	59	59	60	60	60	60	59	59	60	60	60	58	59	60	60	60
16000-32000	59	60	60	60	59	59	60	60	60	60	59	59	60	60	60	58	59	60	60	60
	2000-2000	2000-4000	2000-8000	2000-16000	2000-32000	4000-2000	4000-4000	4000-8000	4000-16000	4000-32000	8000-2000	8000-4000	8000-8000	8000-16000	8000-32000	16000-2000	16000-4000	16000-8000	16000-16000	16000-32000
	Target																			

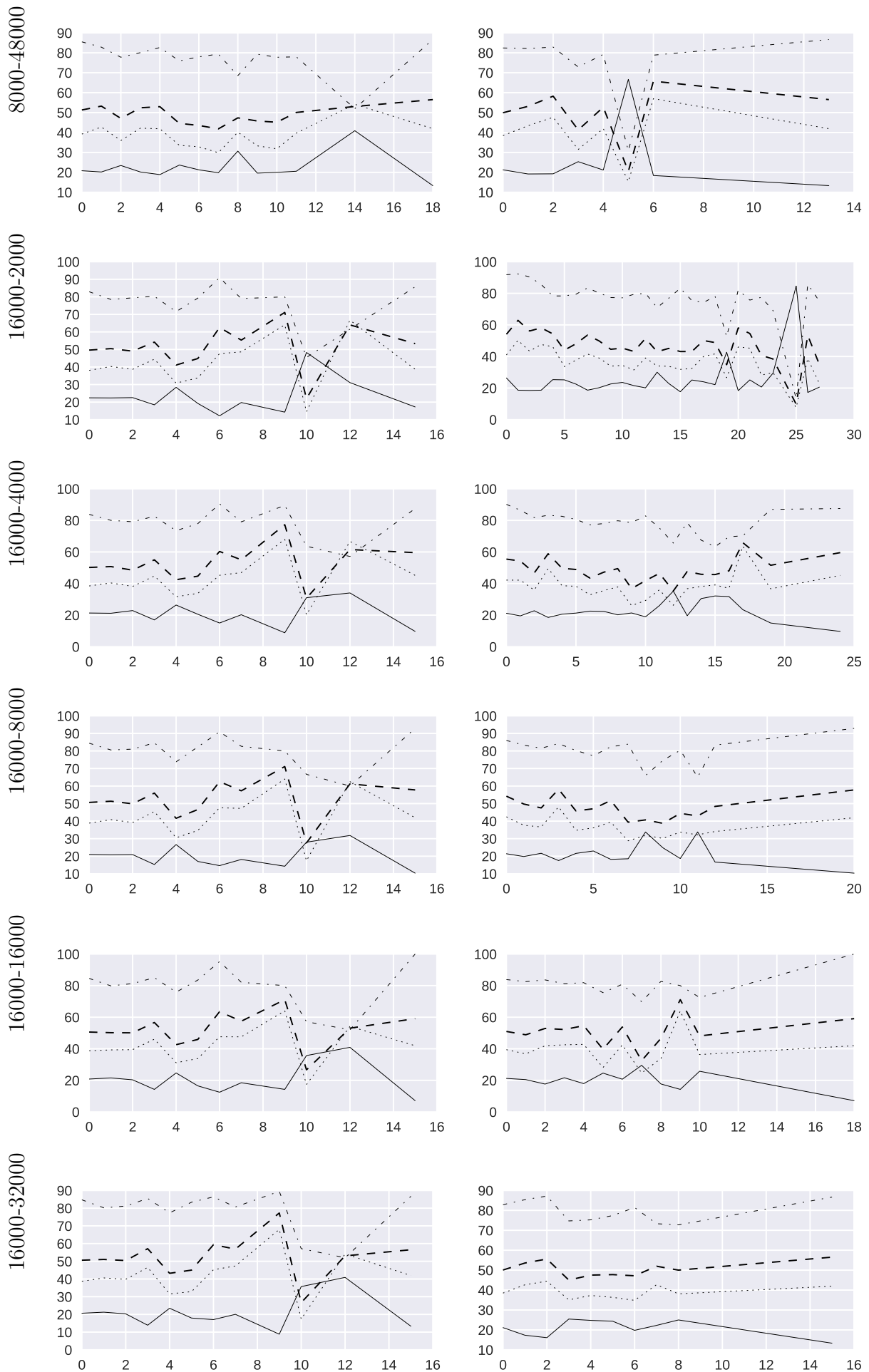
Figure E.137: GDF-after: BPE-based **Fastalign** for English-Vietnamese. We show precision (PRE) of symmetrical alignment. Note that we first transform asymmetrical BPE-based alignments to word-based alignments, and then apply GDF to the word-based alignments for obtaining the symmetrical alignments.

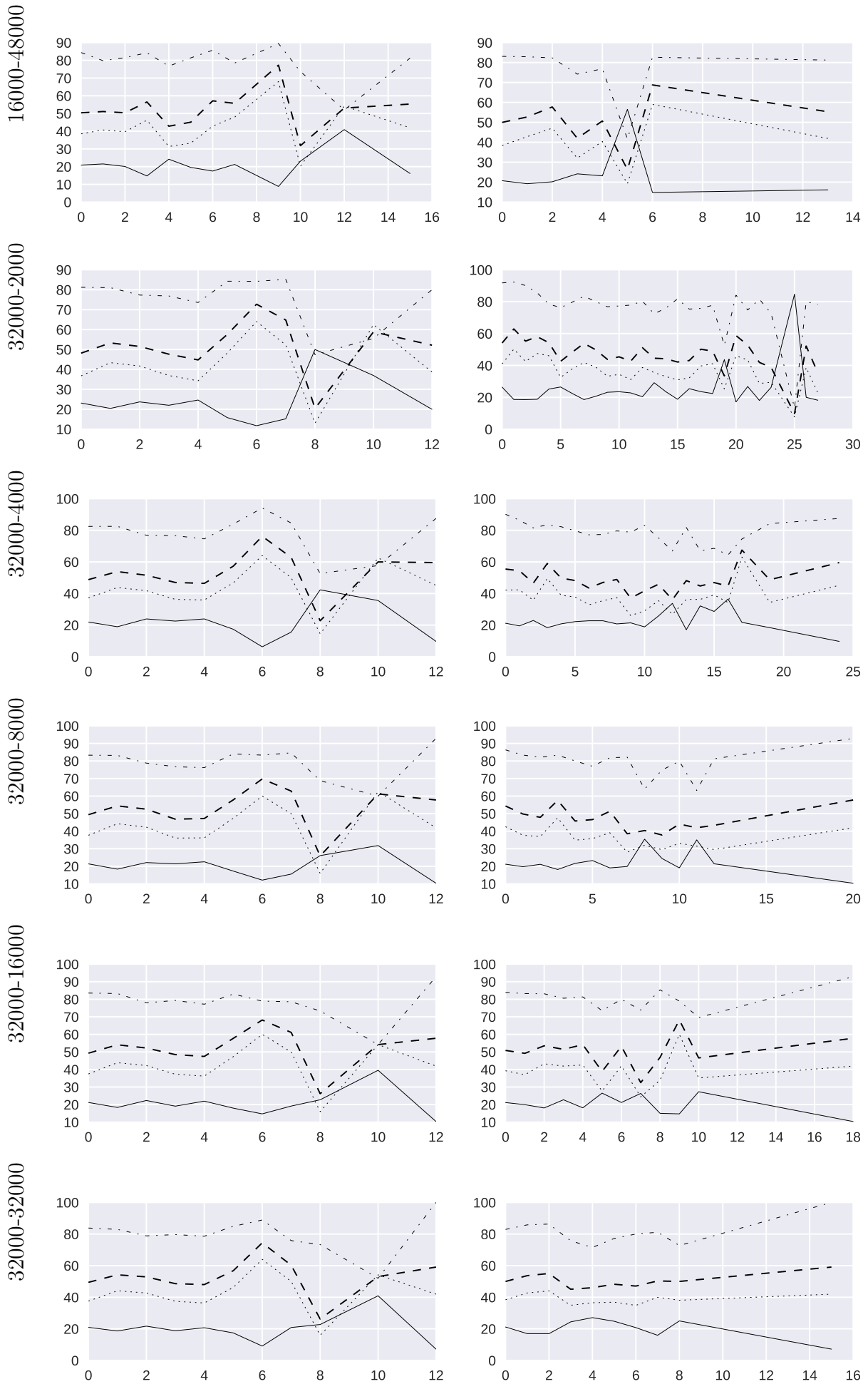
E.10 Scores as a function of sentence length



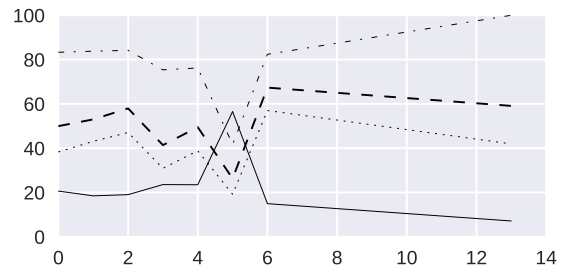
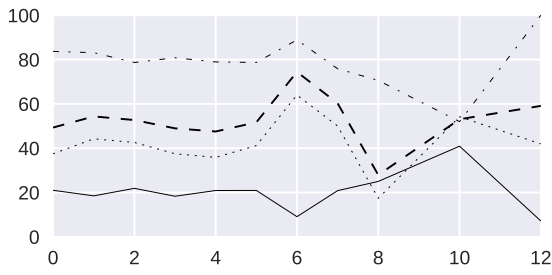




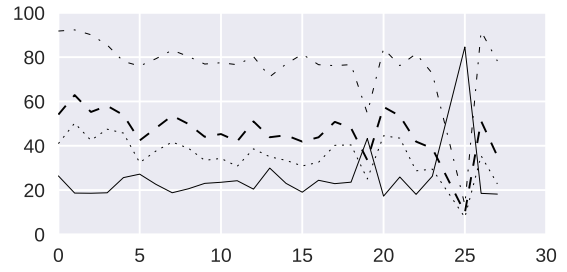
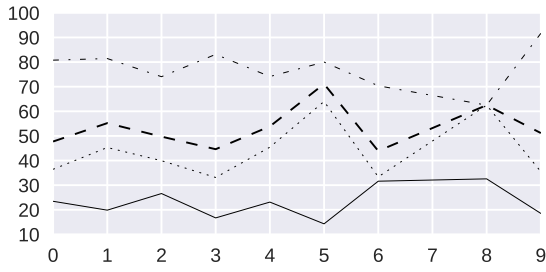




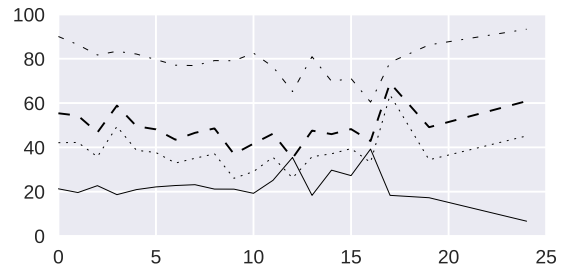
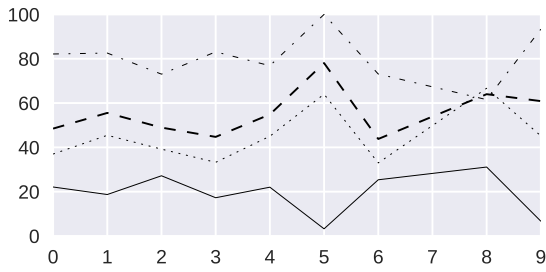
32000-48000



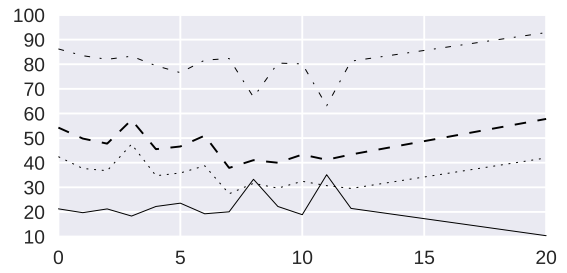
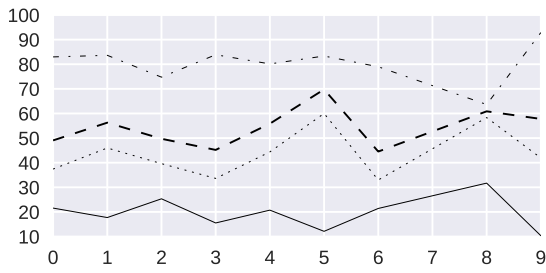
48000-20000



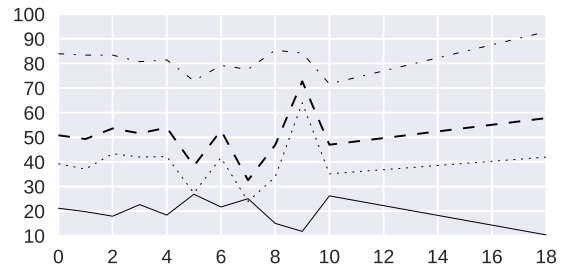
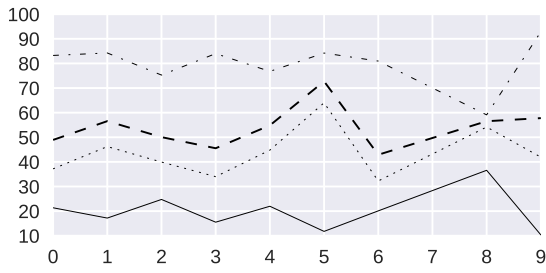
48000-4000



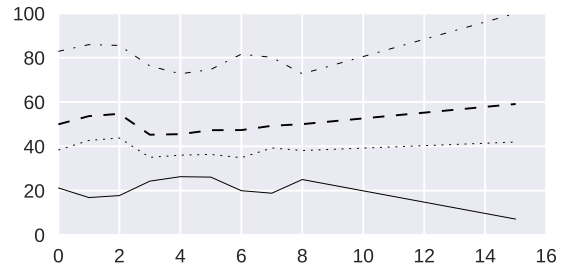
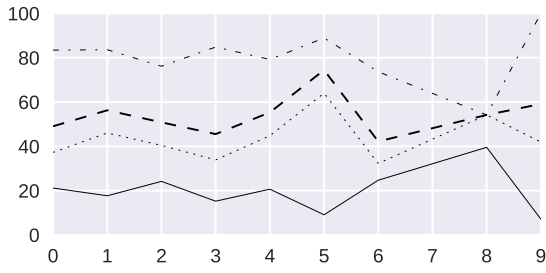
48000-8000



48000-16000



48000-32000



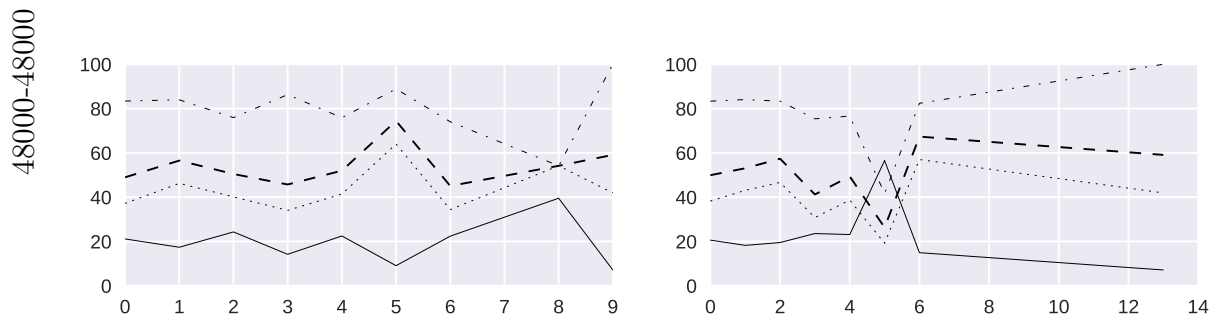
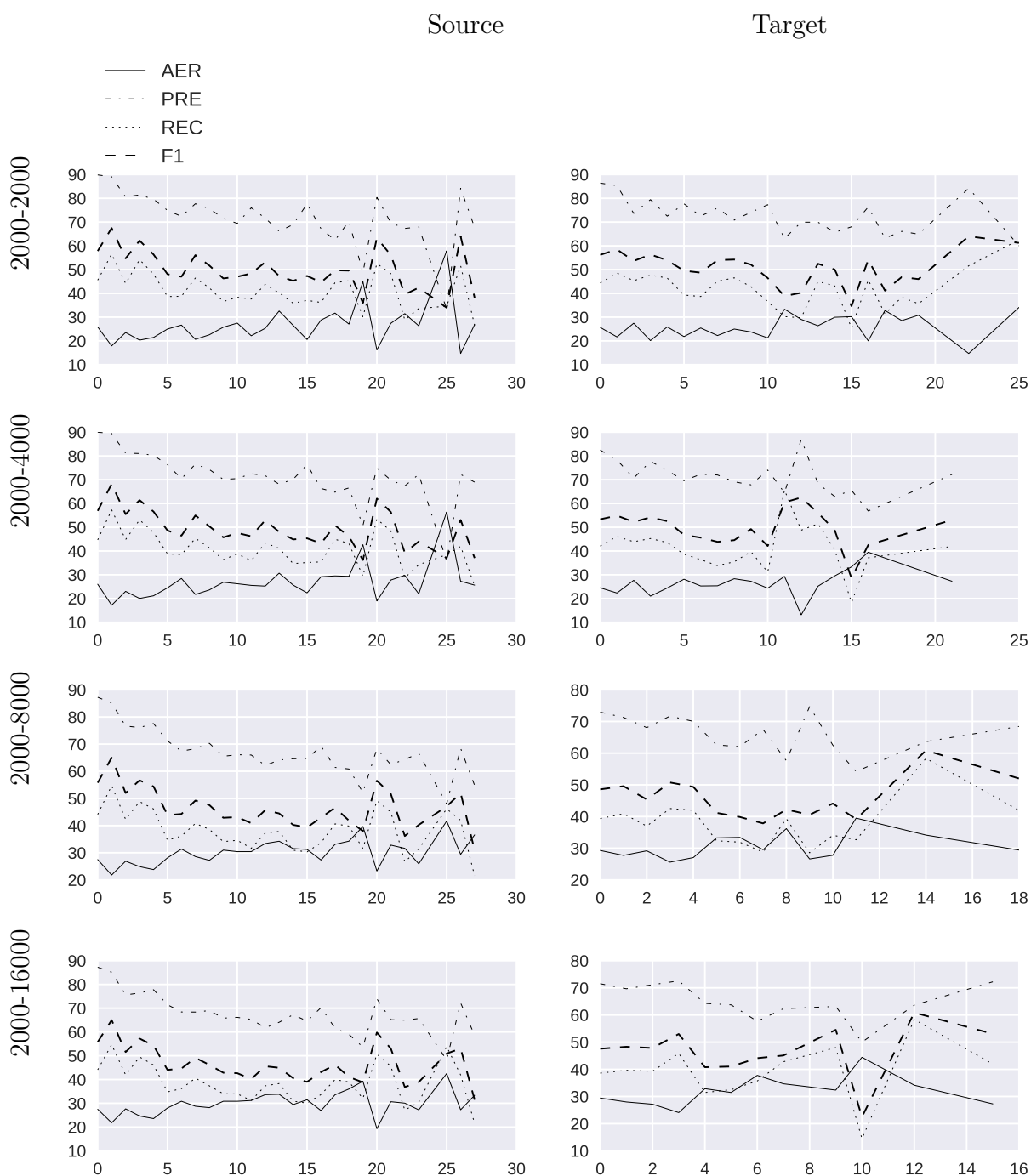
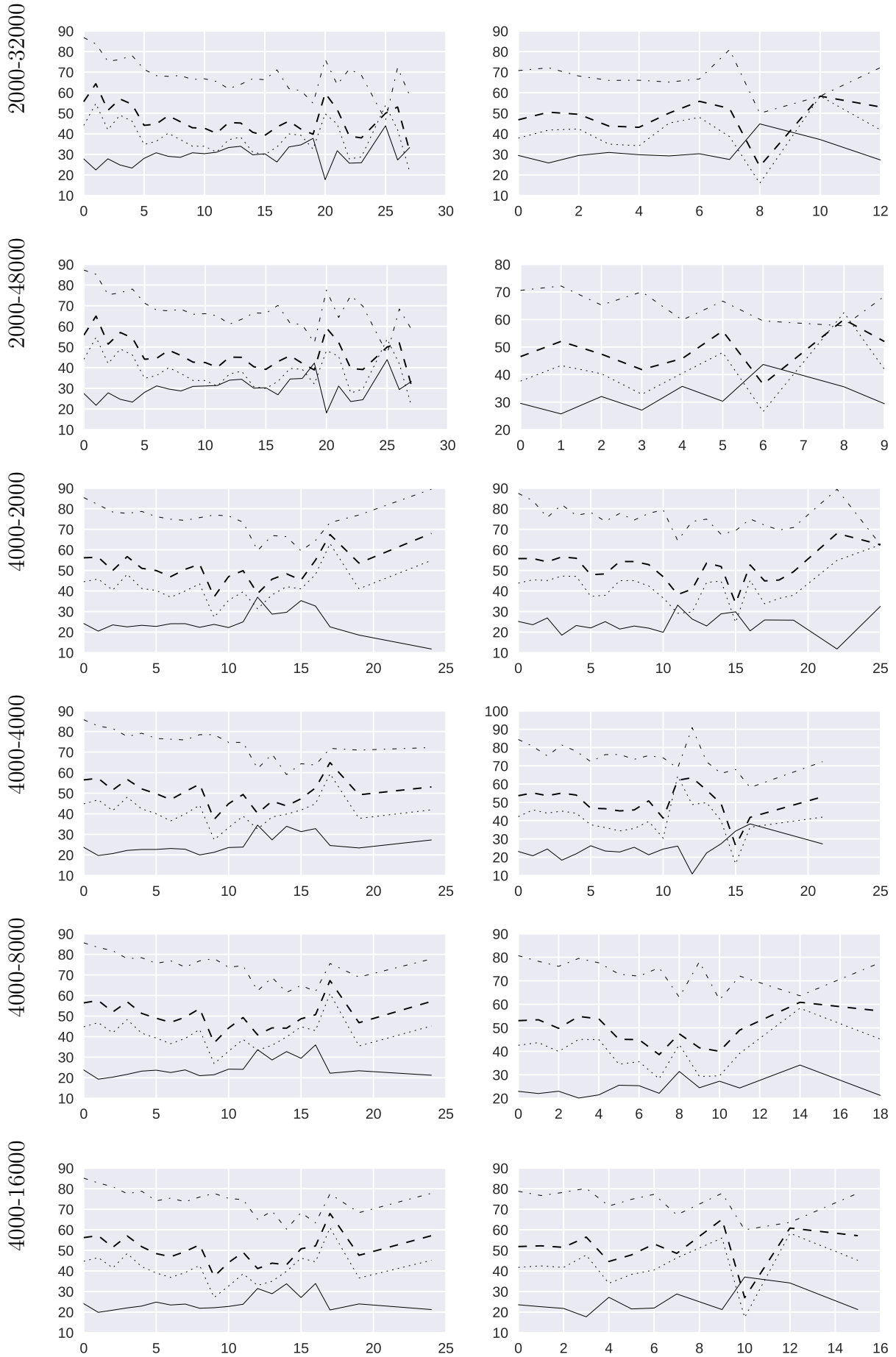
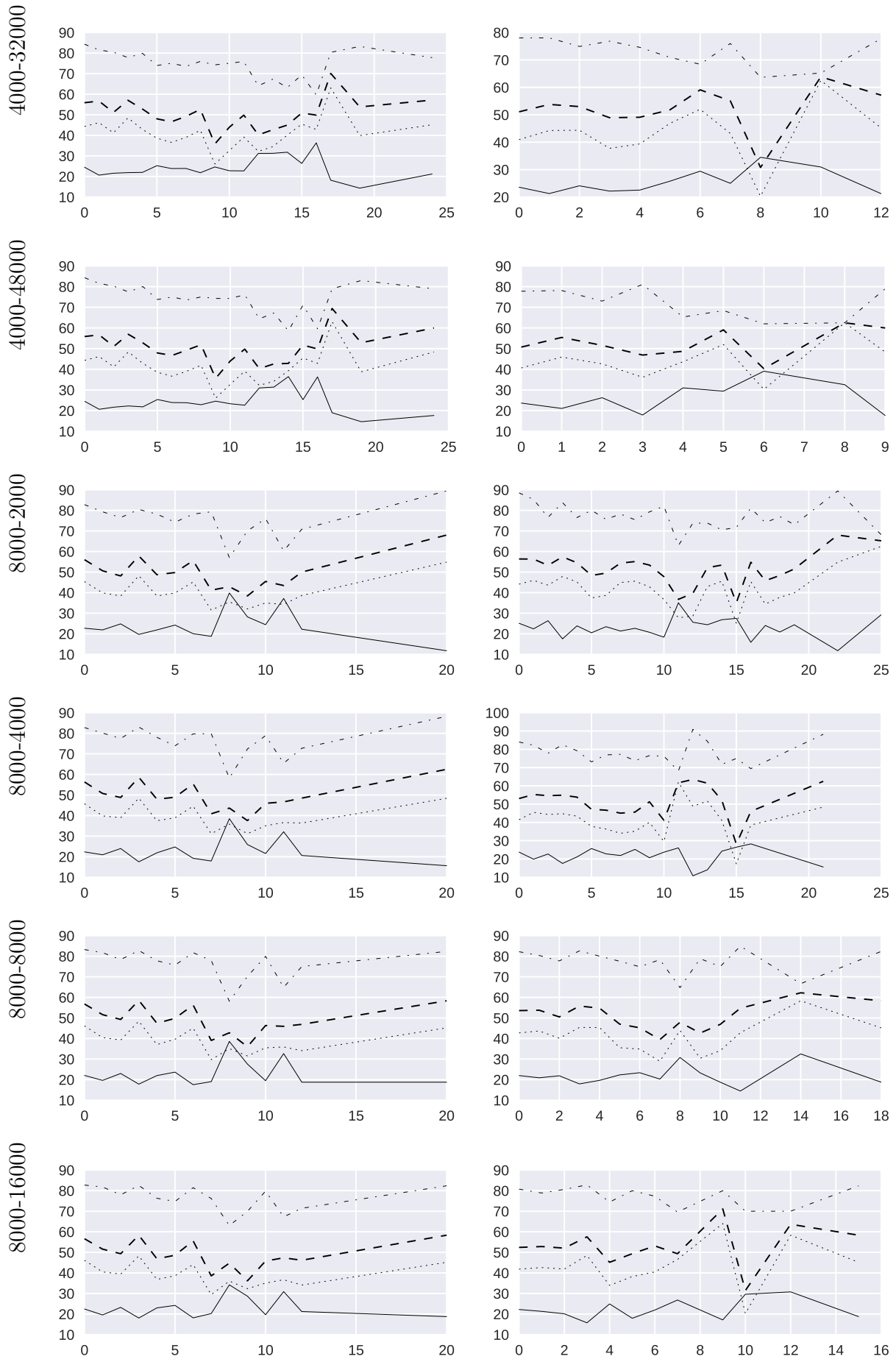
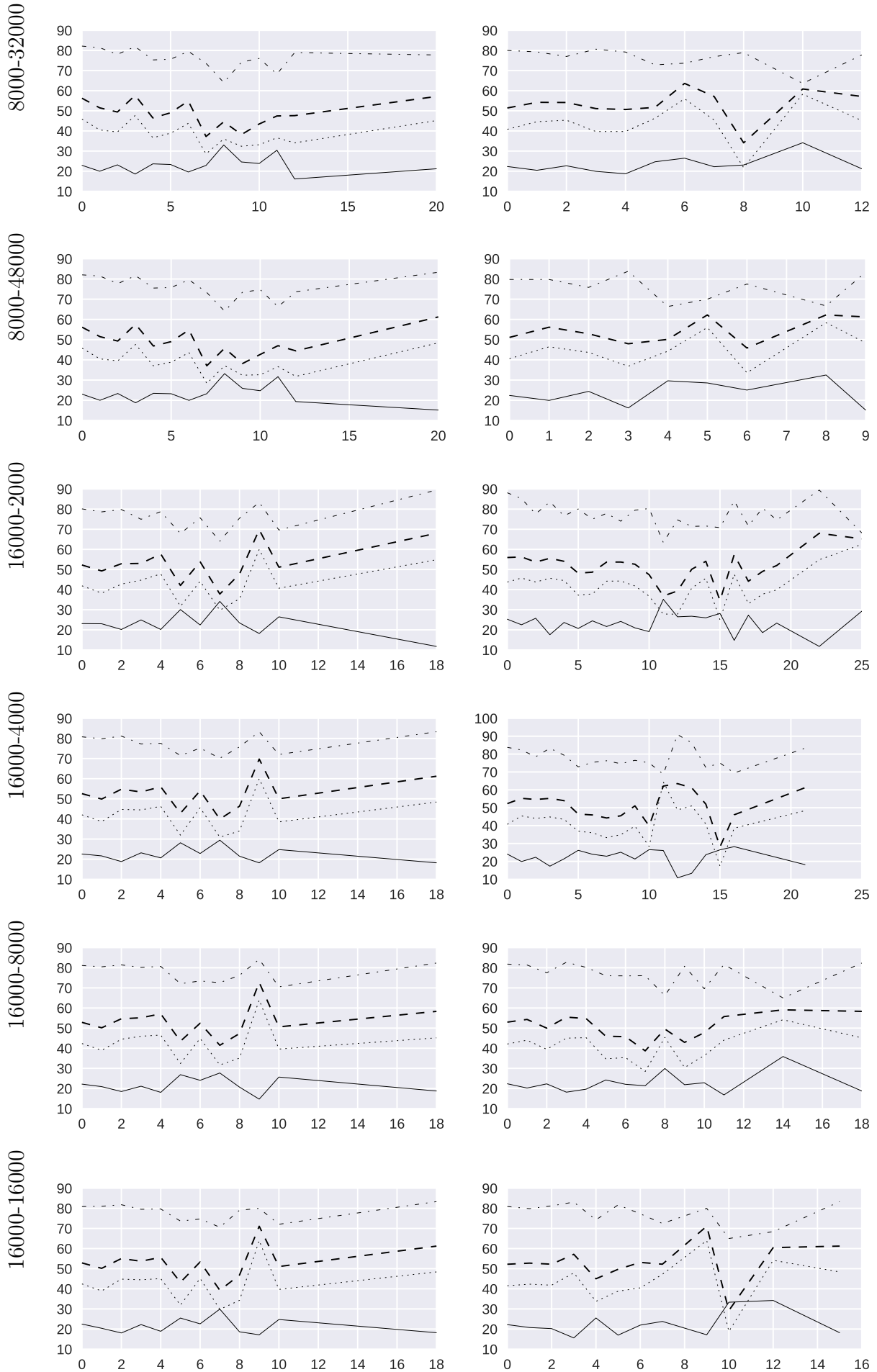


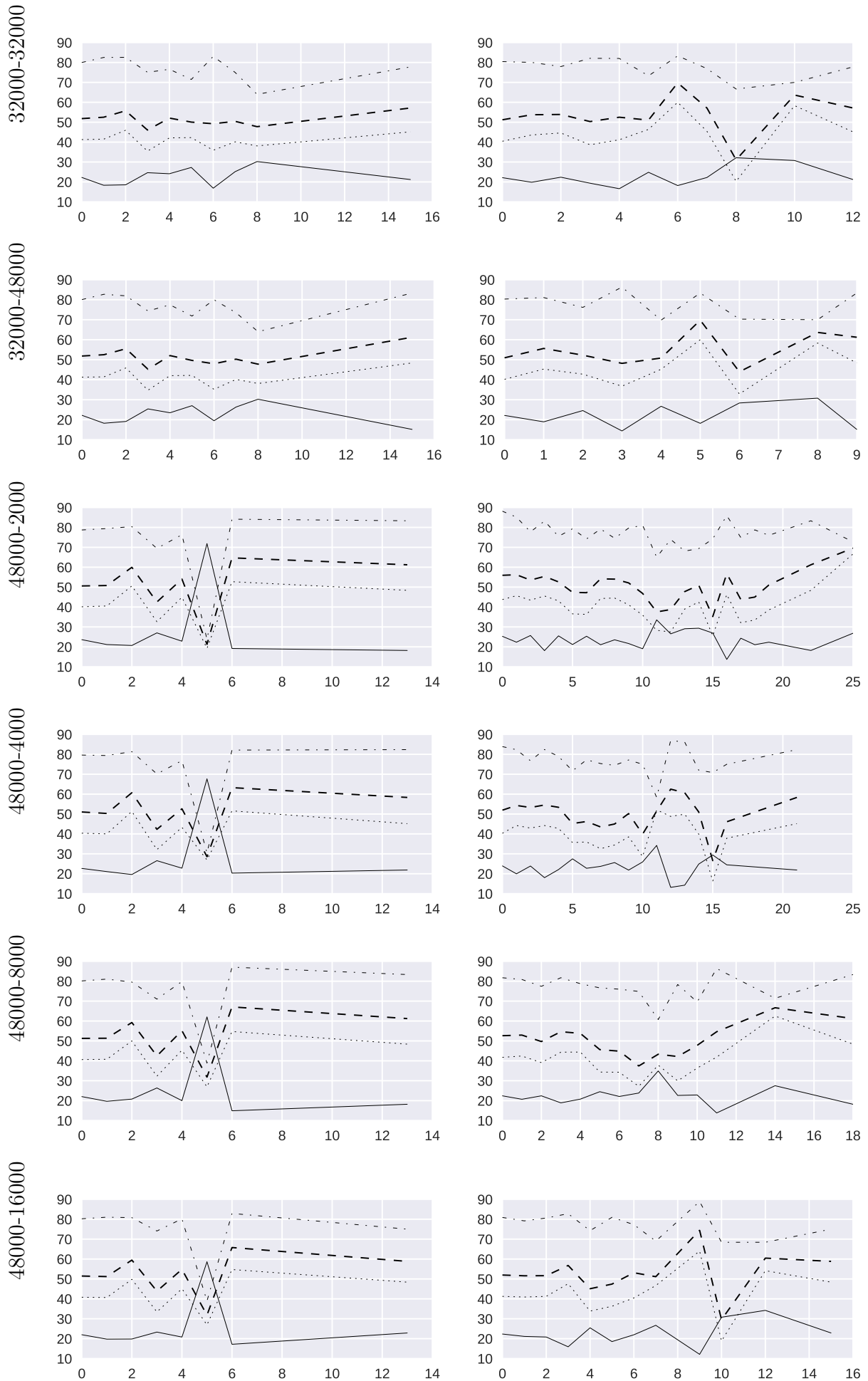
Figure E.139: BPE-based **Fastalign** for the direction English-French: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.











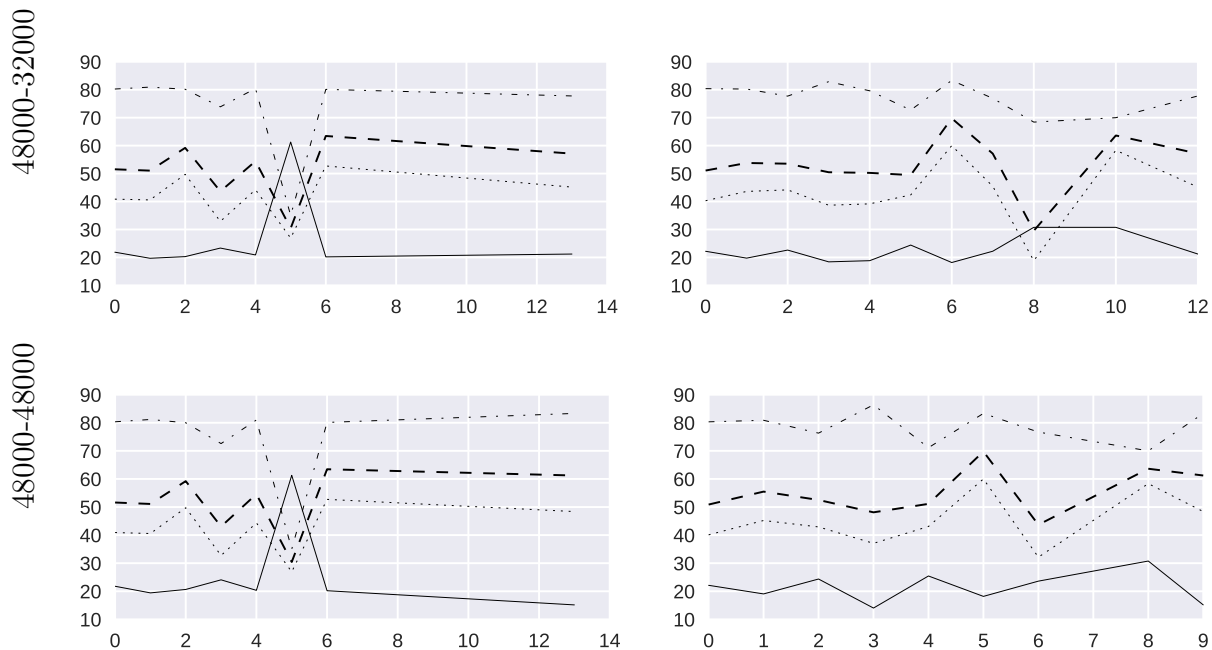
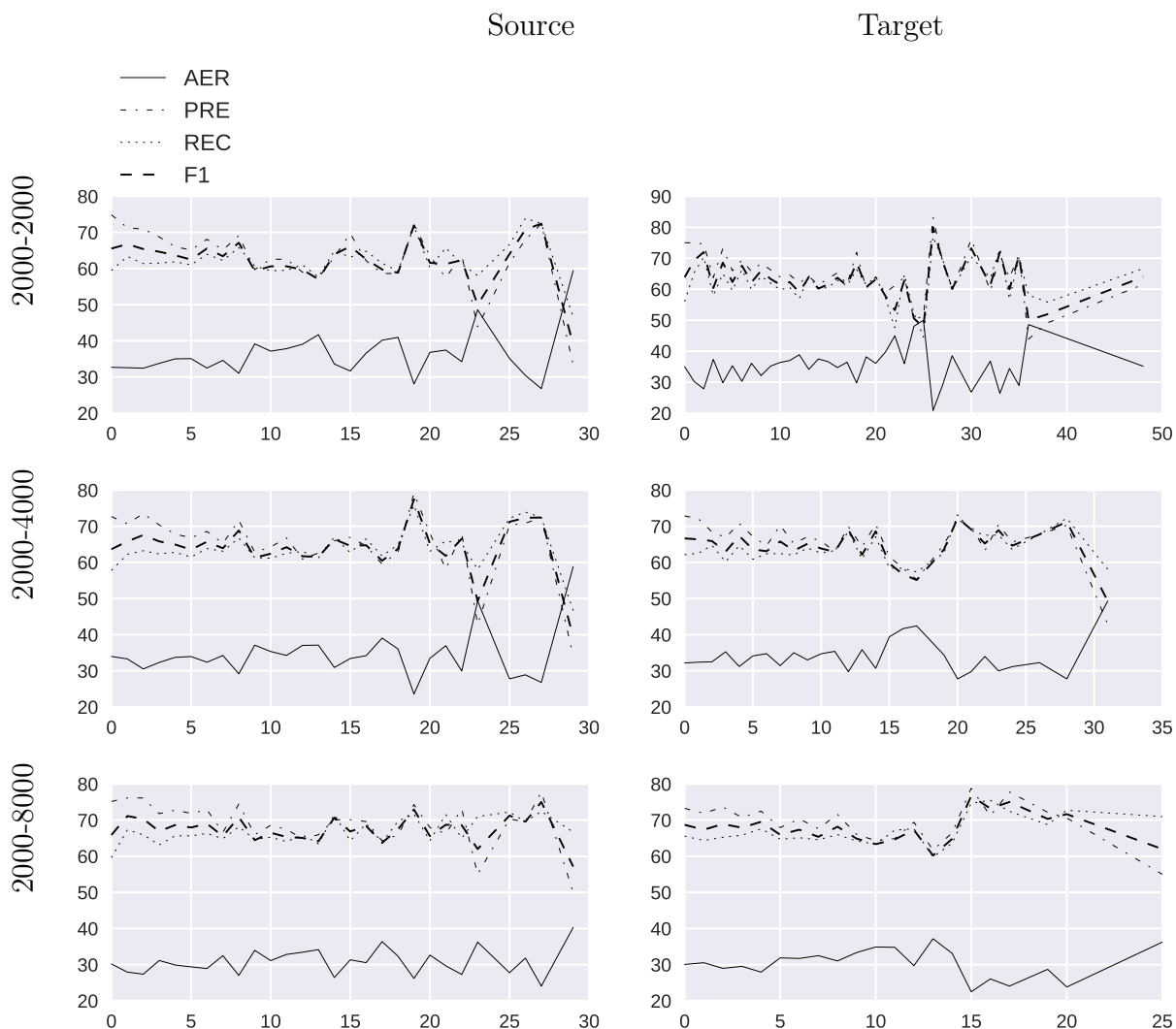
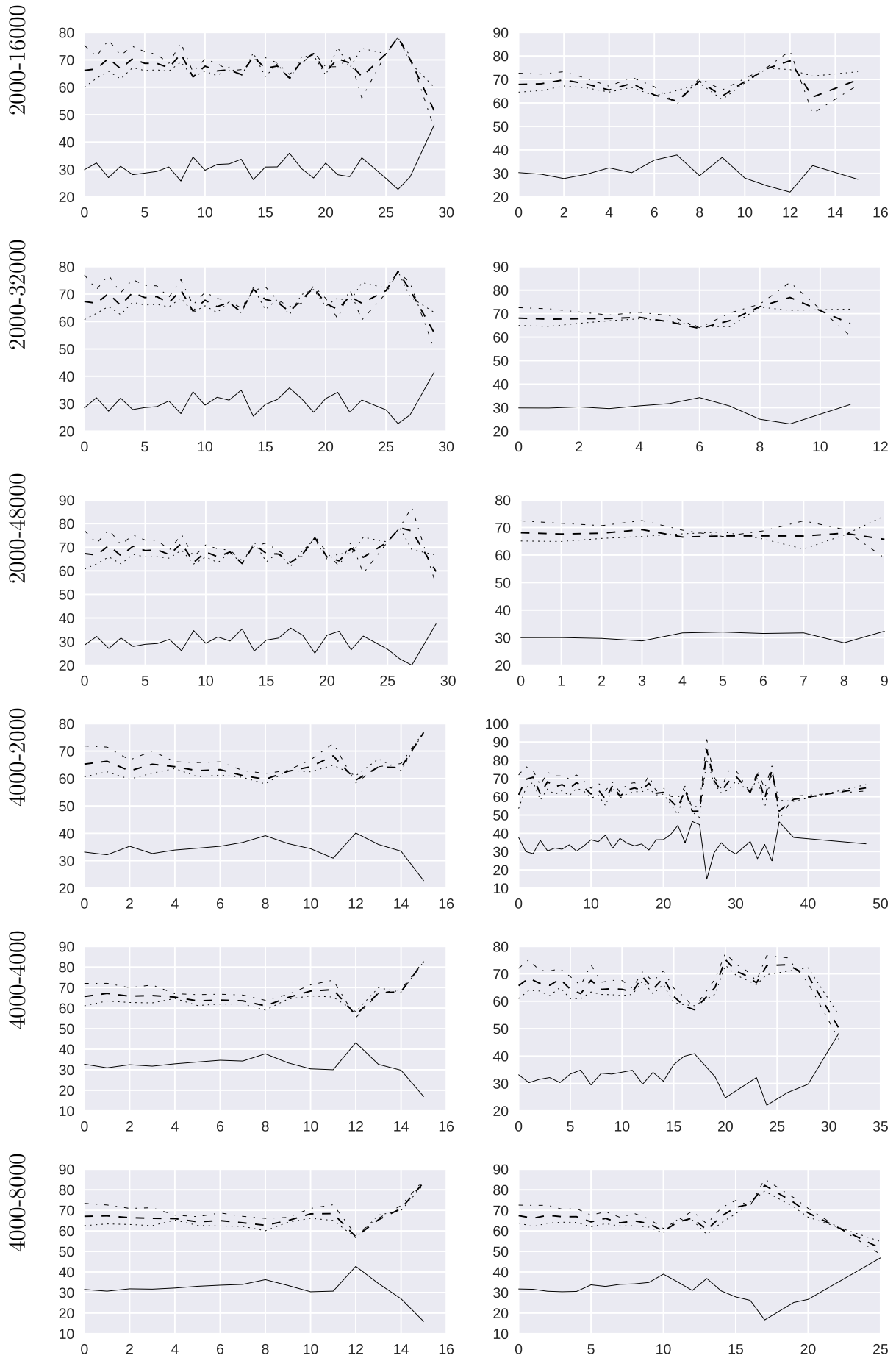
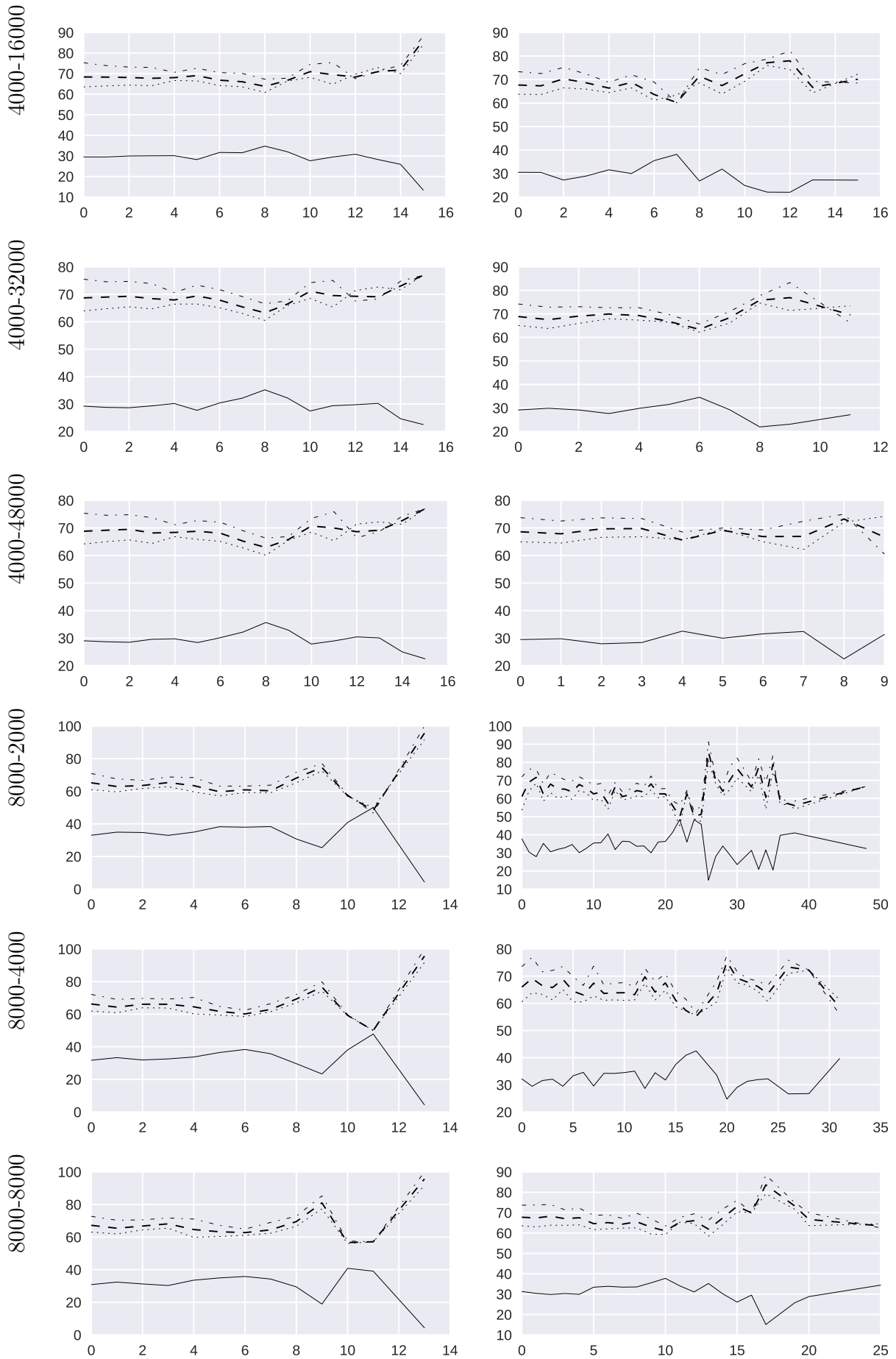
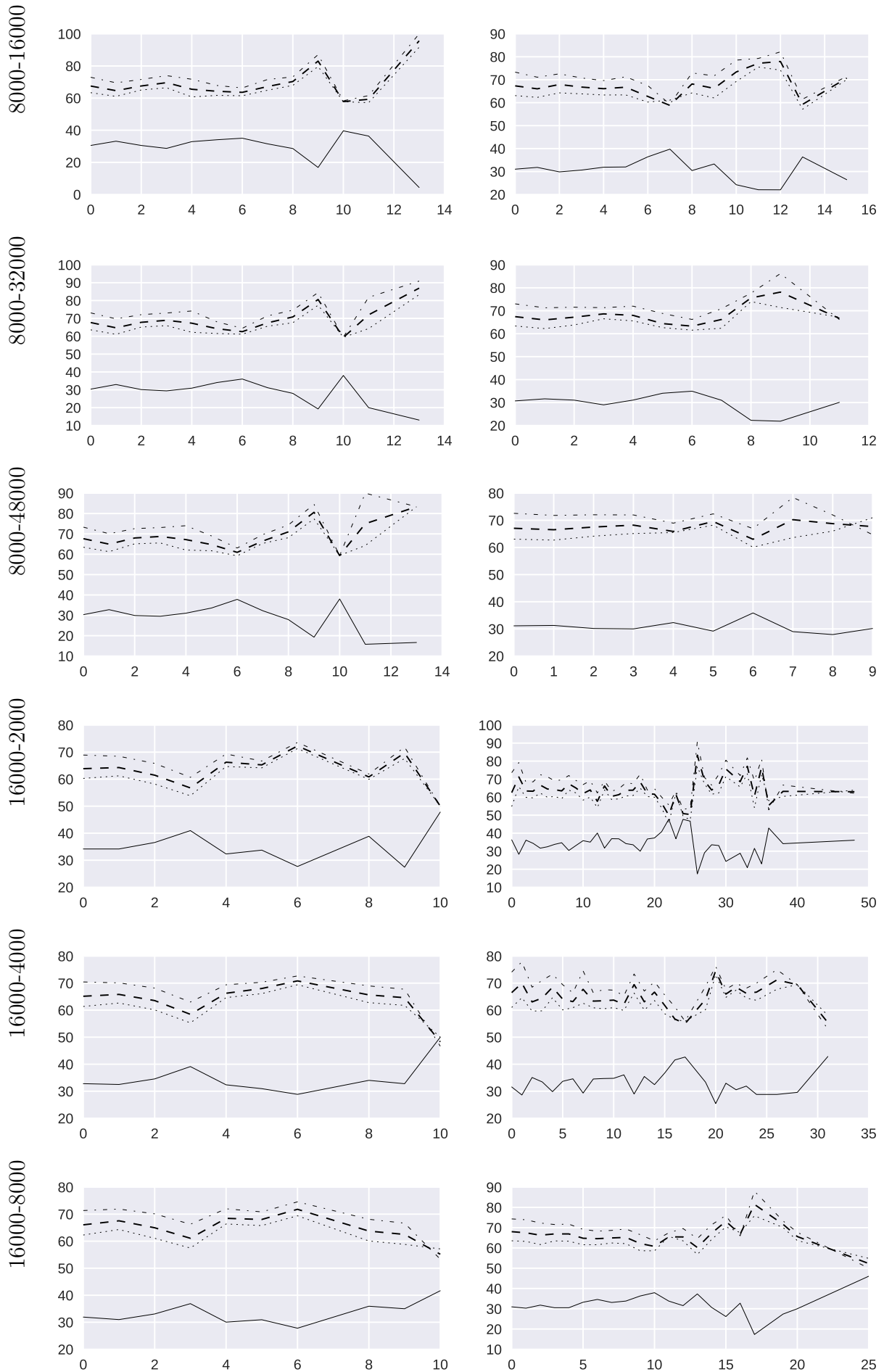


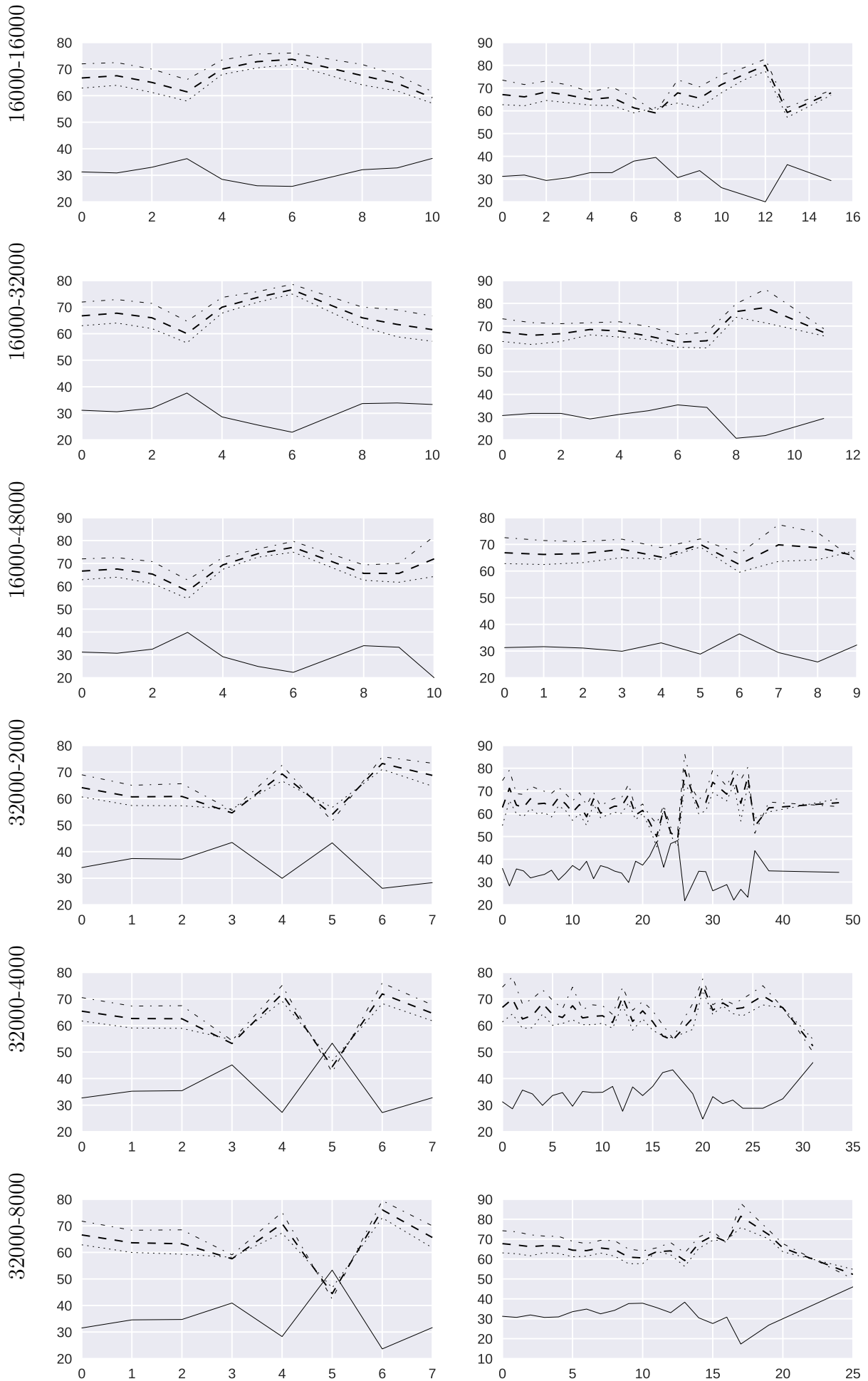
Figure E.140: BPE-based `Fastalign` for the direction French-English: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.

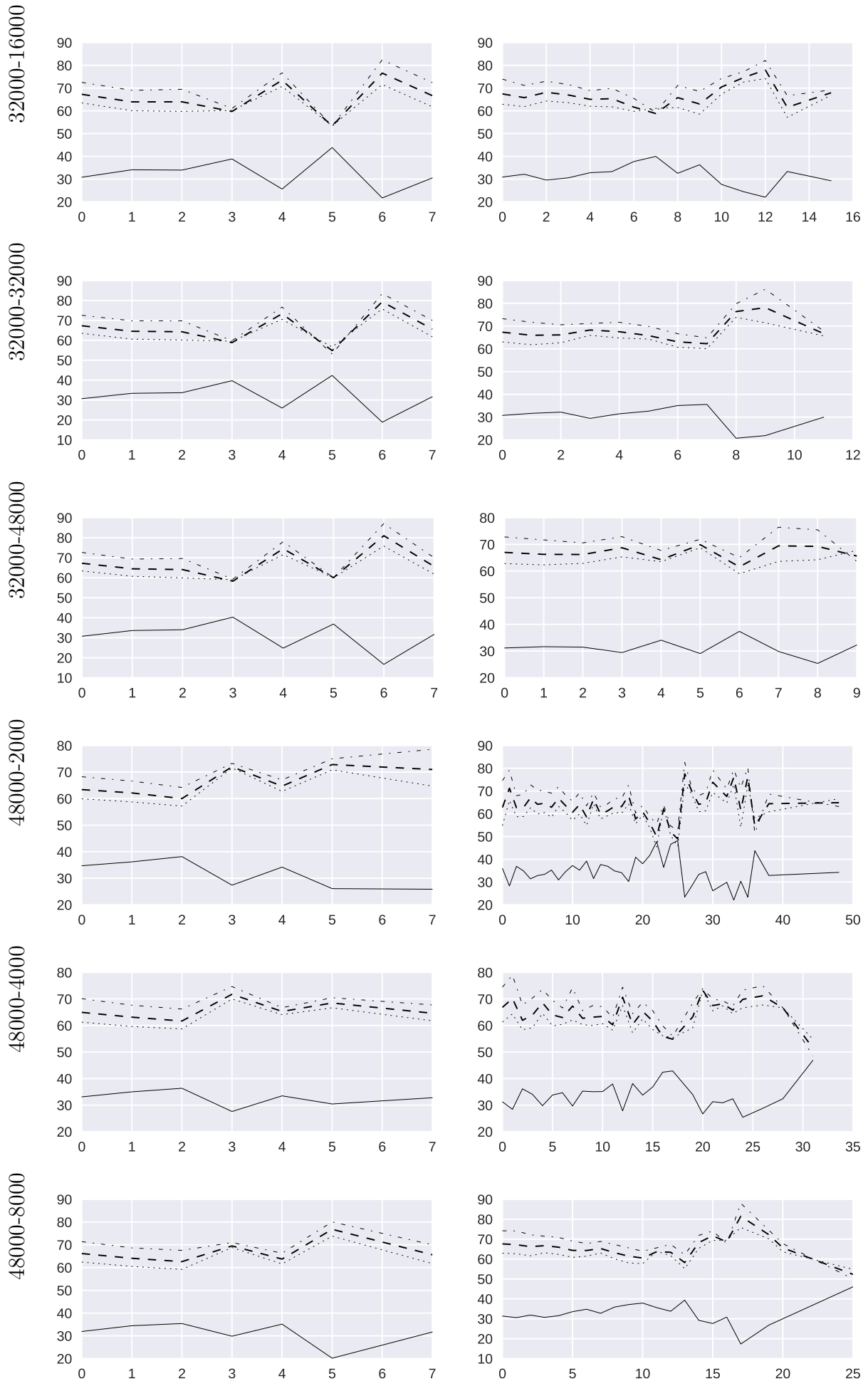












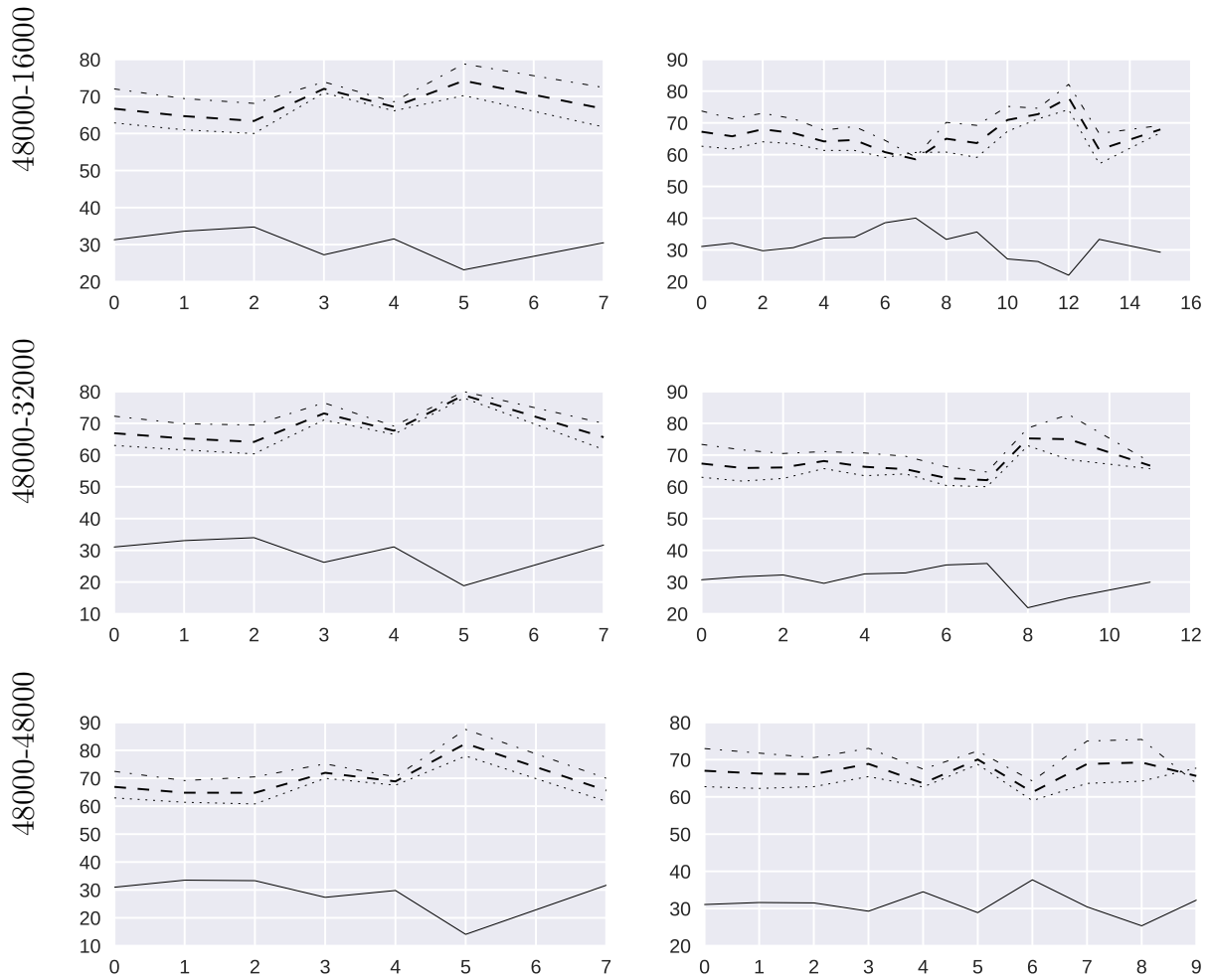
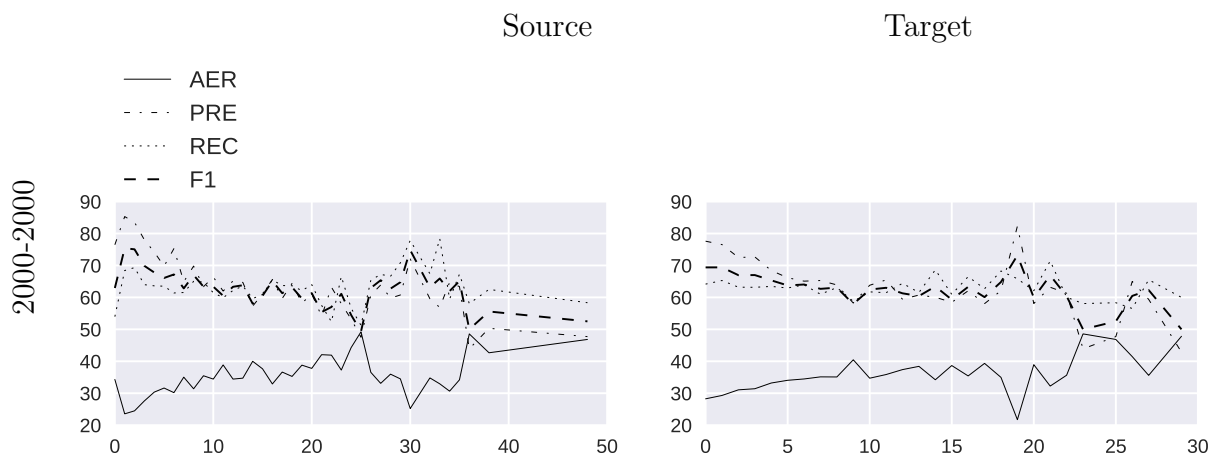
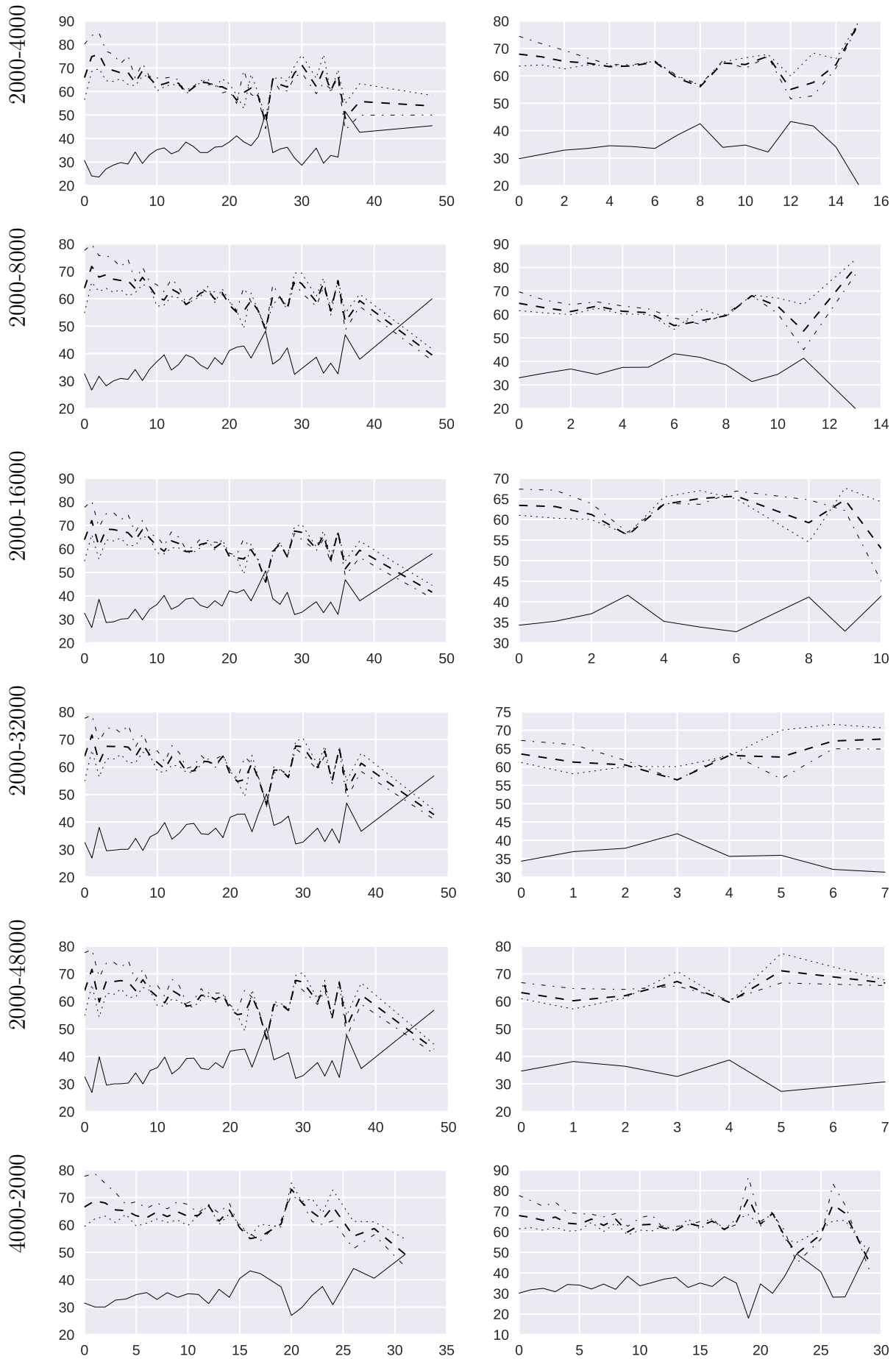
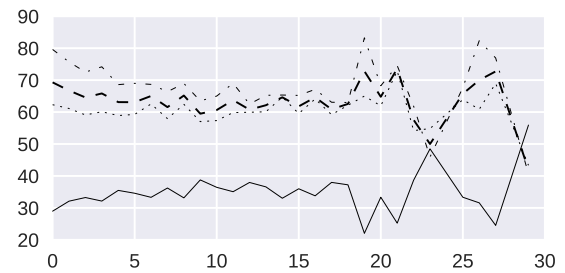
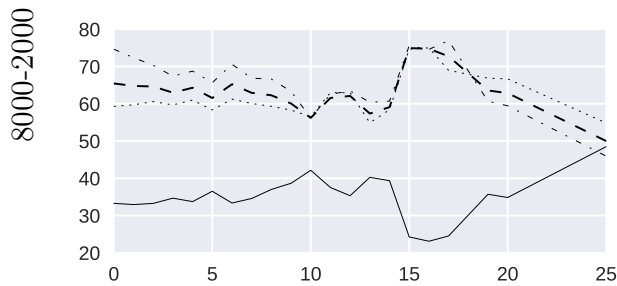
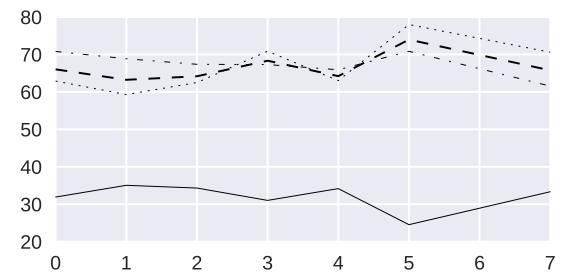
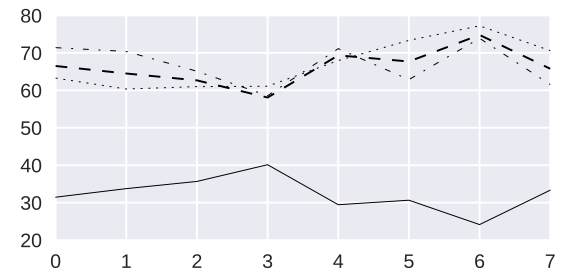
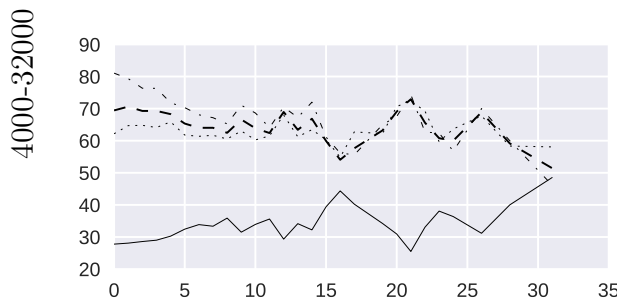
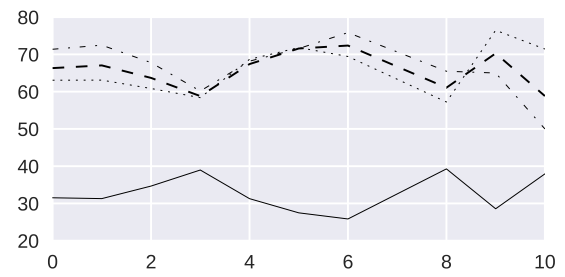
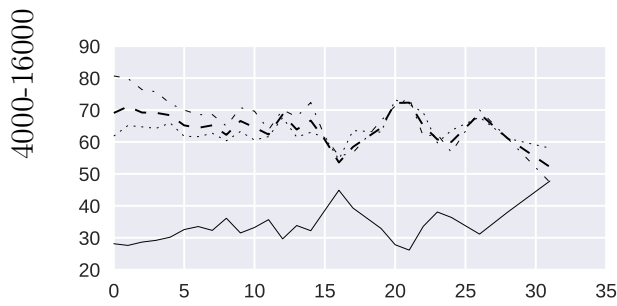
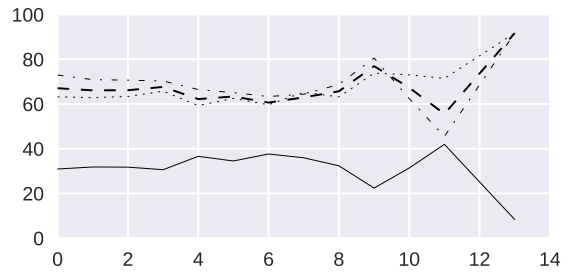
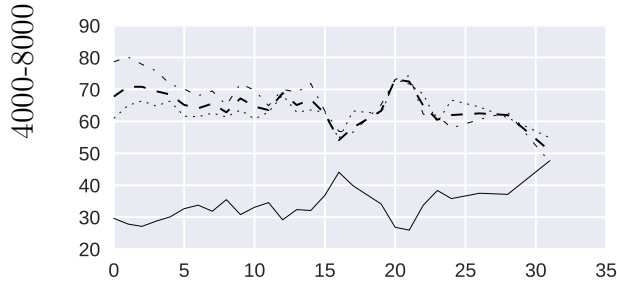
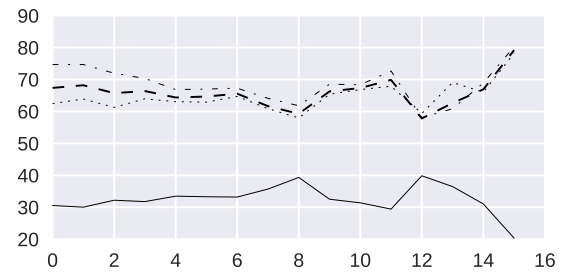
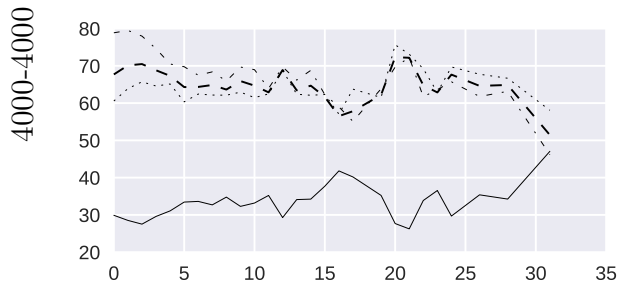
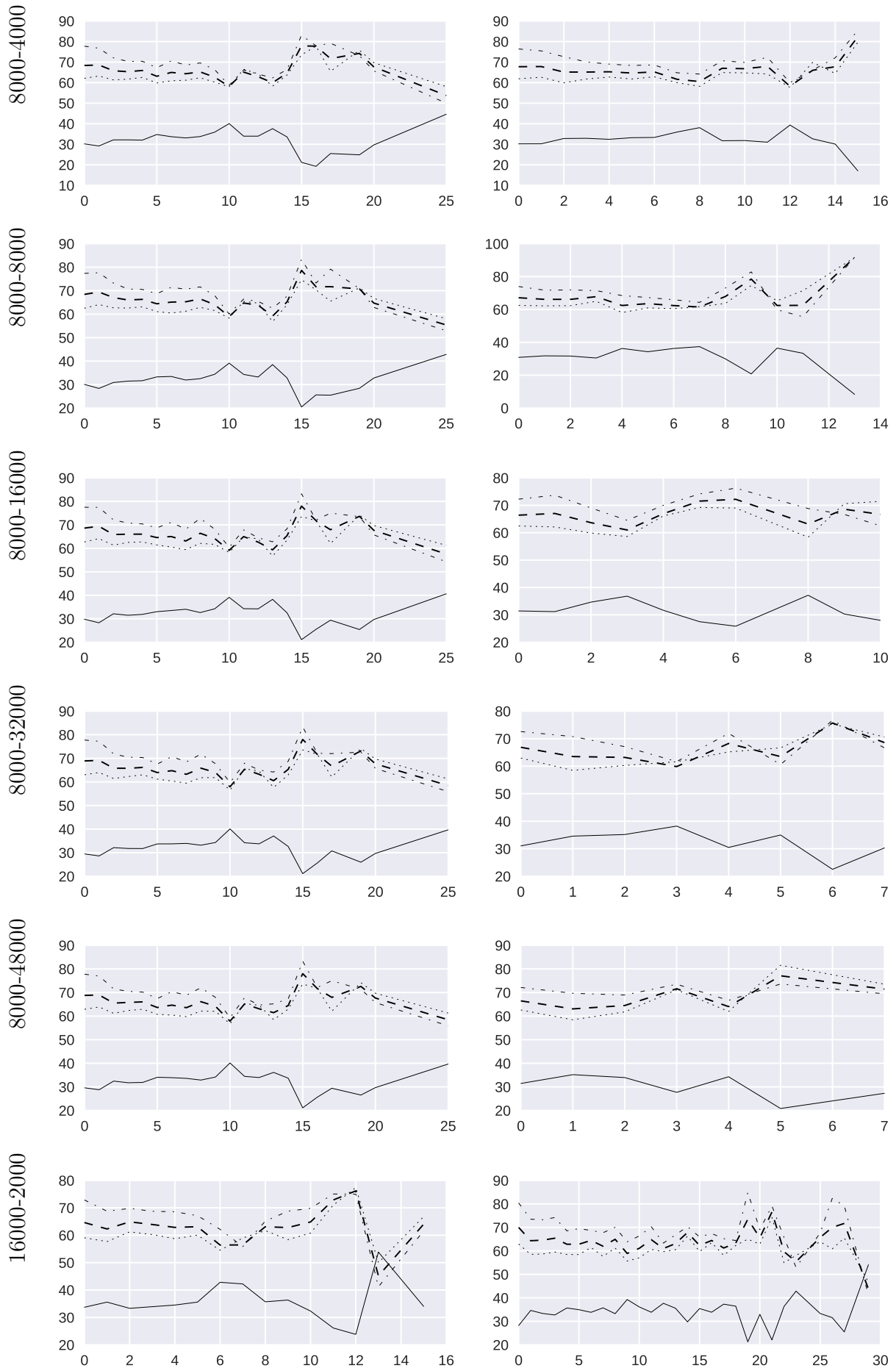


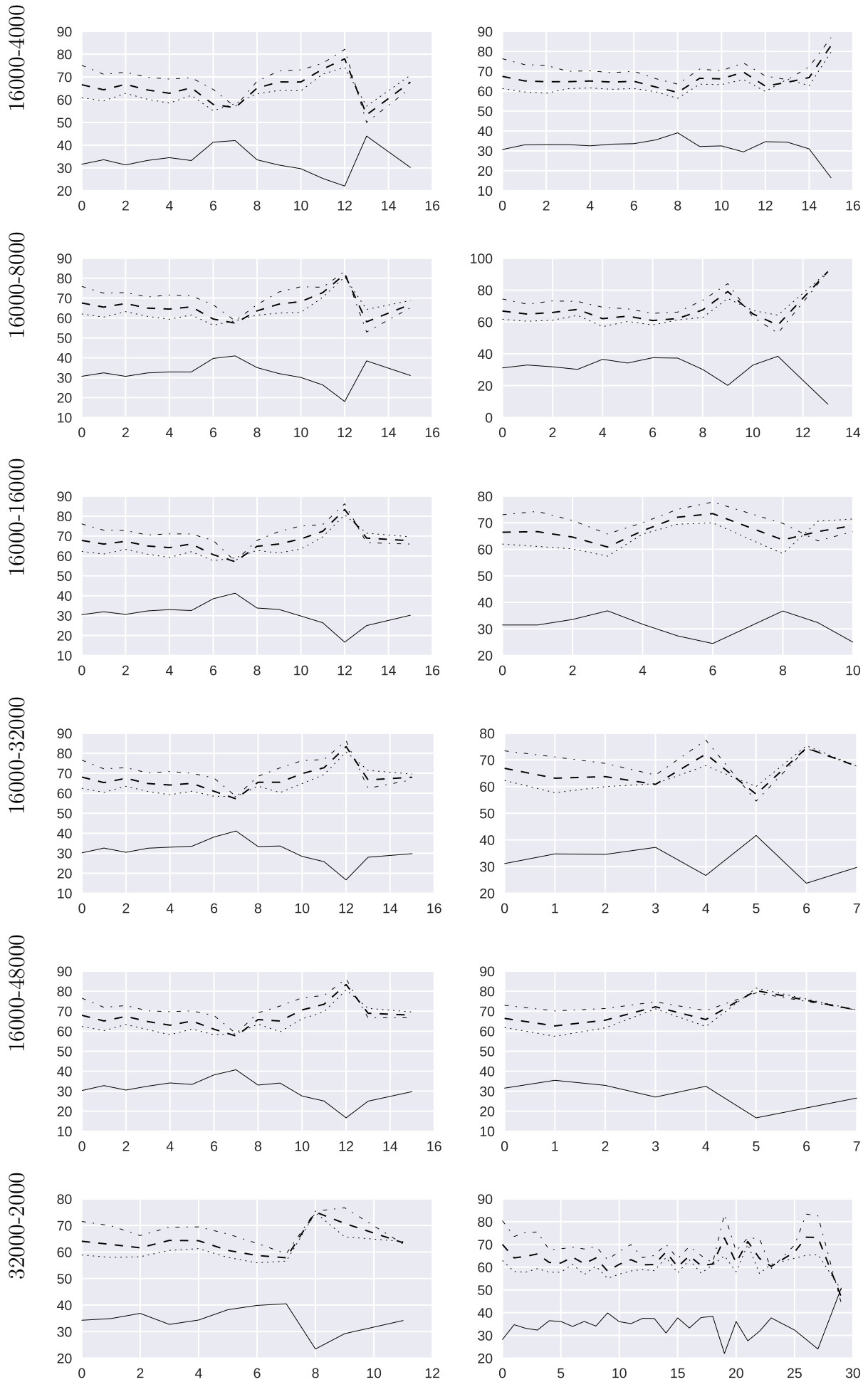
Figure E.141: BPE-based *Fastalign* for the direction English-German: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.

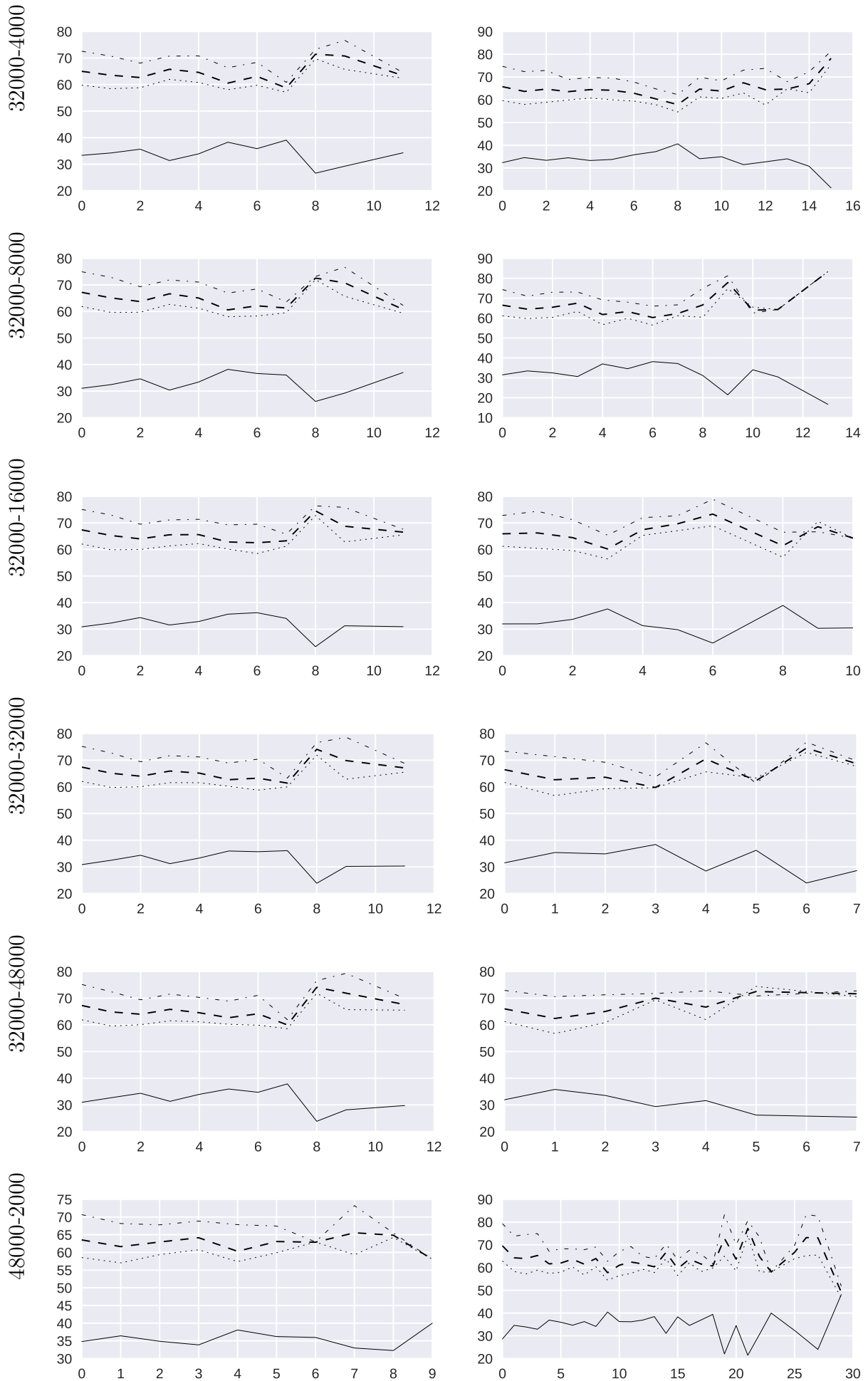












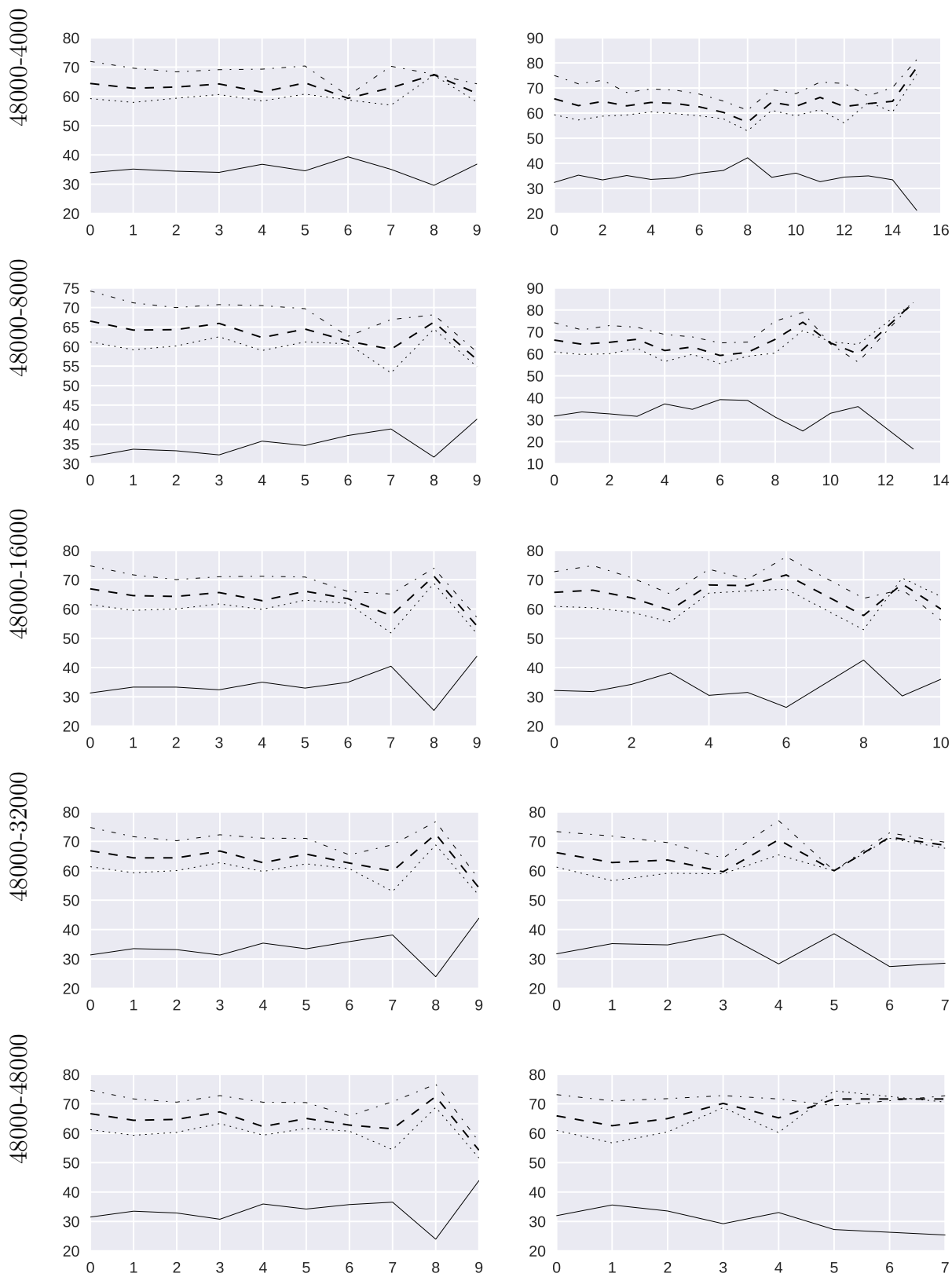
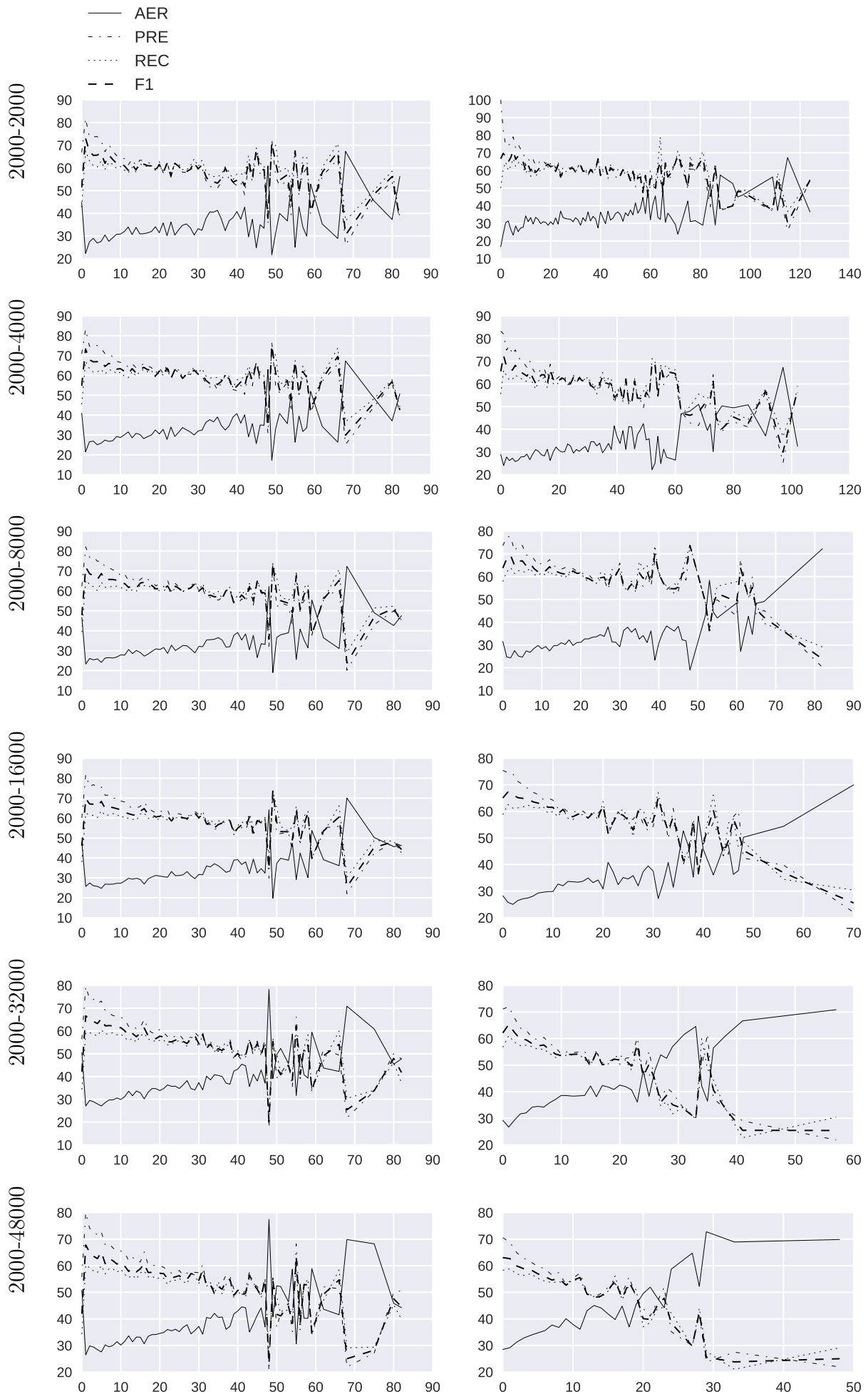
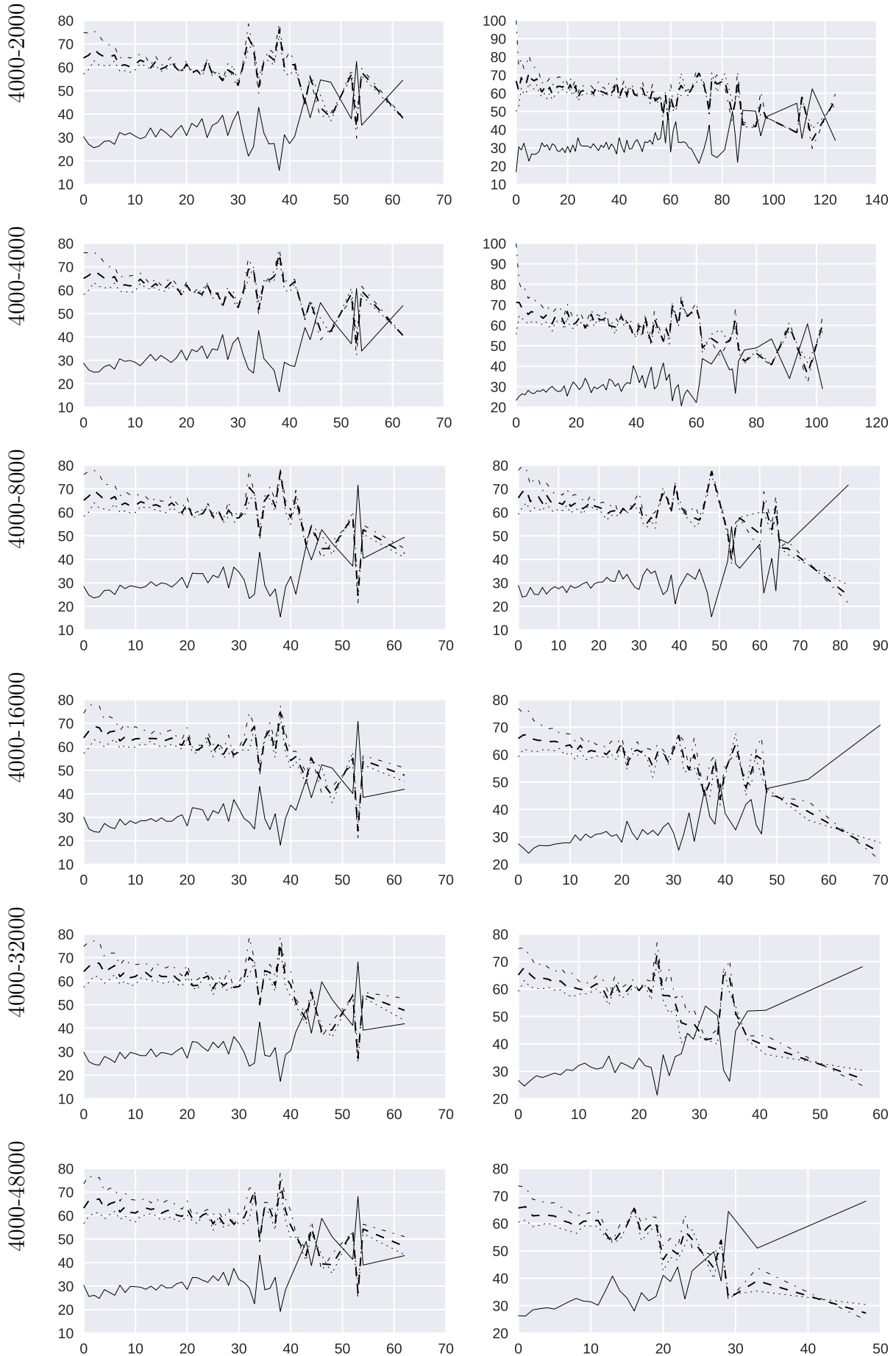


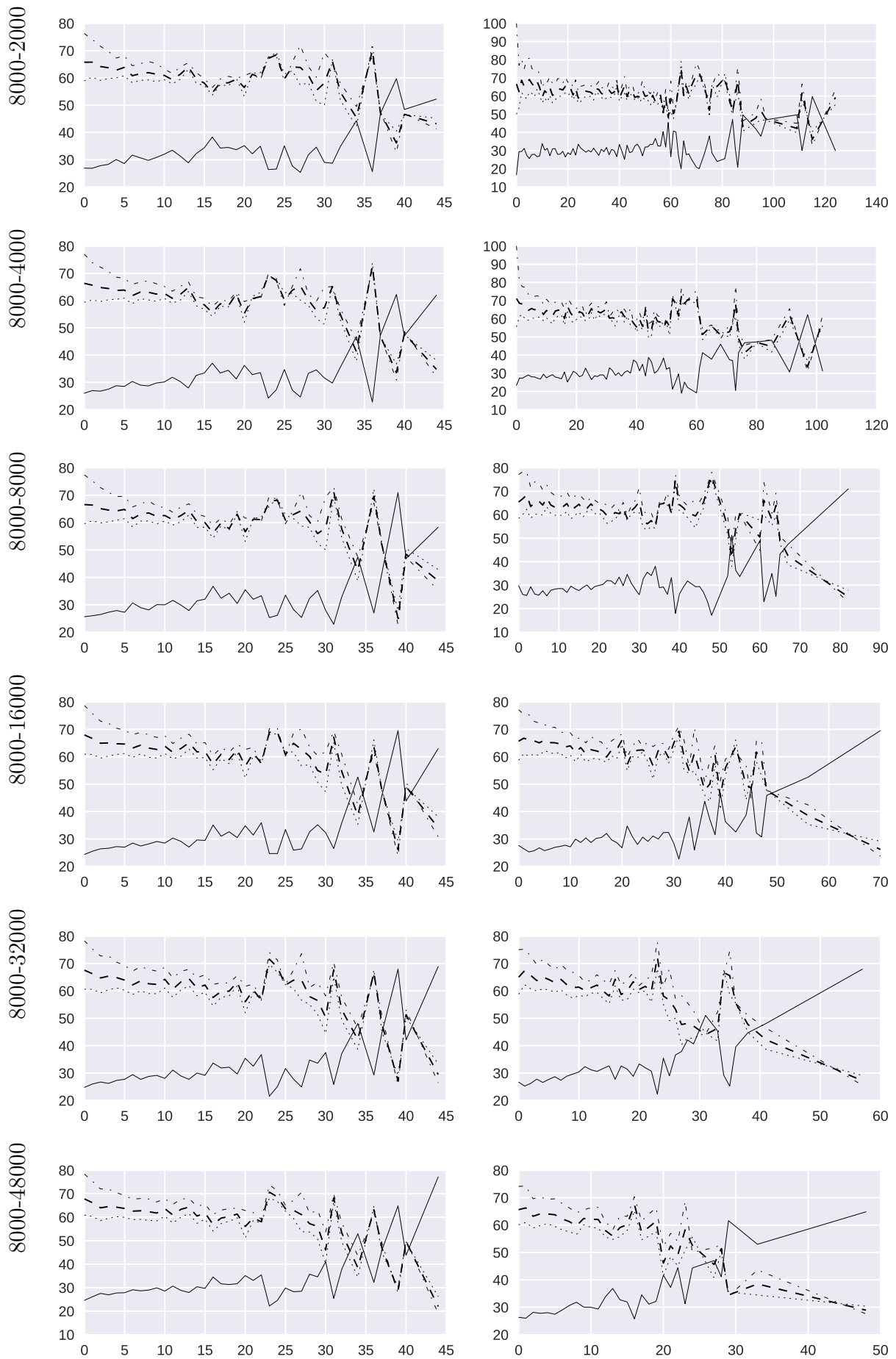
Figure E.142: BPE-based Fastalign for the direction German-English: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.

Source

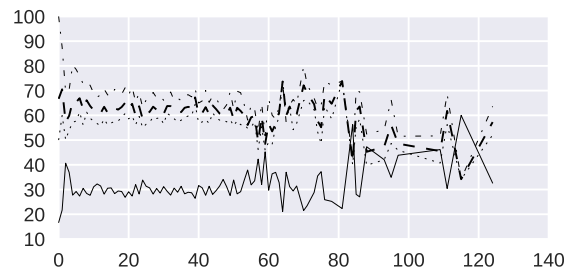
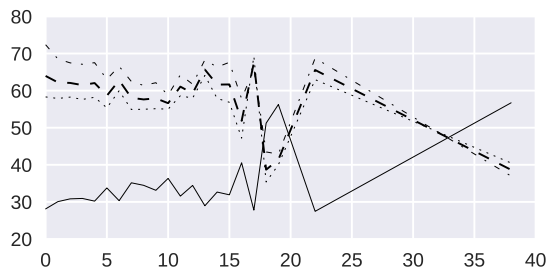
Target



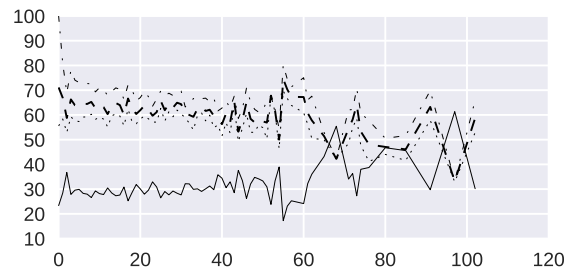
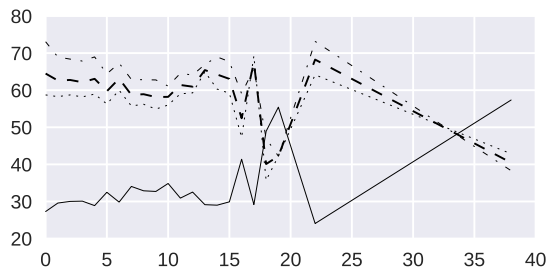




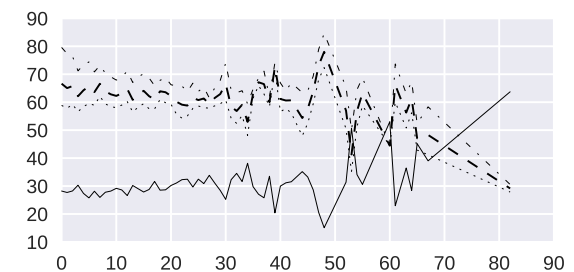
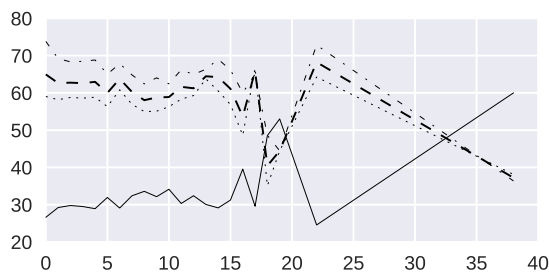
32000-2000



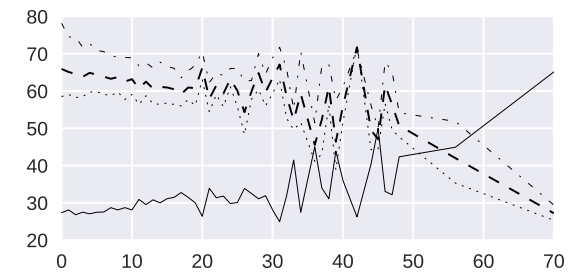
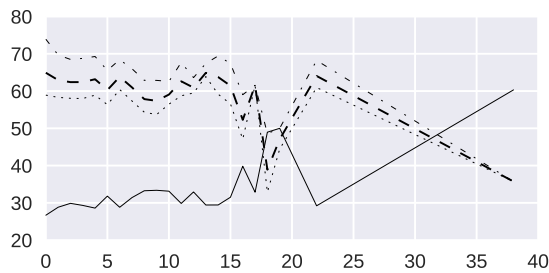
32000-4000



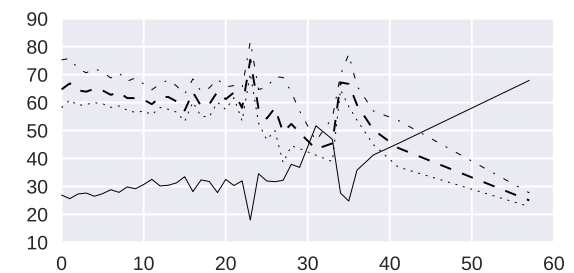
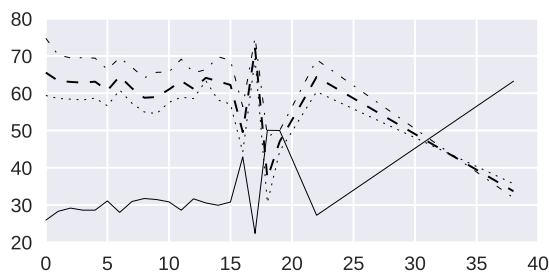
32000-8000



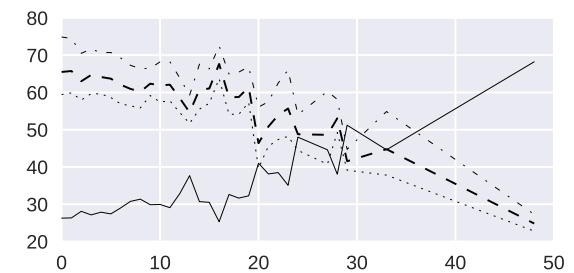
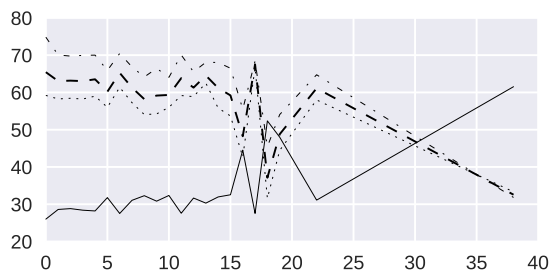
32000-16000



32000-32000



32000-48000



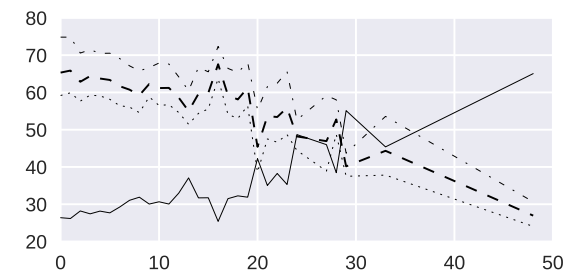
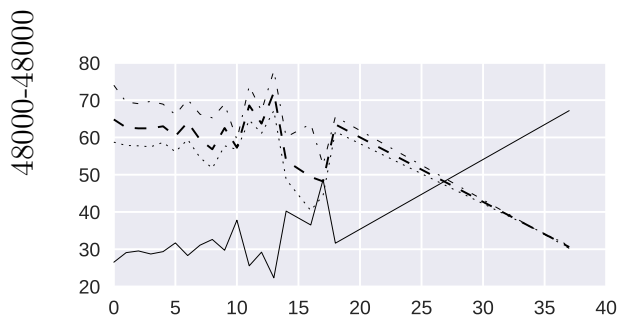
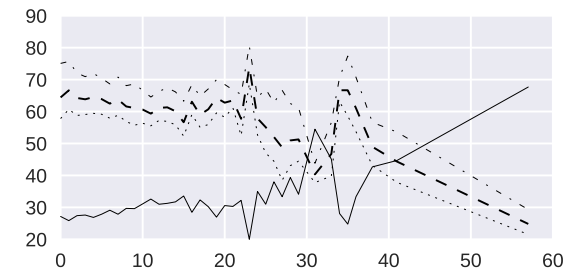
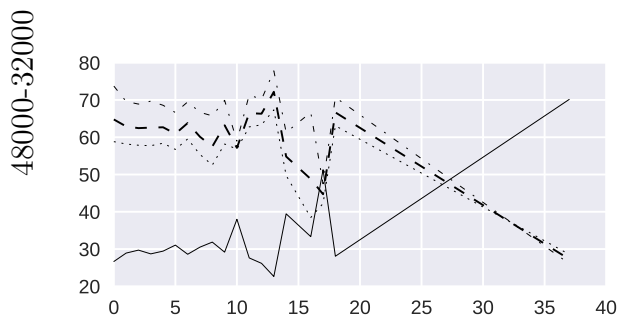
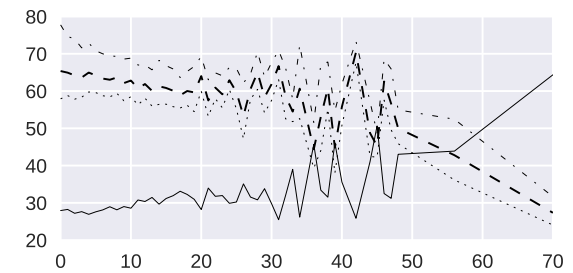
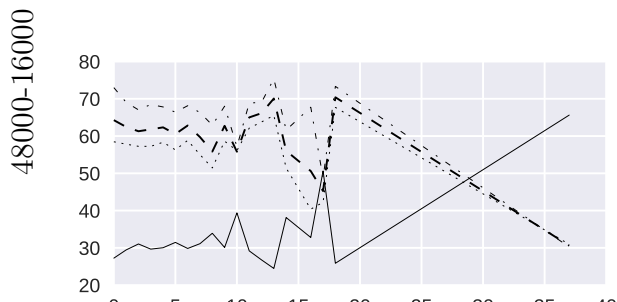
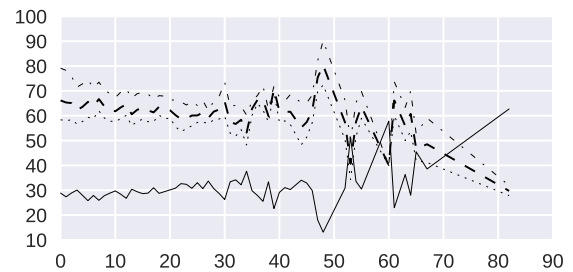
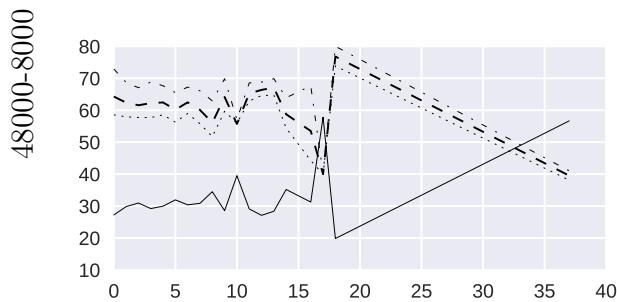
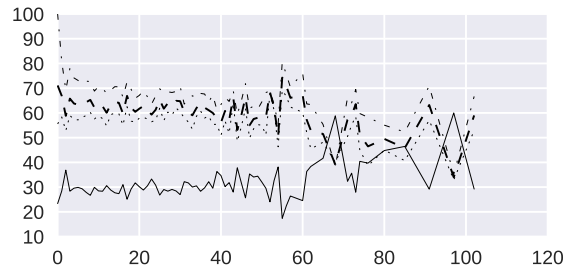
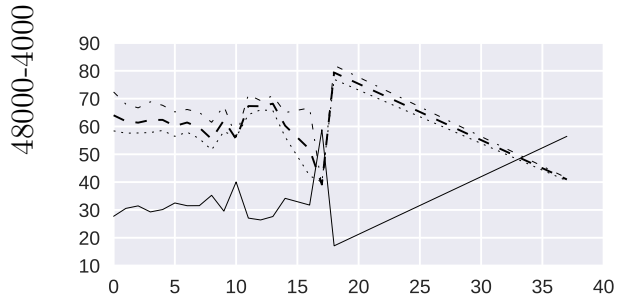
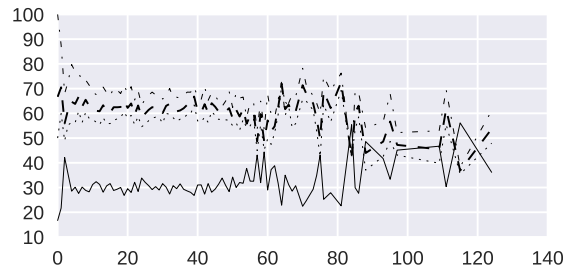
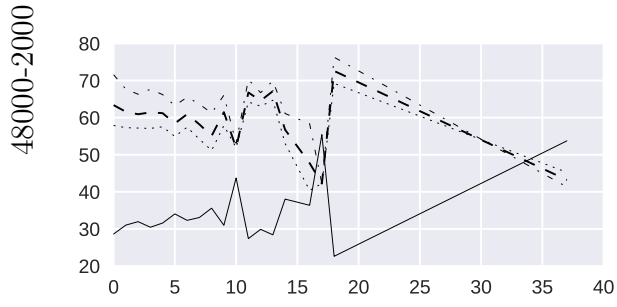
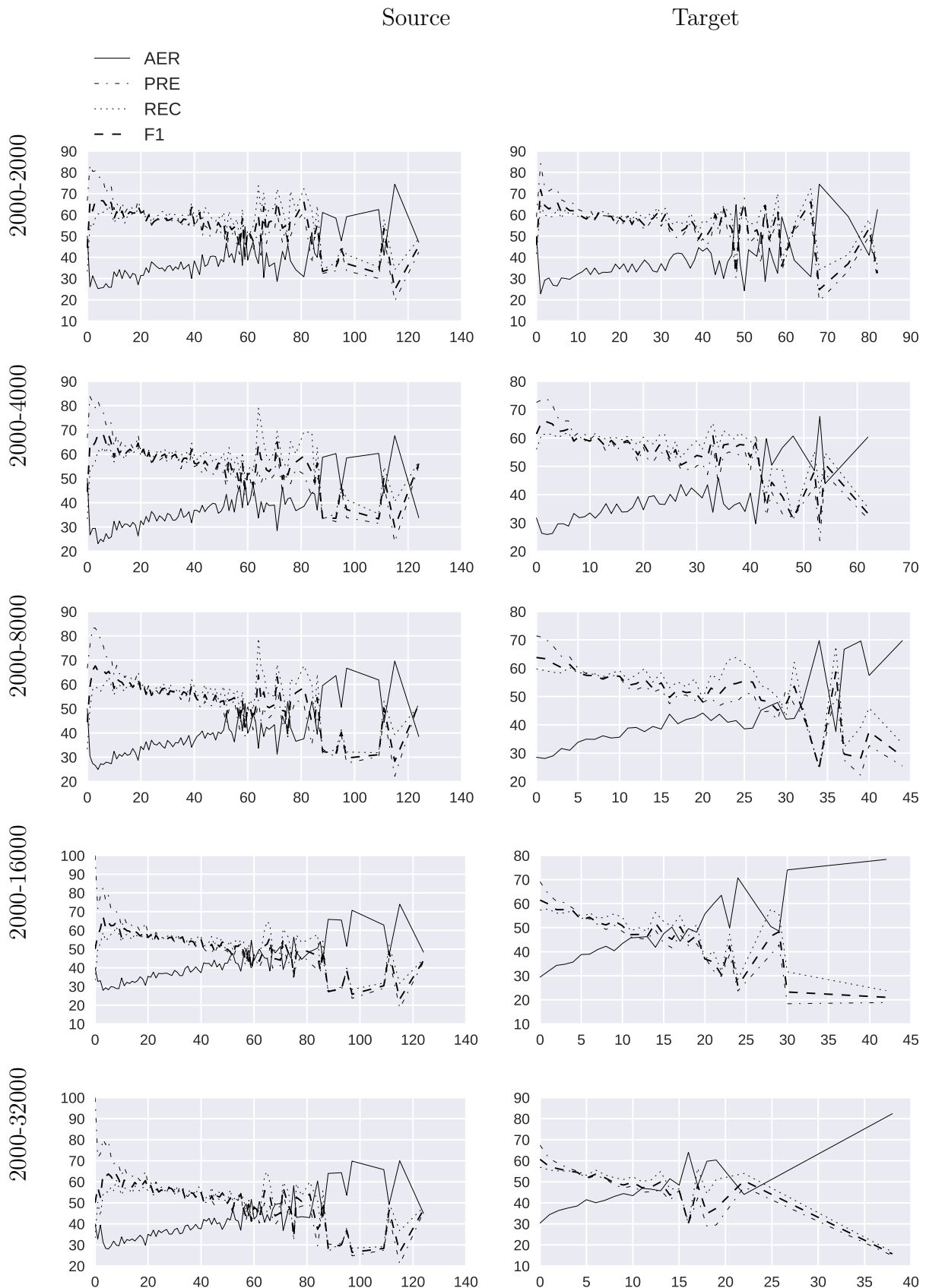
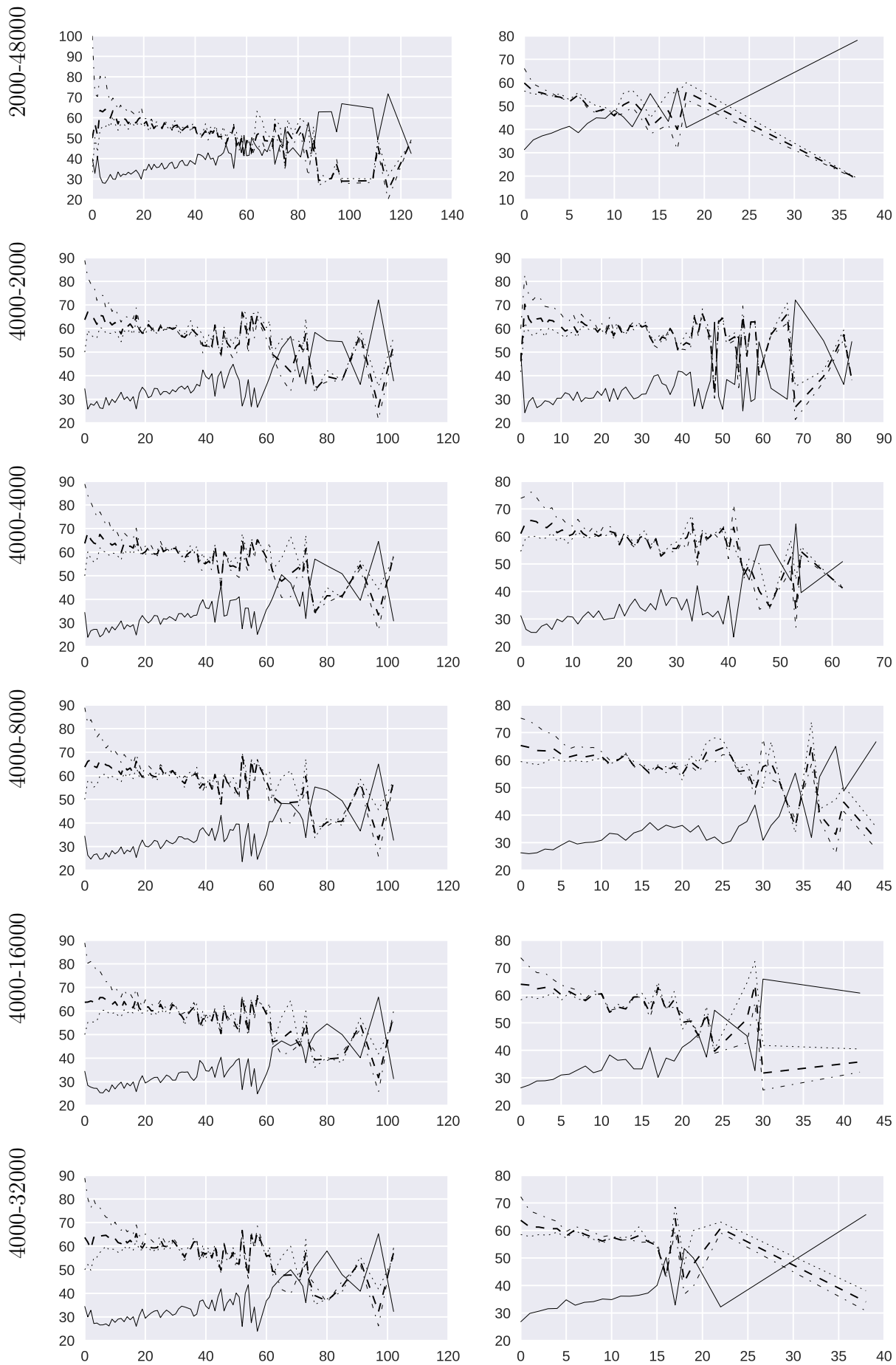
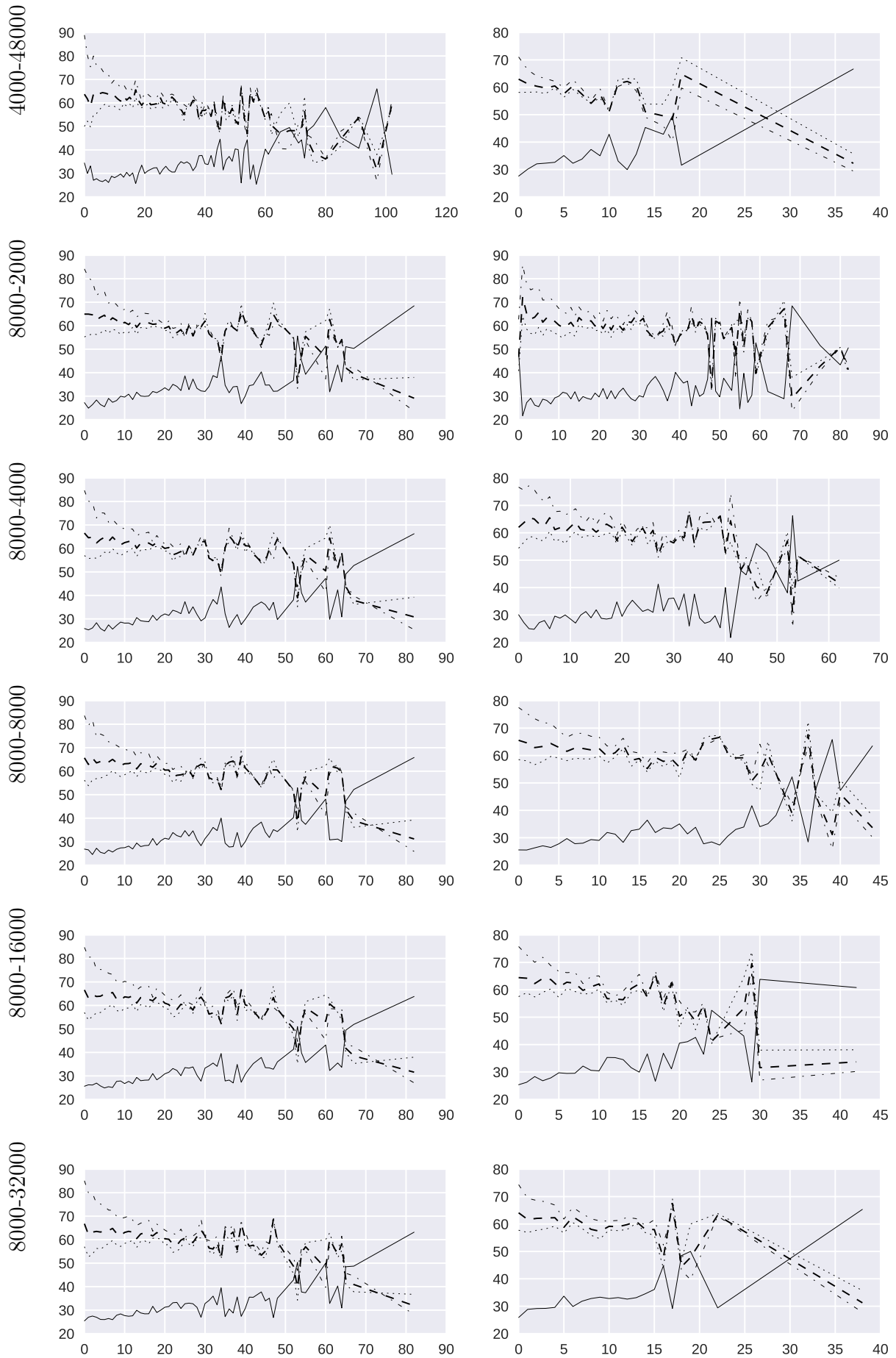
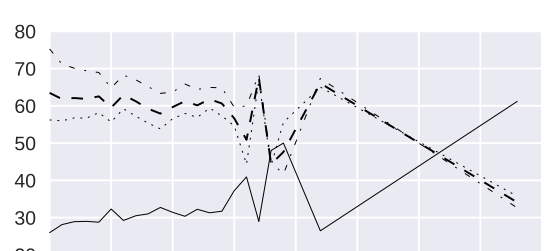
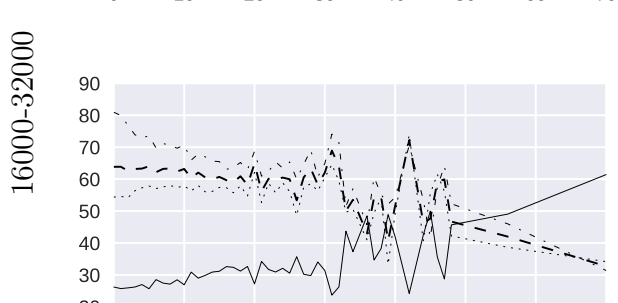
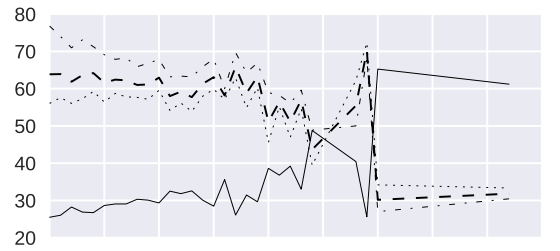
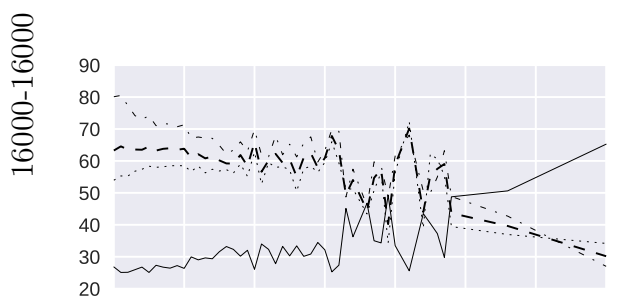
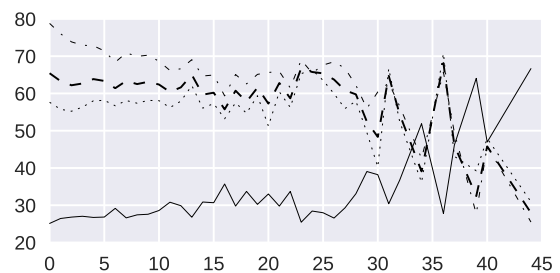
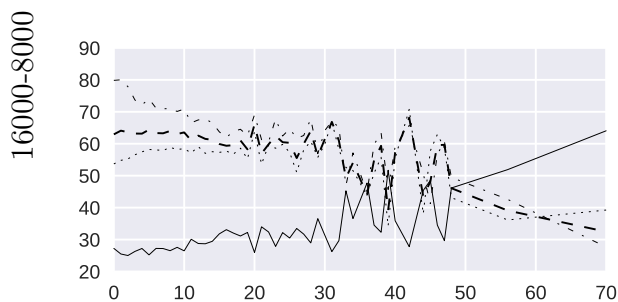
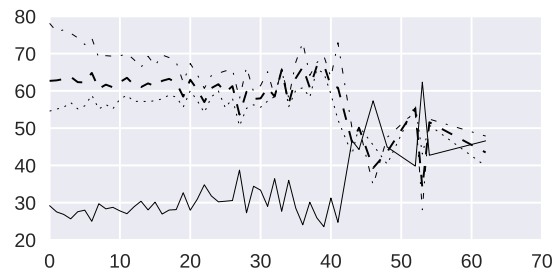
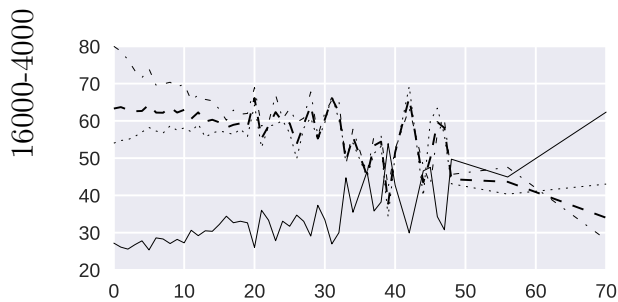
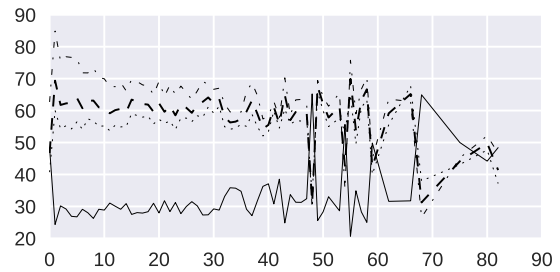
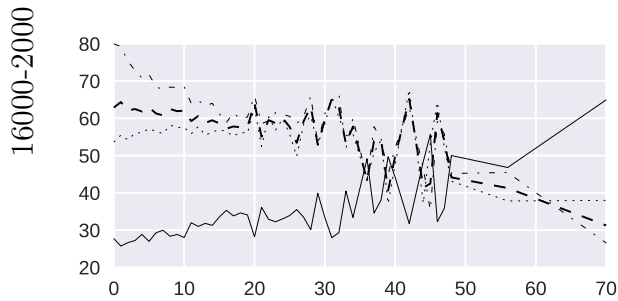
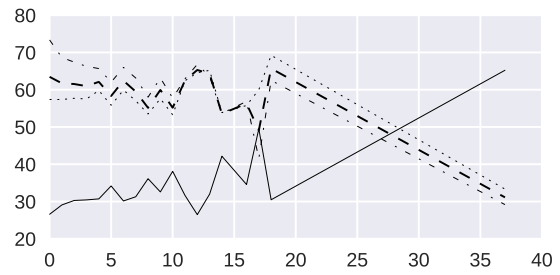
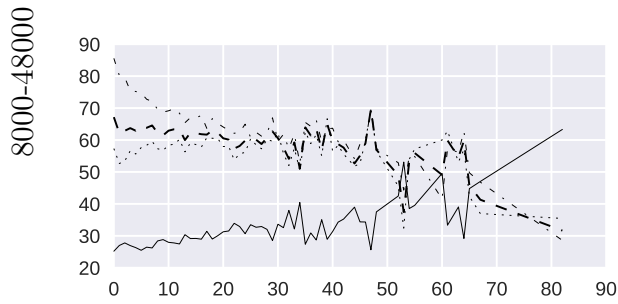


Figure E.143: BPE-based Fastalign for the direction English-Czech: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.

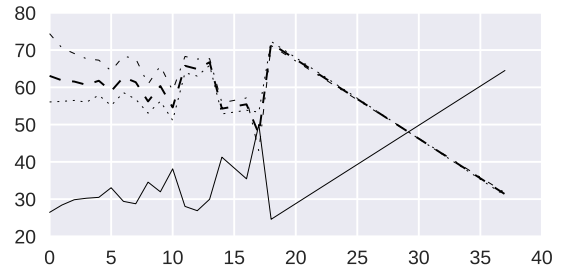
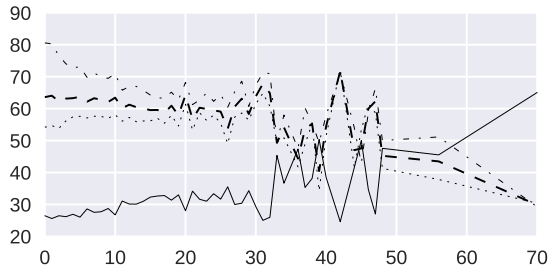




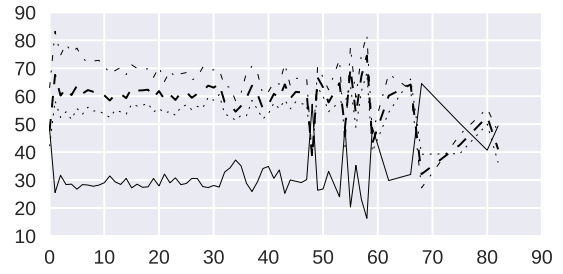
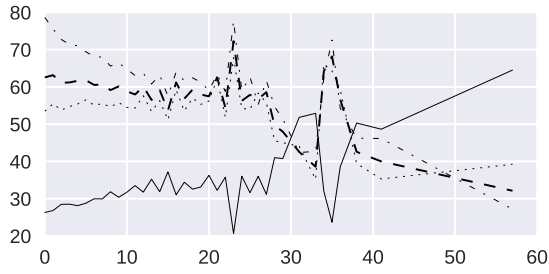




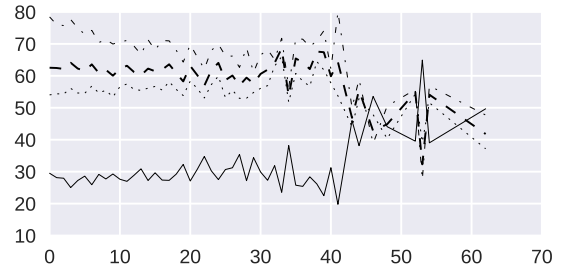
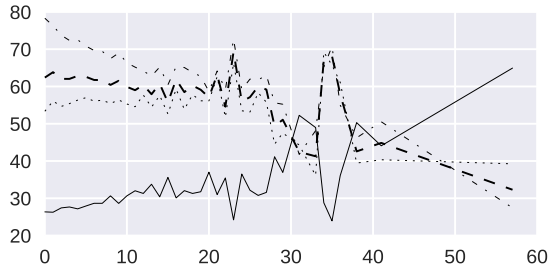
16000-48000



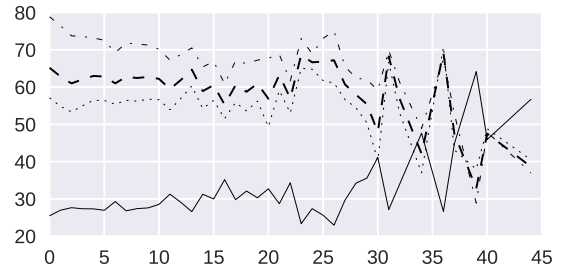
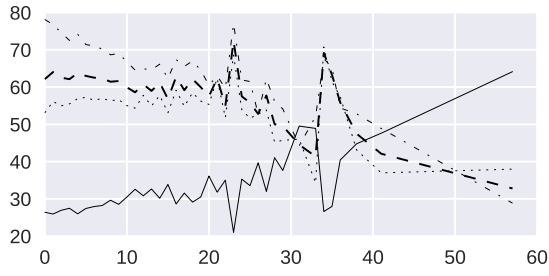
32000-20000



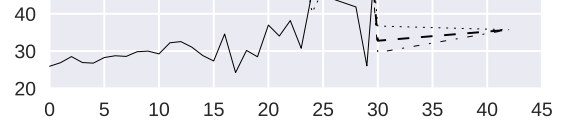
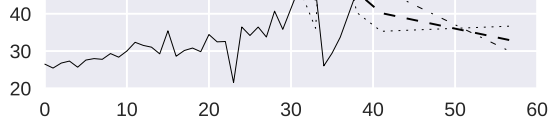
32000-4000



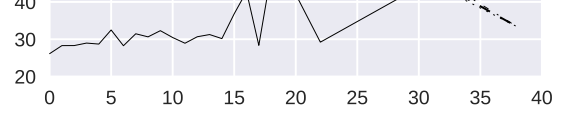
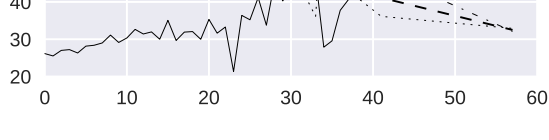
32000-8000

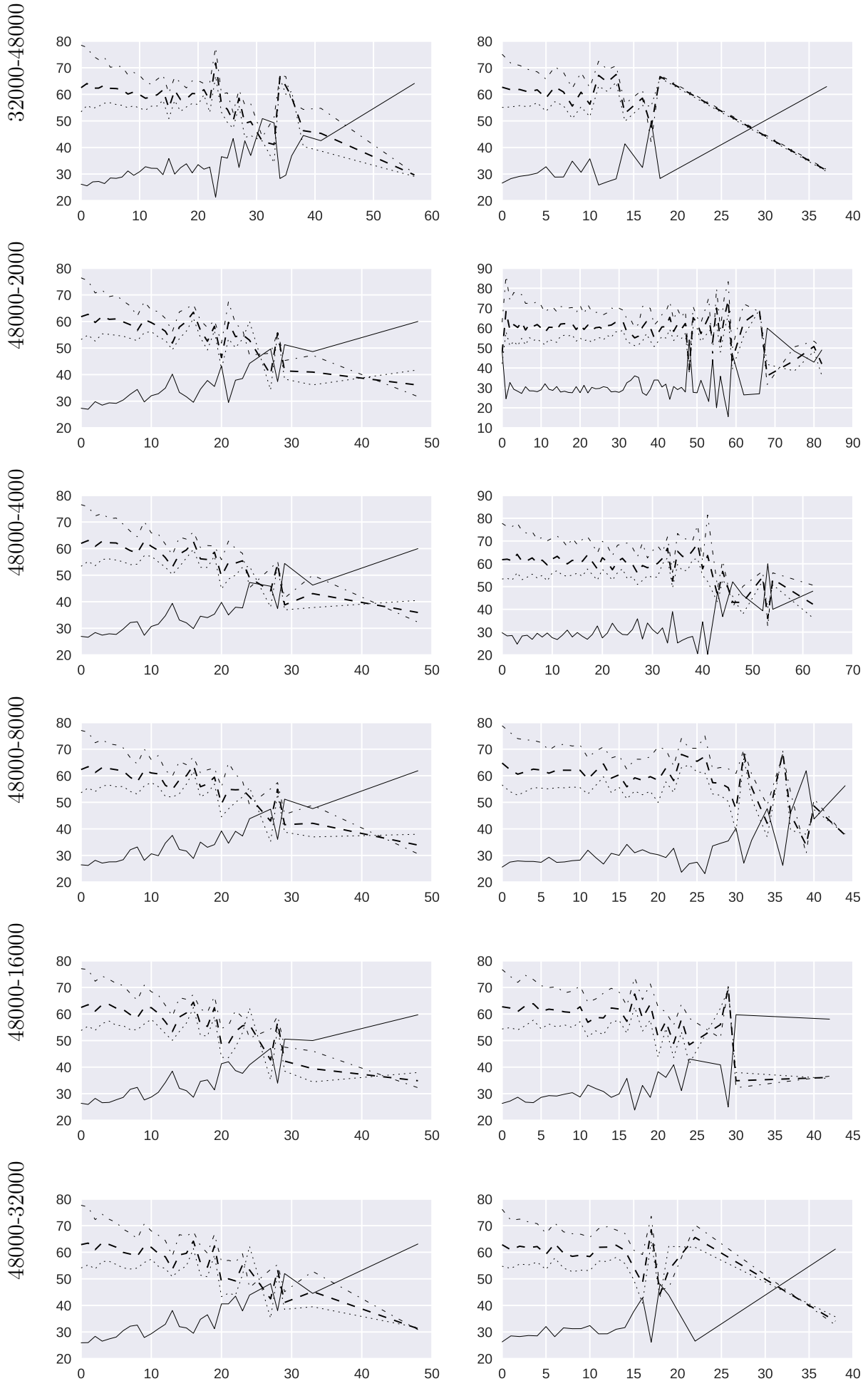


32000-16000



32000-32000





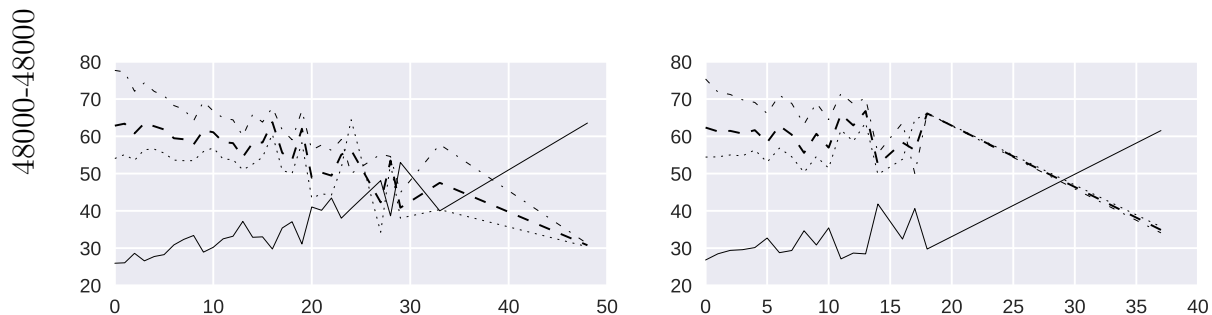
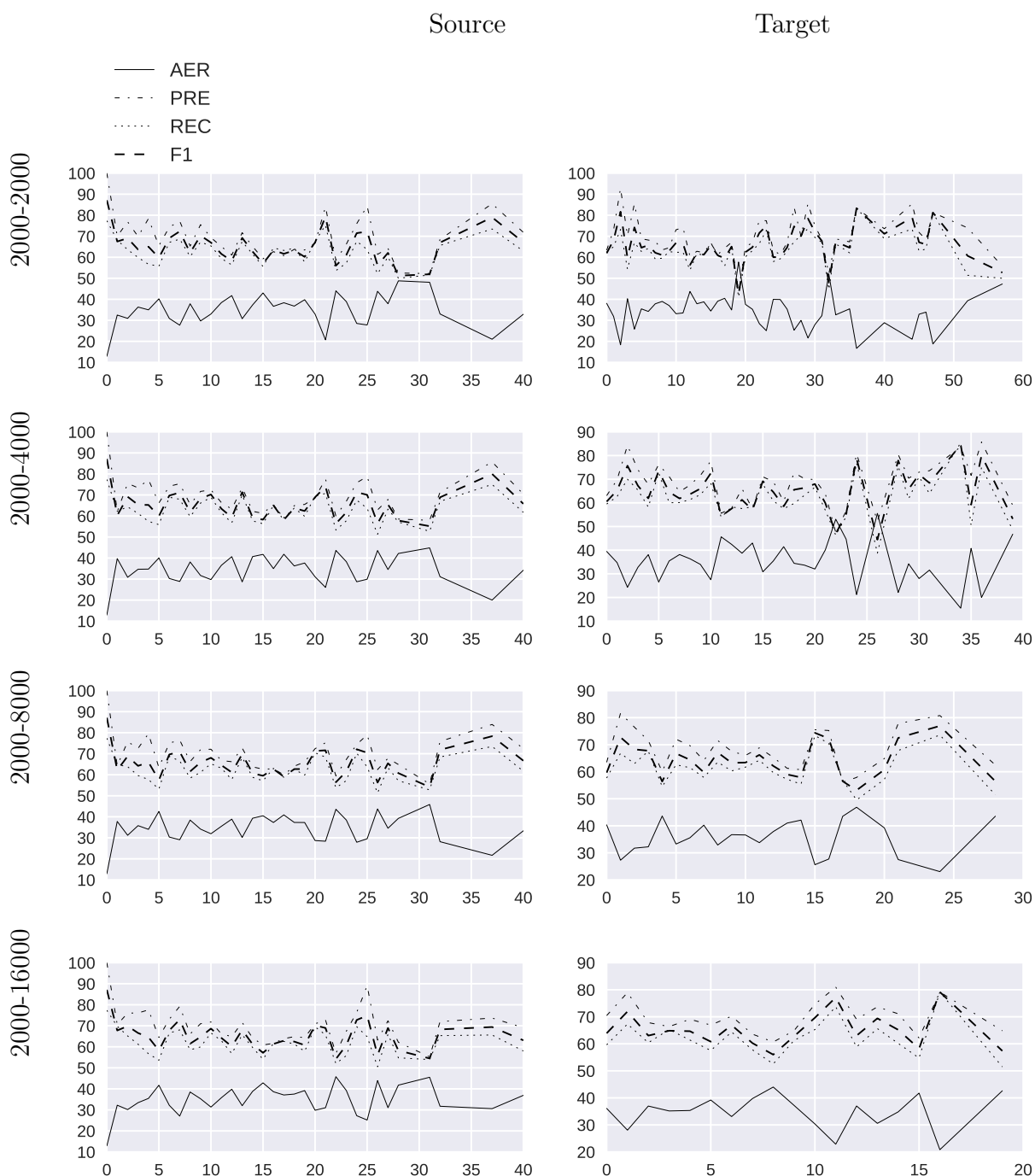
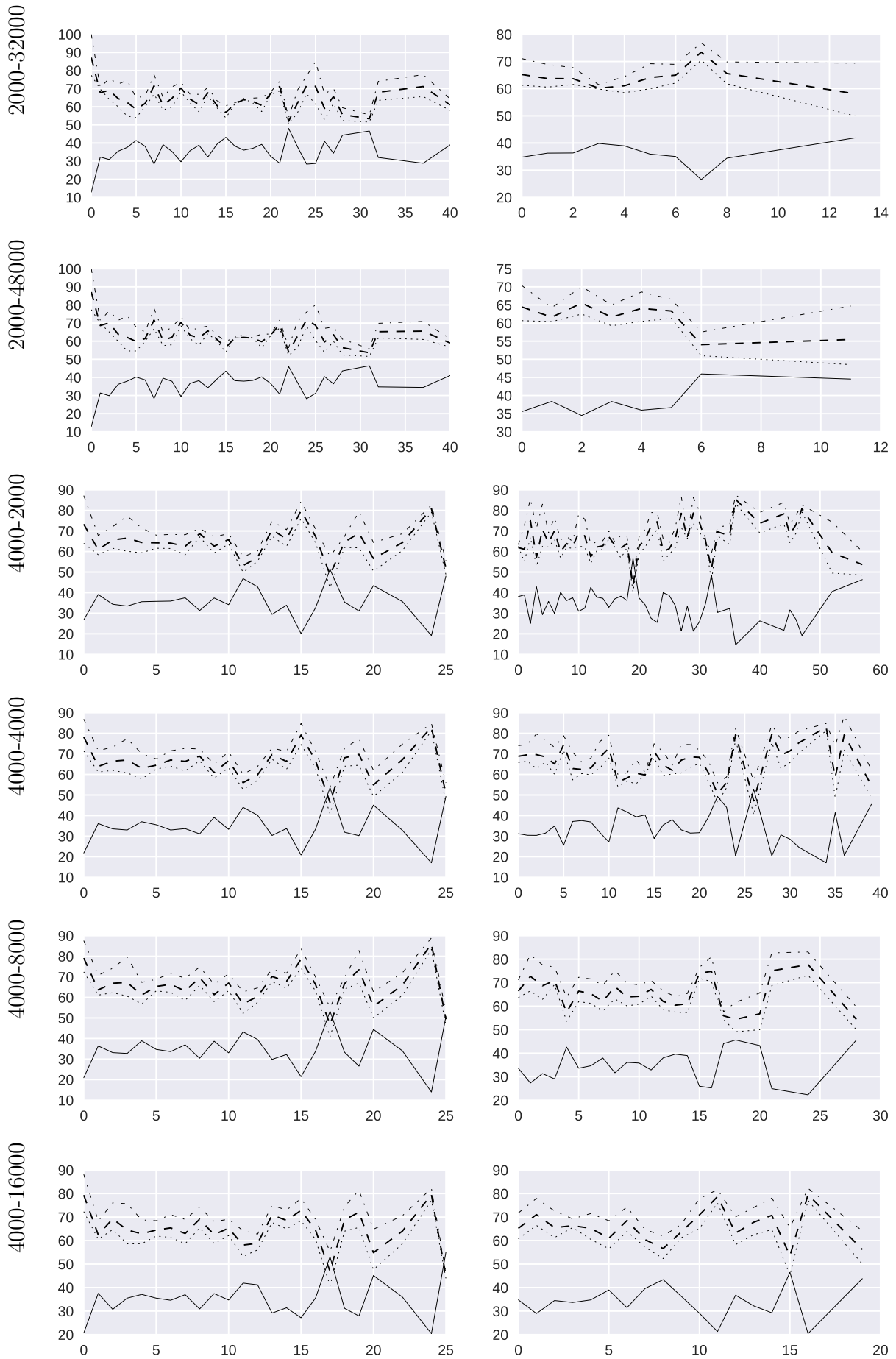
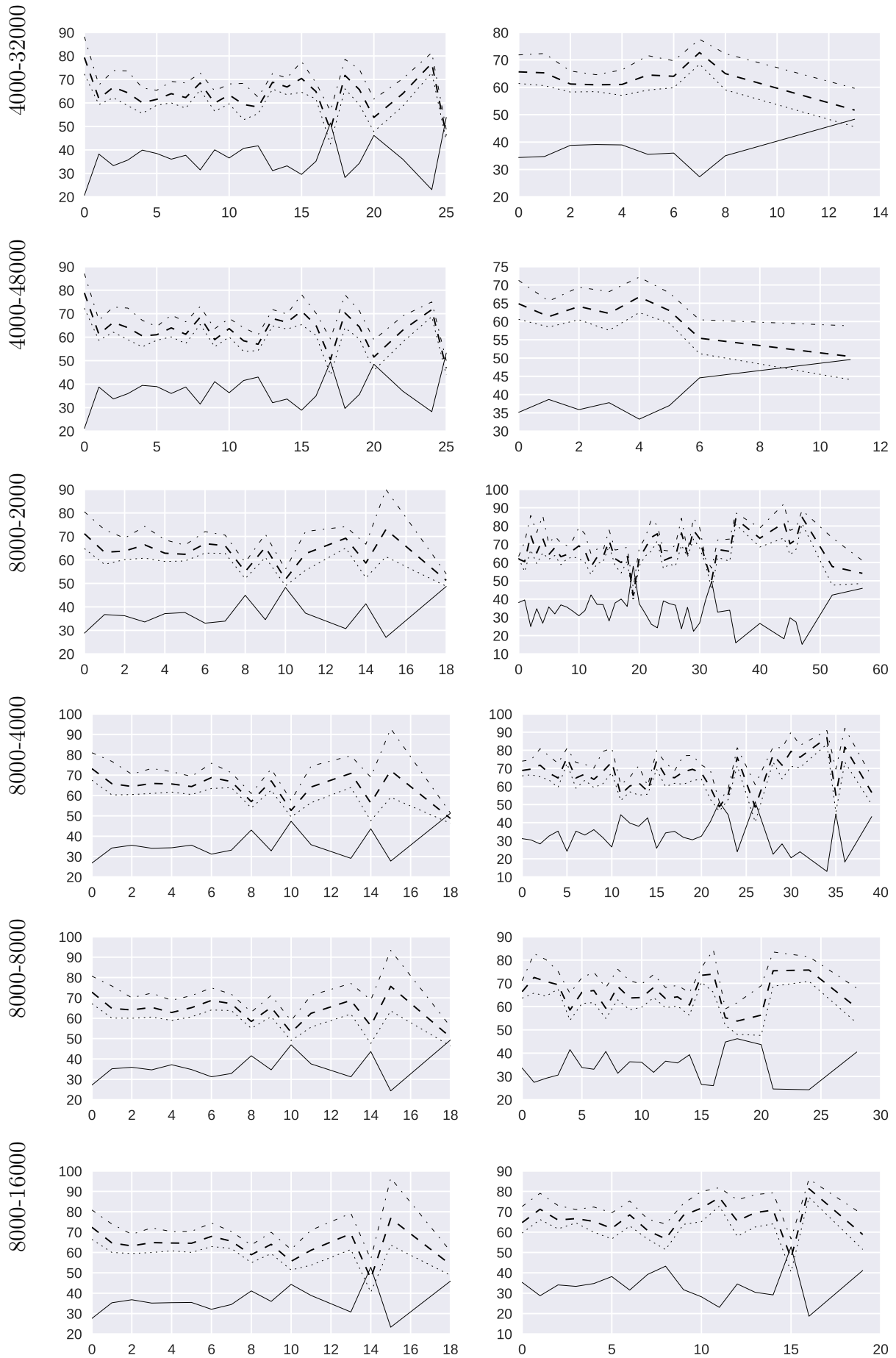
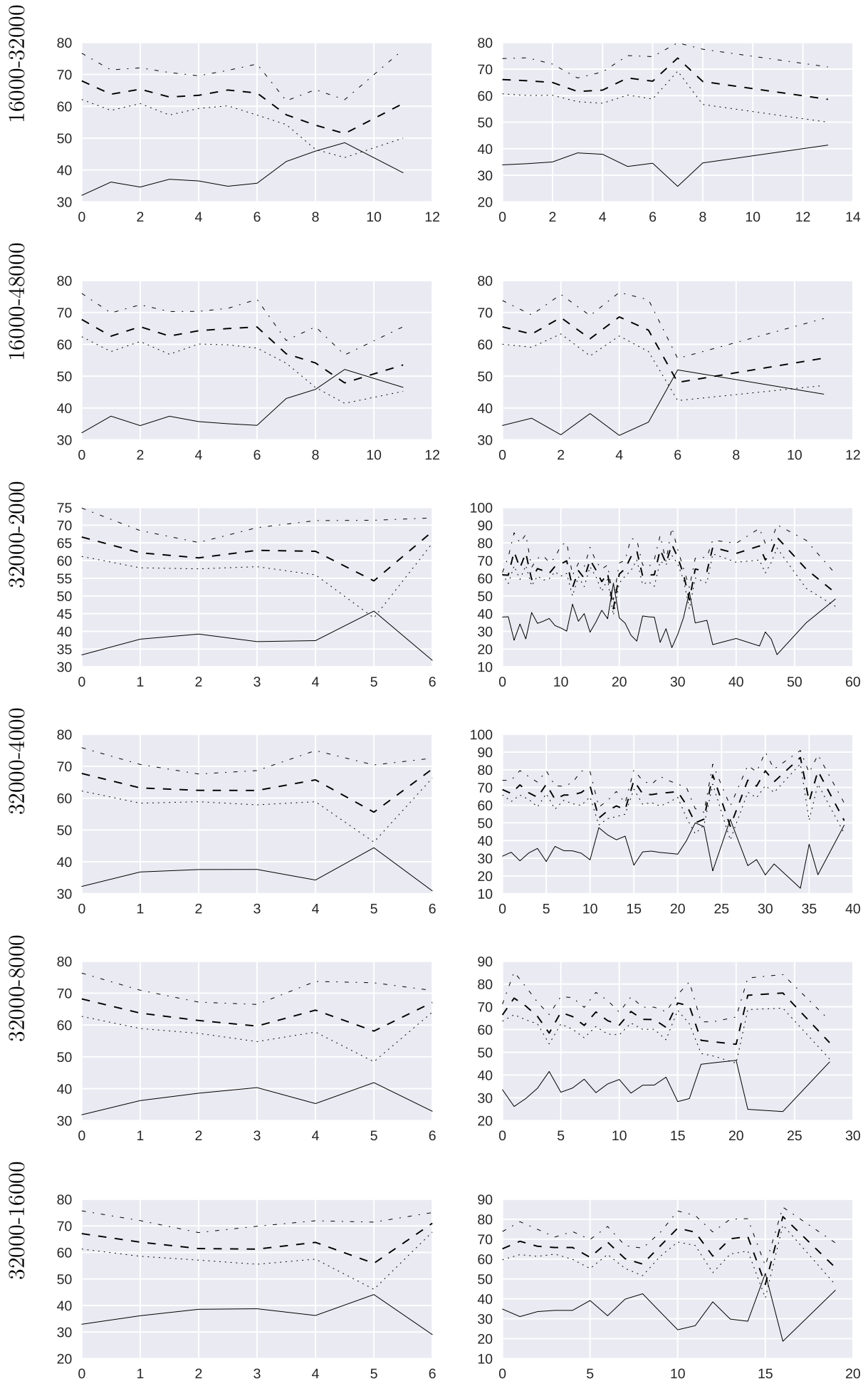


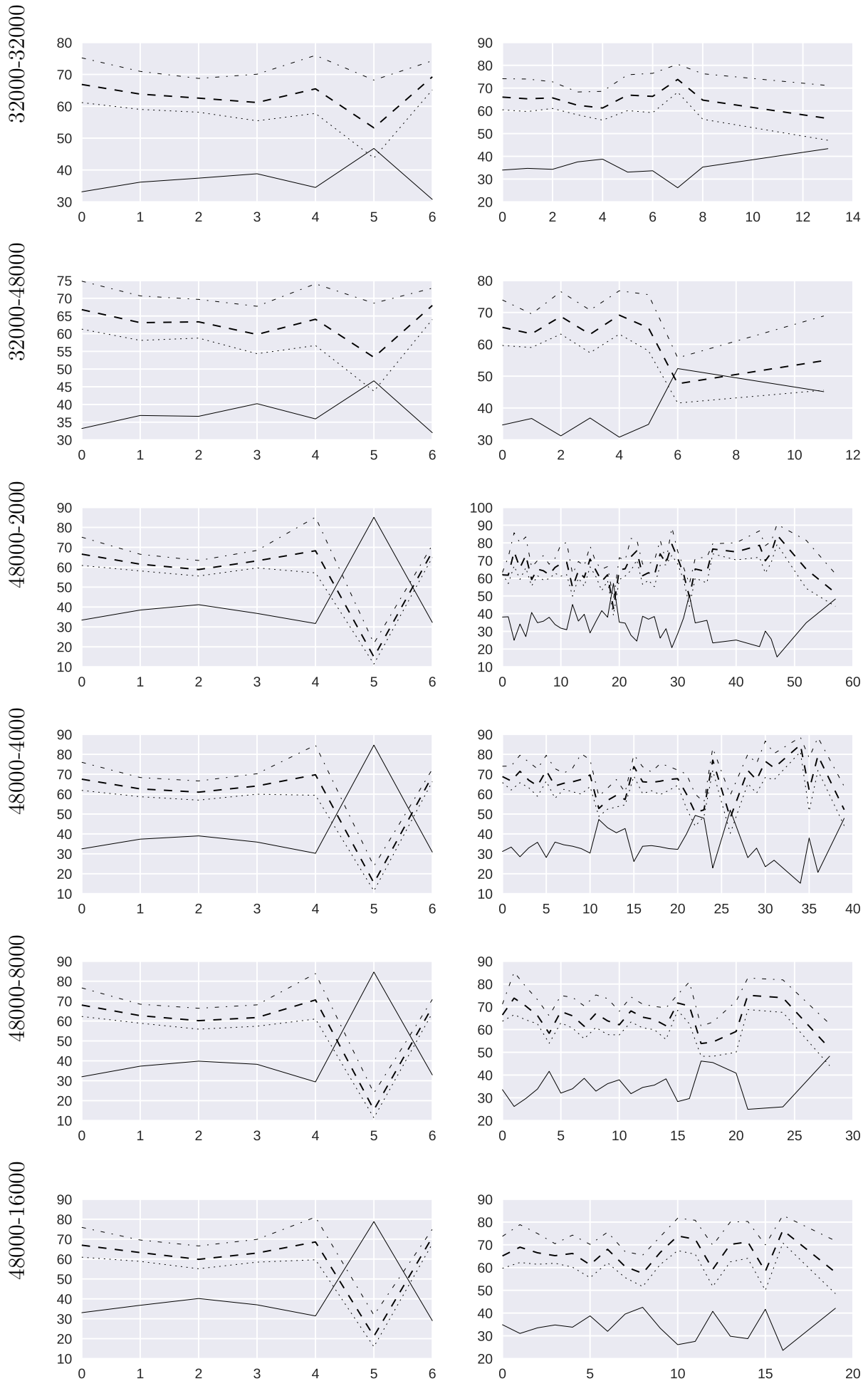
Figure E.144: BPE-based *Fastalign* for the direction Czech-English: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.











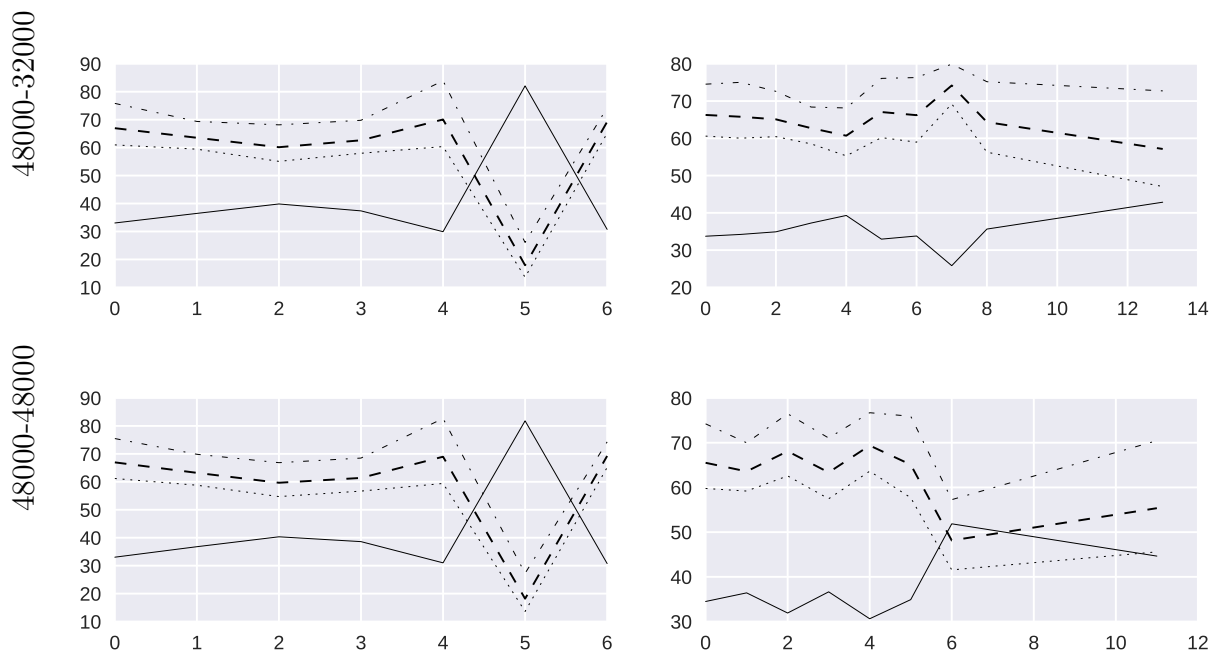
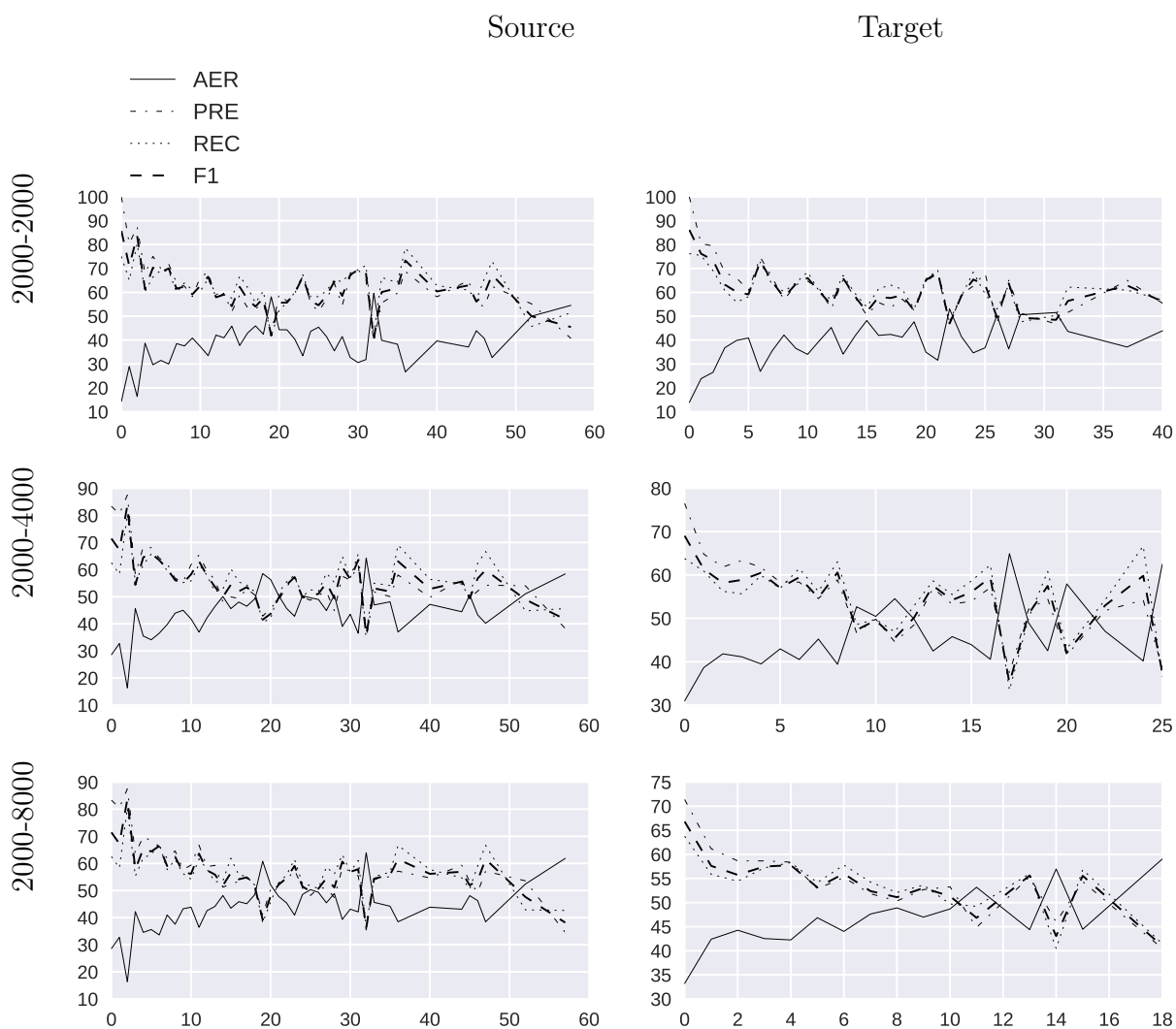
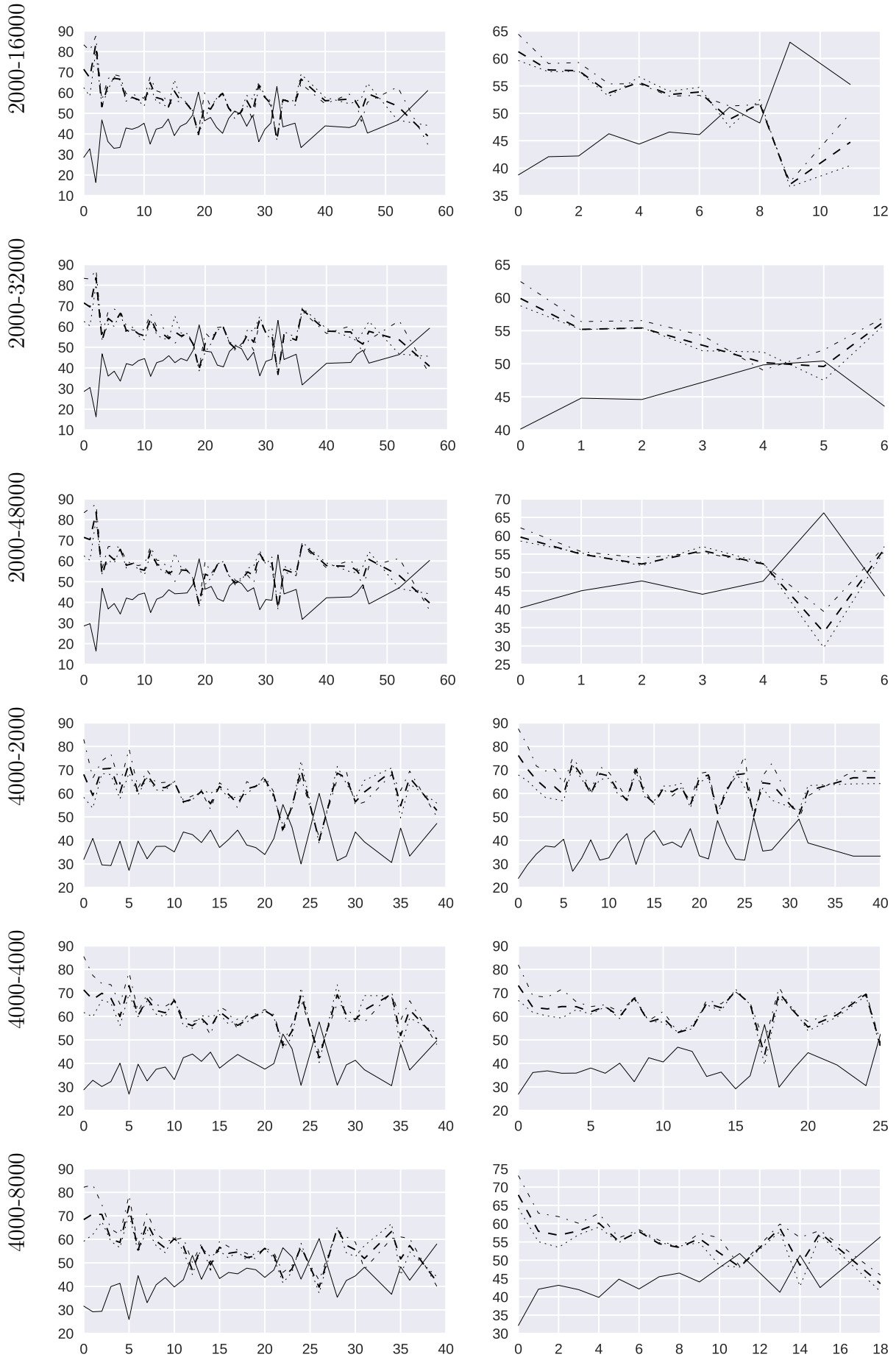
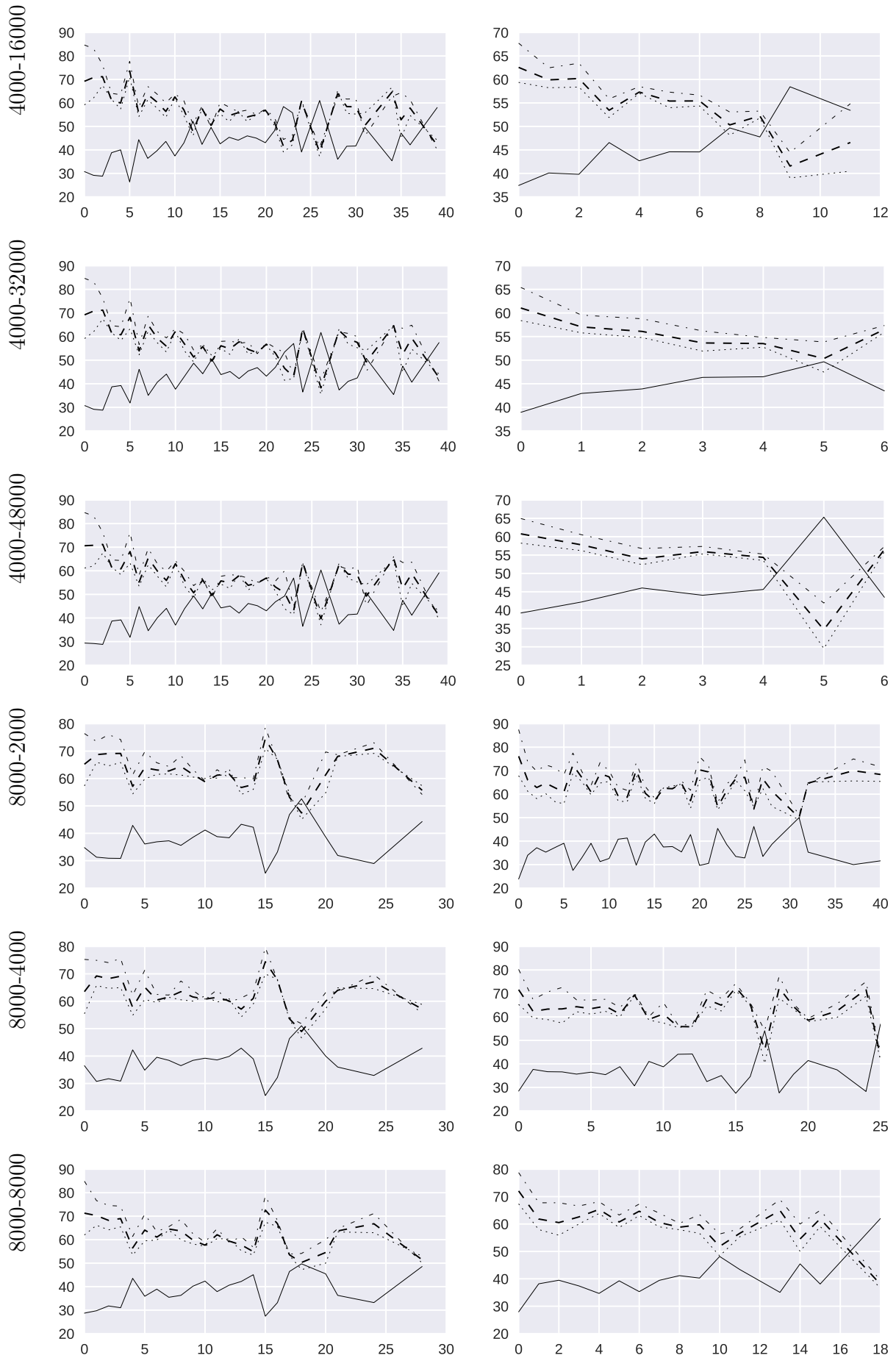
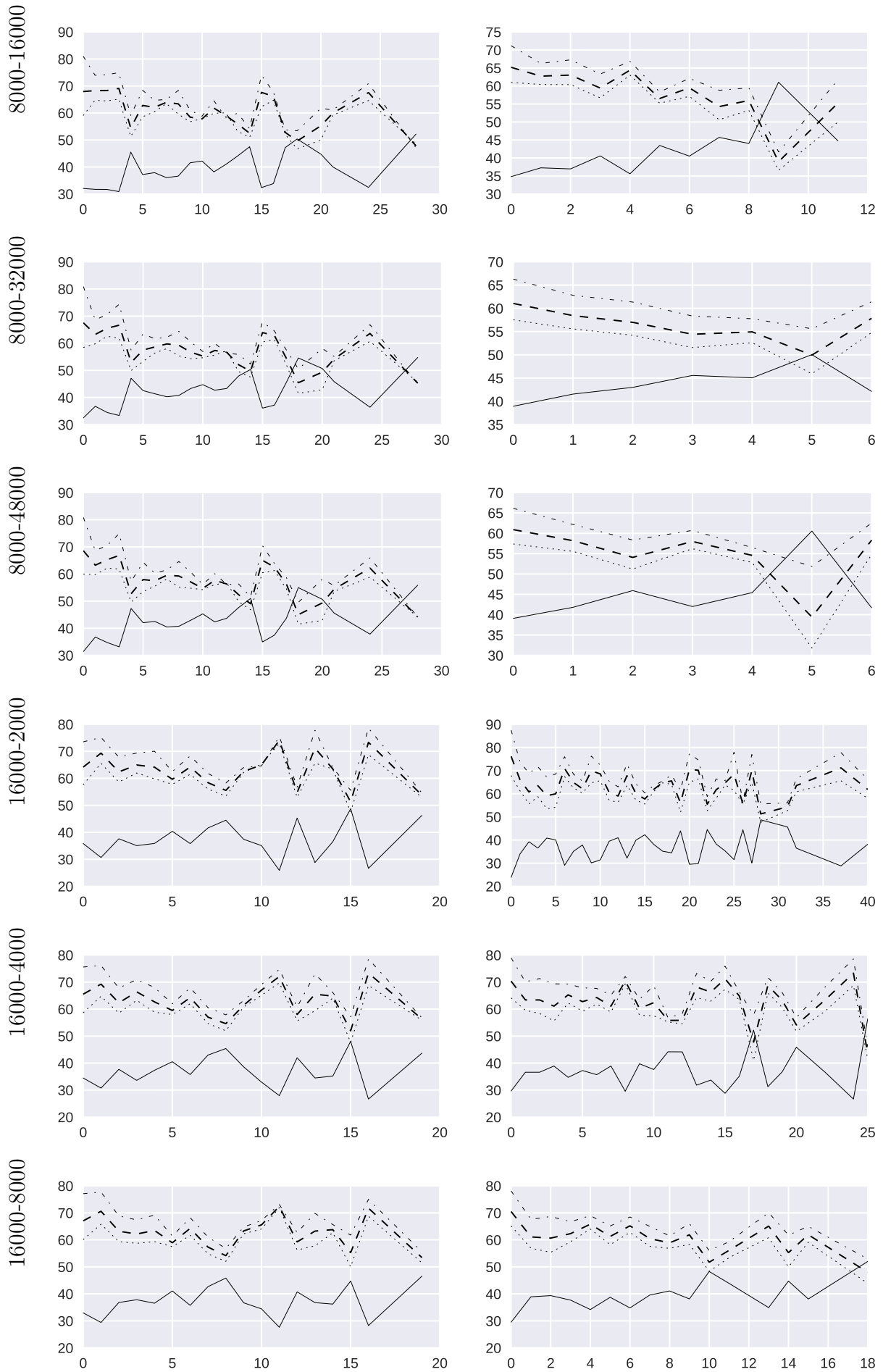


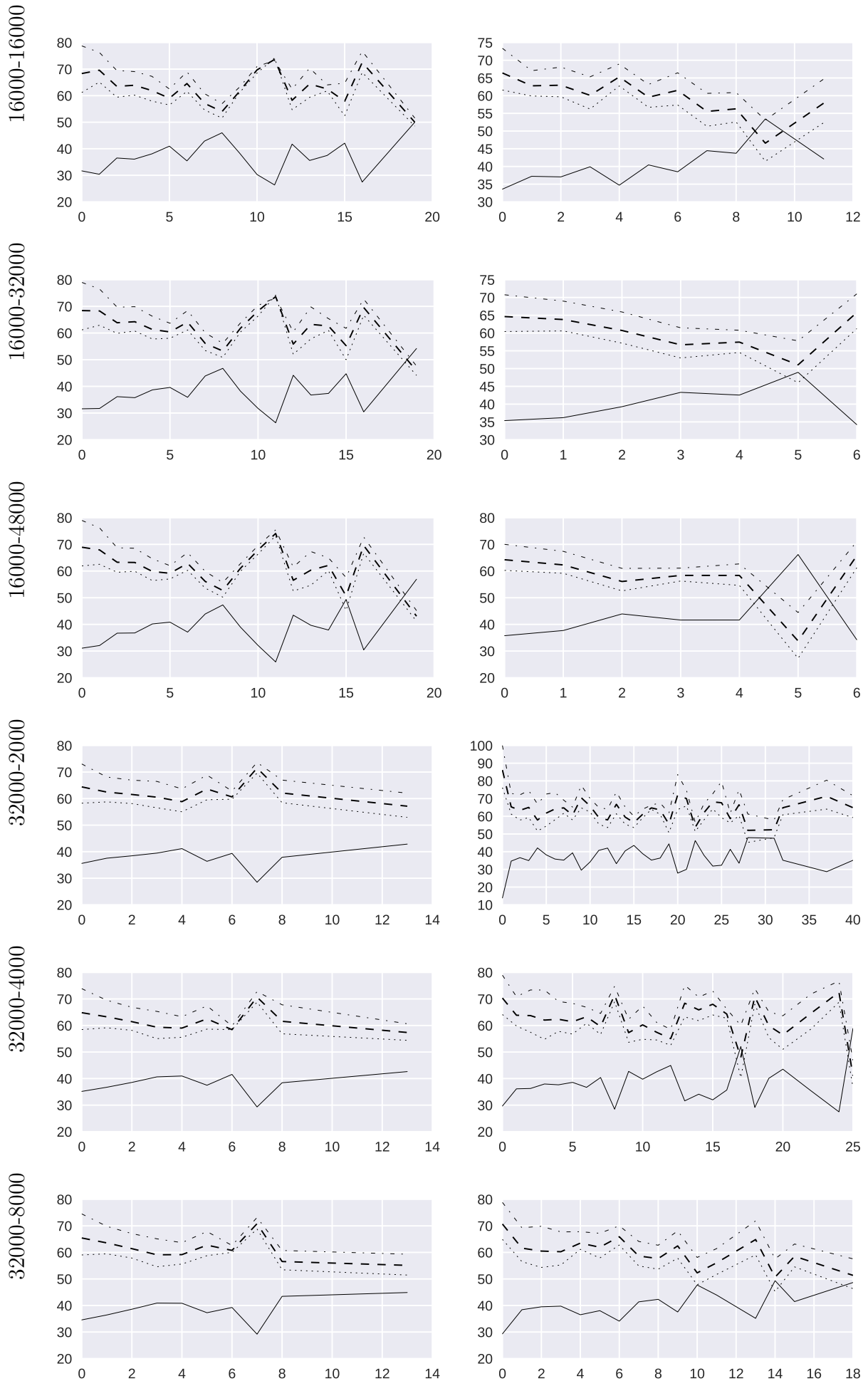
Figure E.145: BPE-based `Fastalign` for the direction English-Romanian: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.

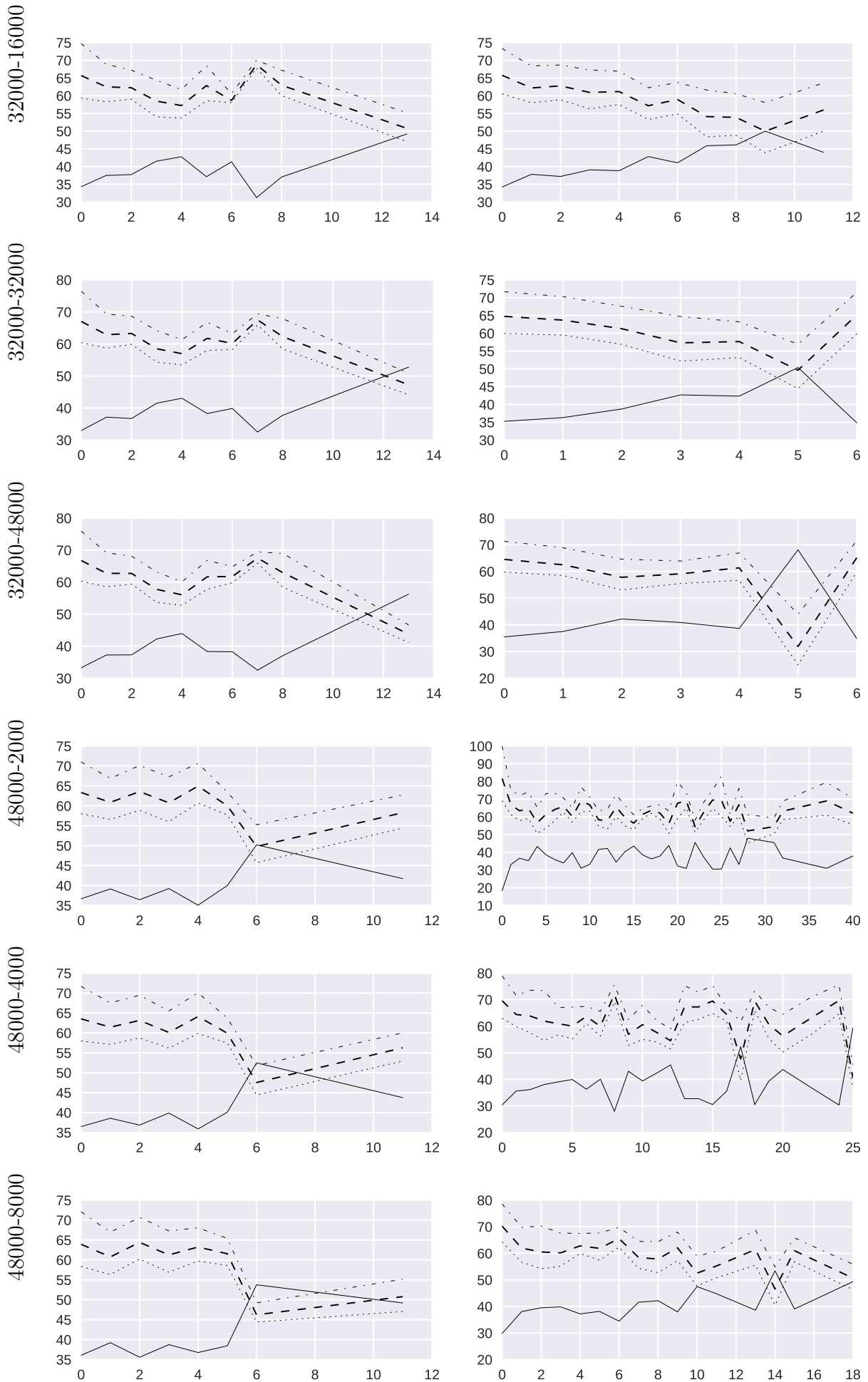












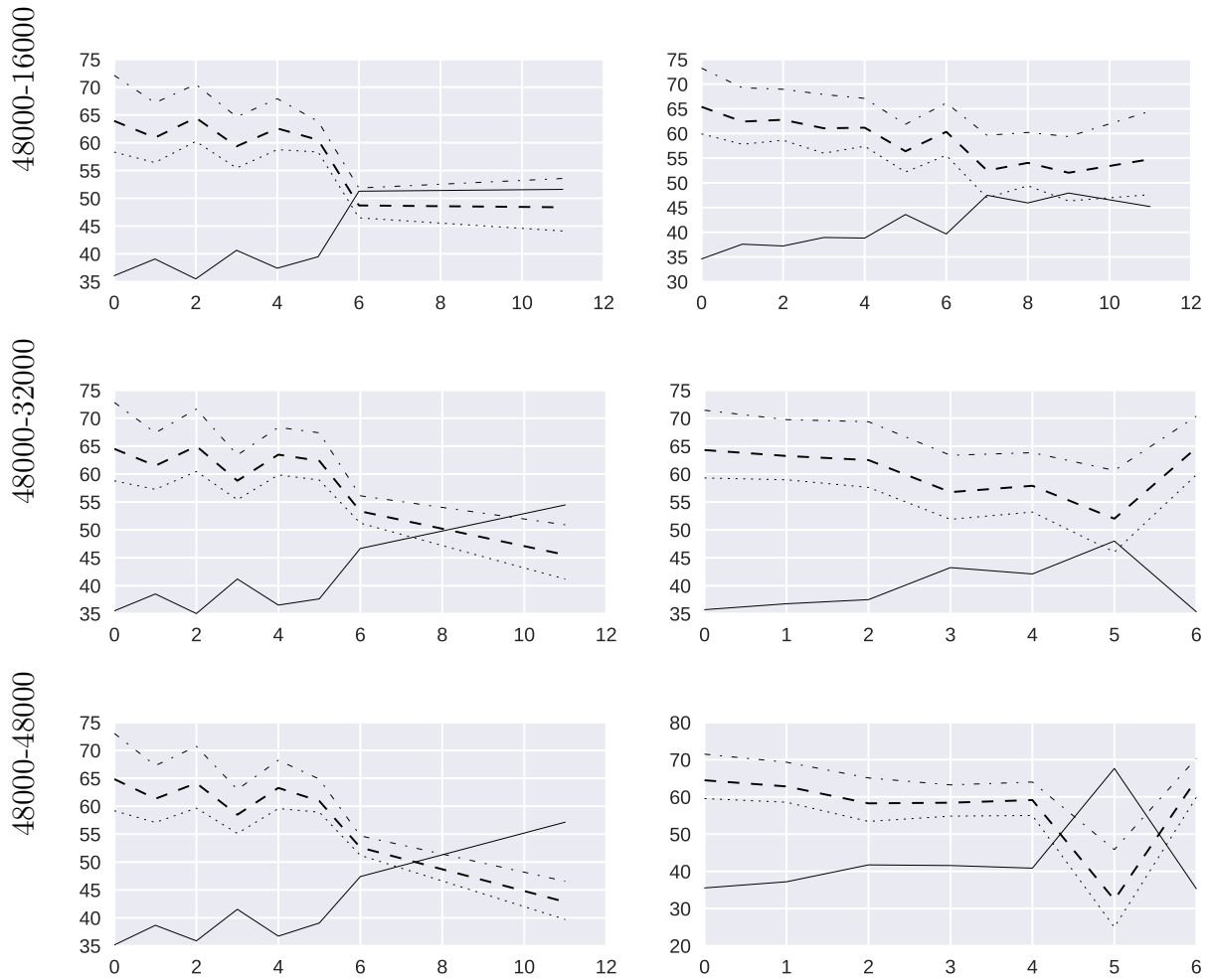
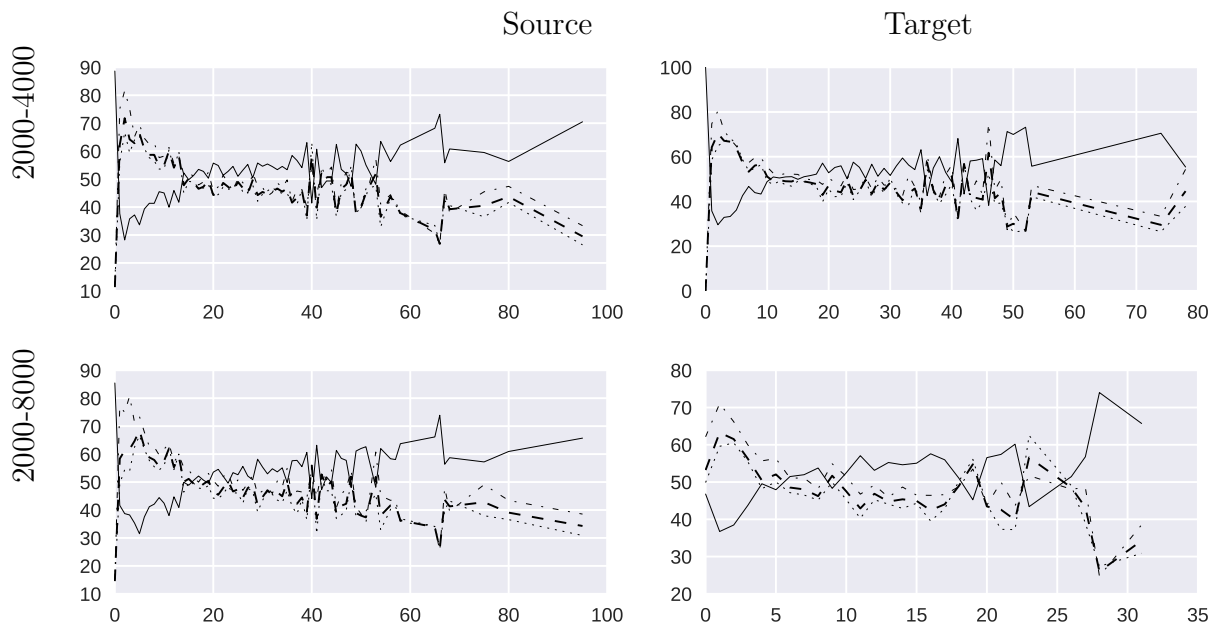
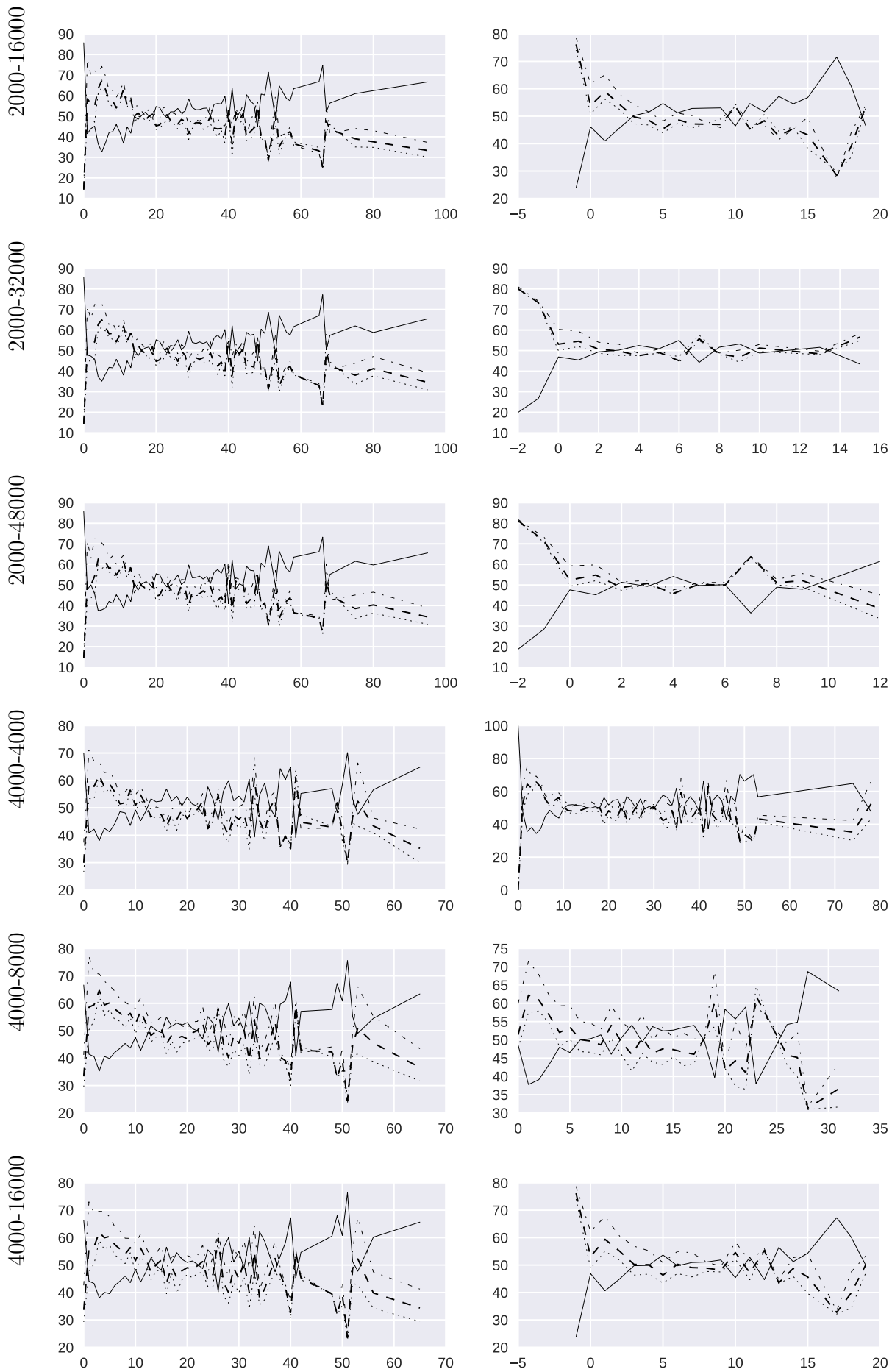
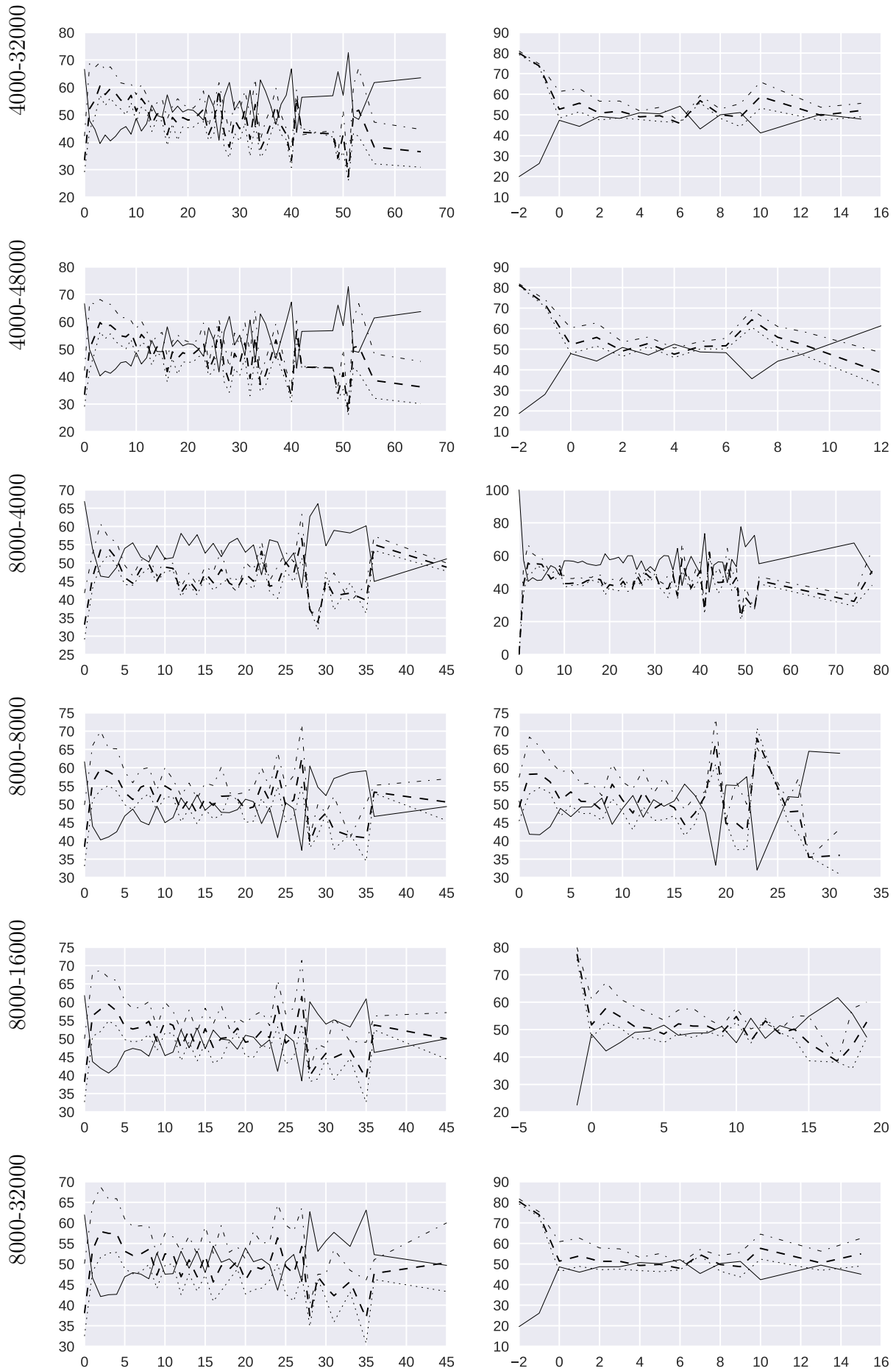
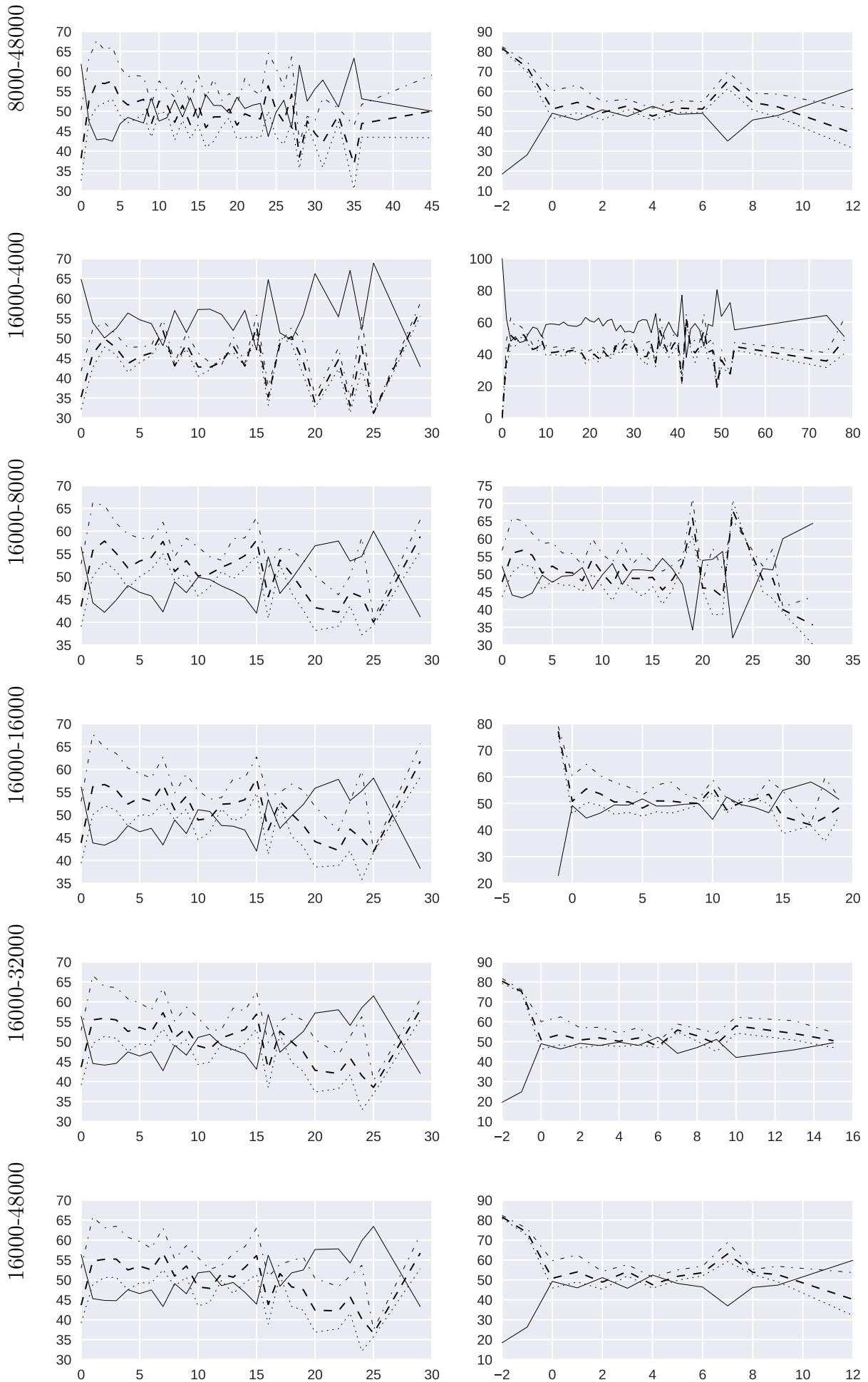


Figure E.146: BPE-based **Fastalign** for the direction Romanian-English: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.

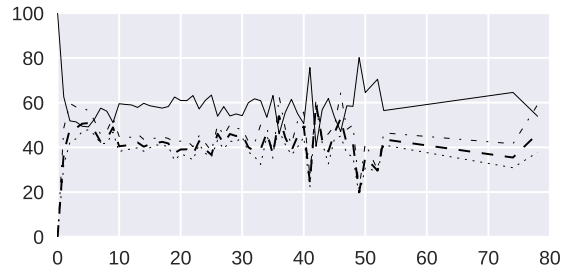
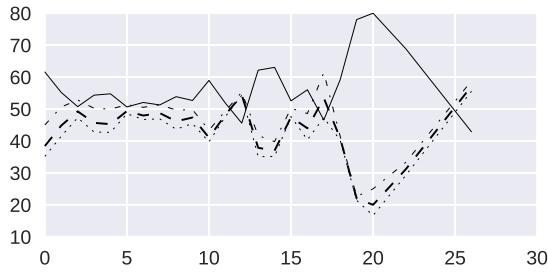




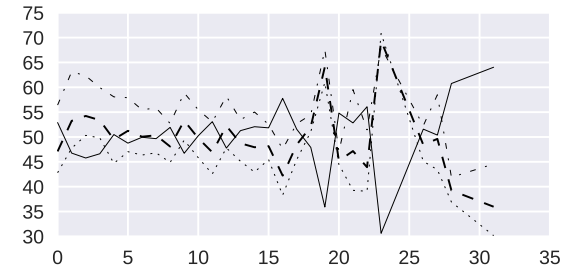
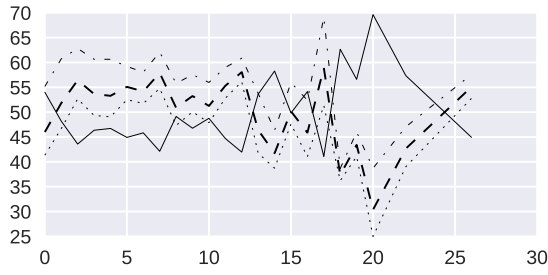




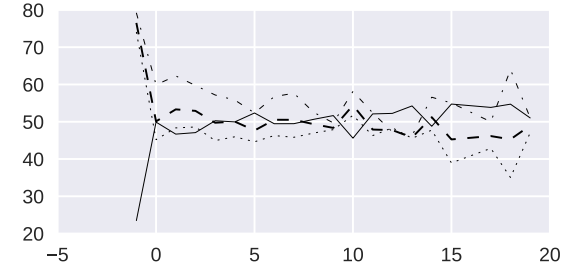
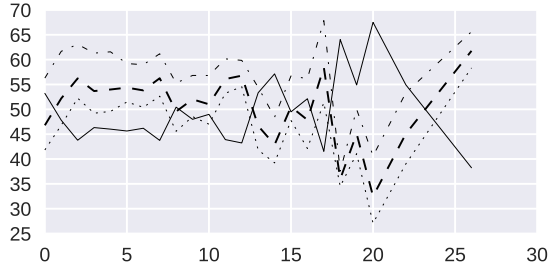
32000-4000



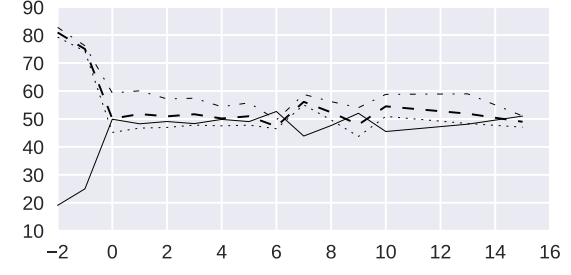
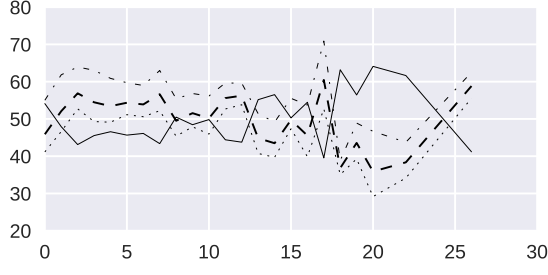
32000-8000



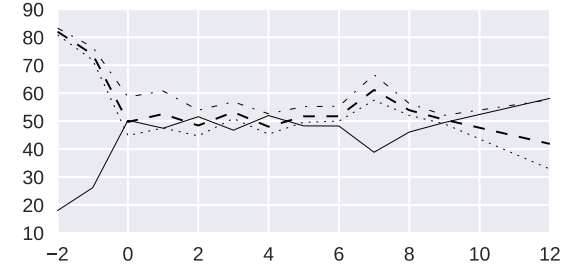
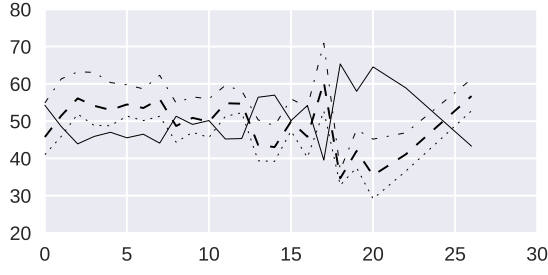
32000-16000



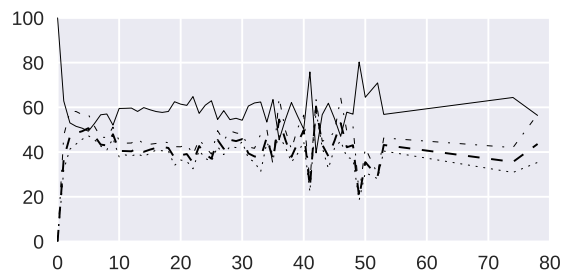
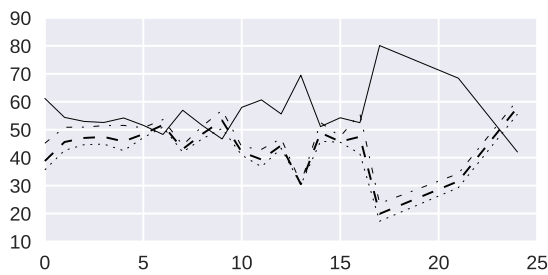
32000-32000



32000-48000



48000-4000



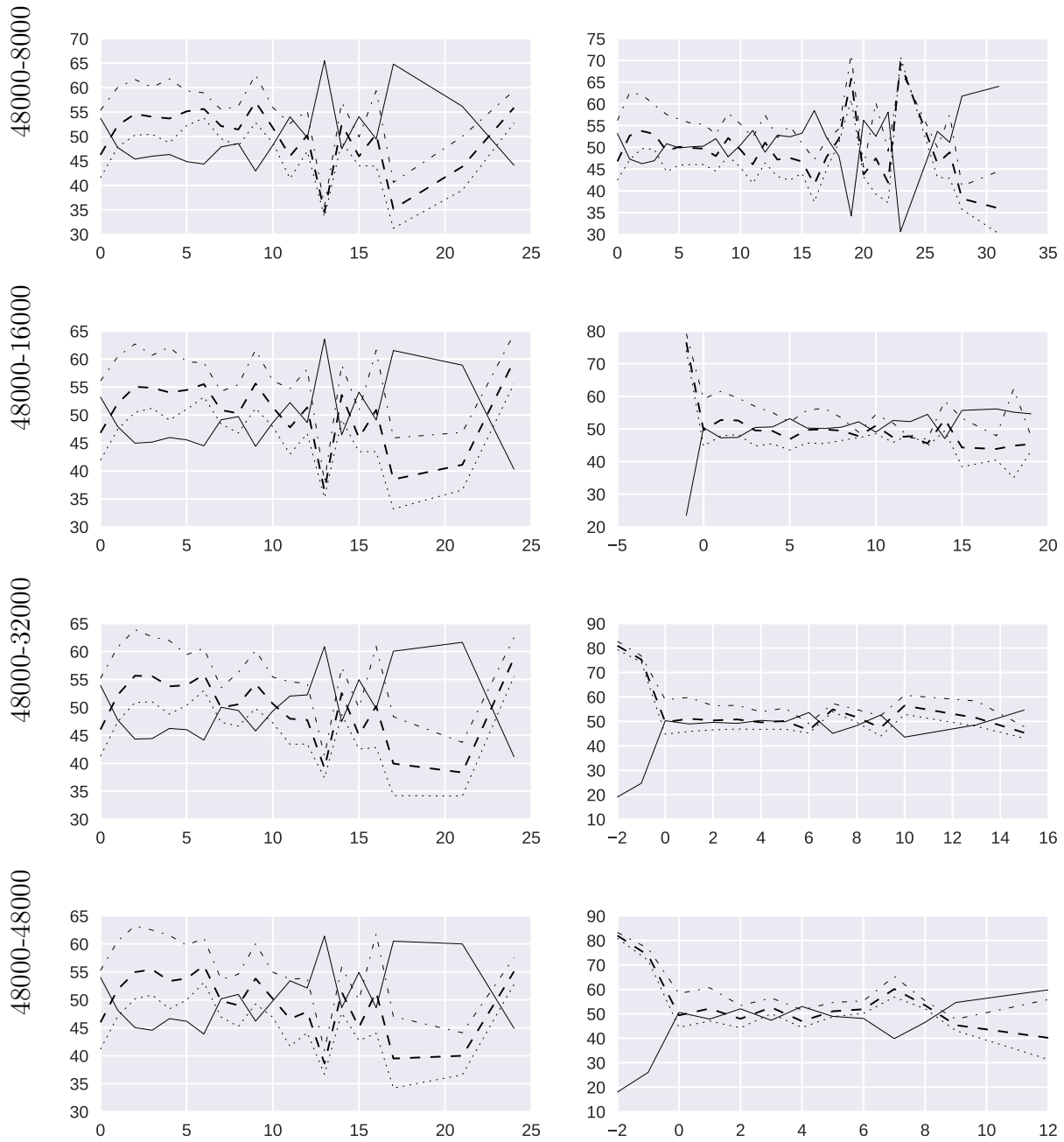
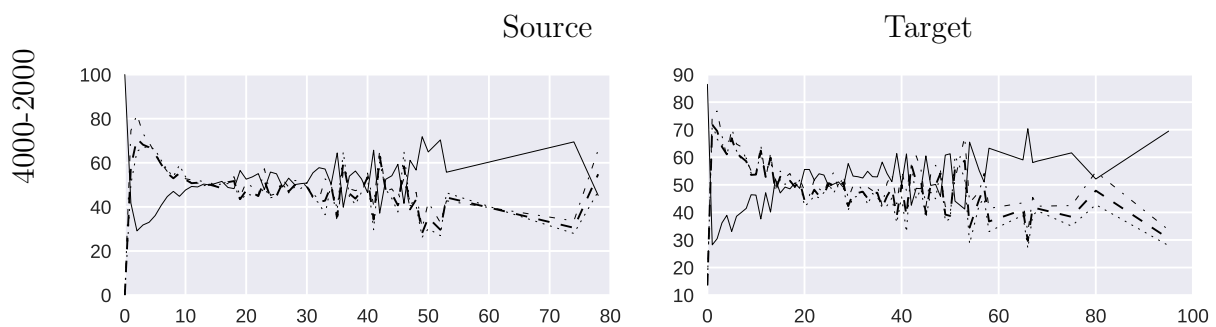
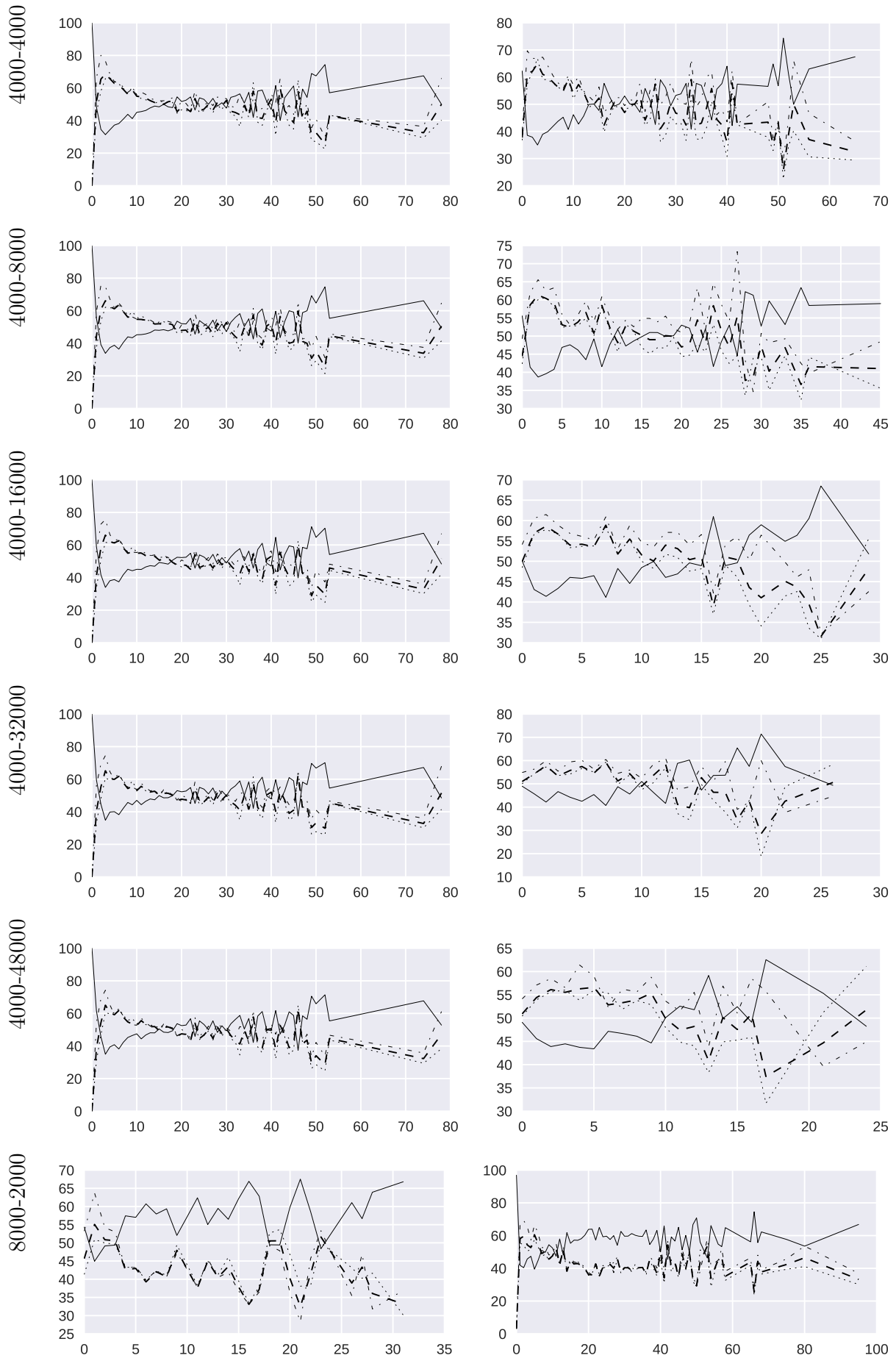
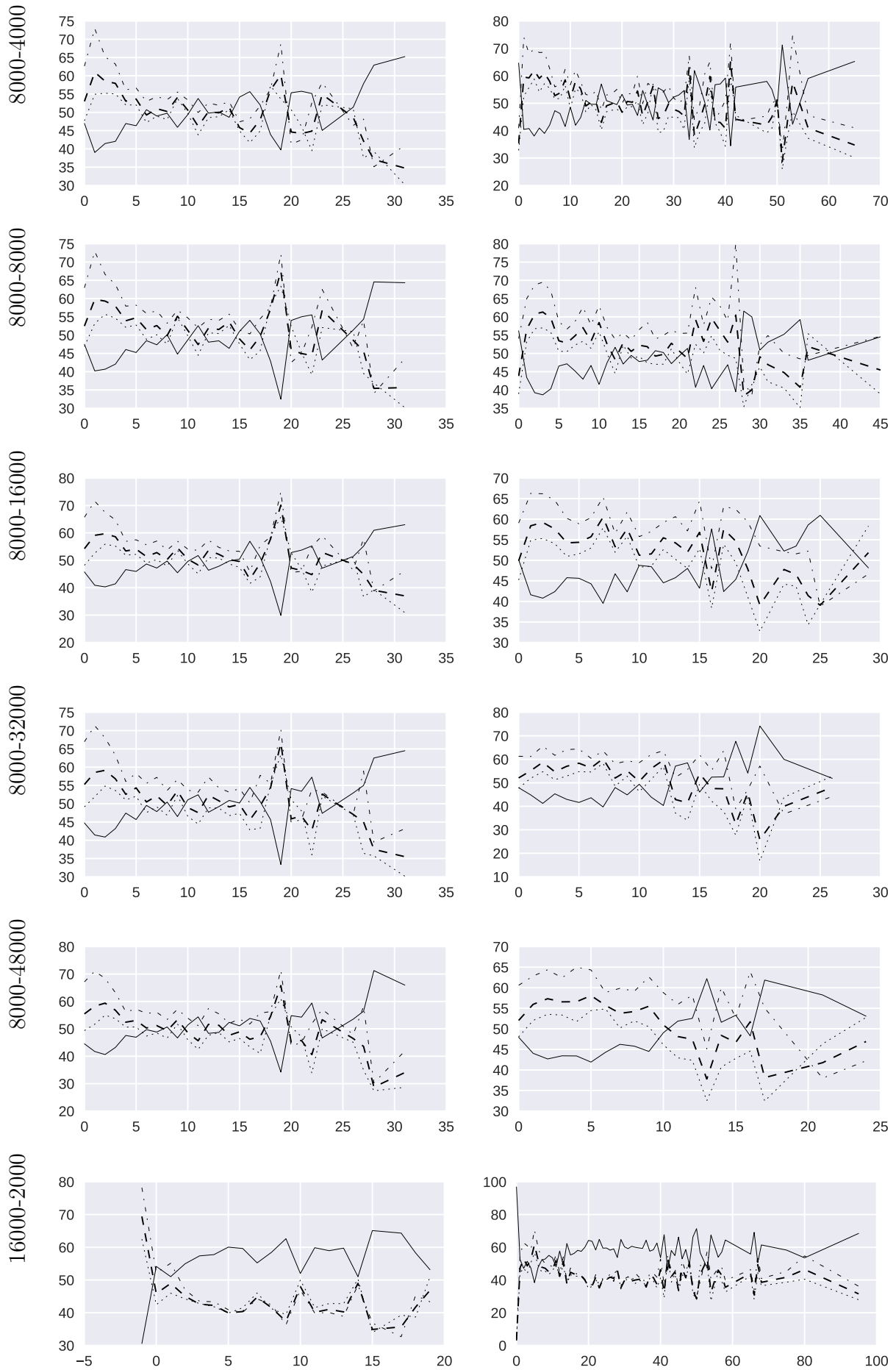
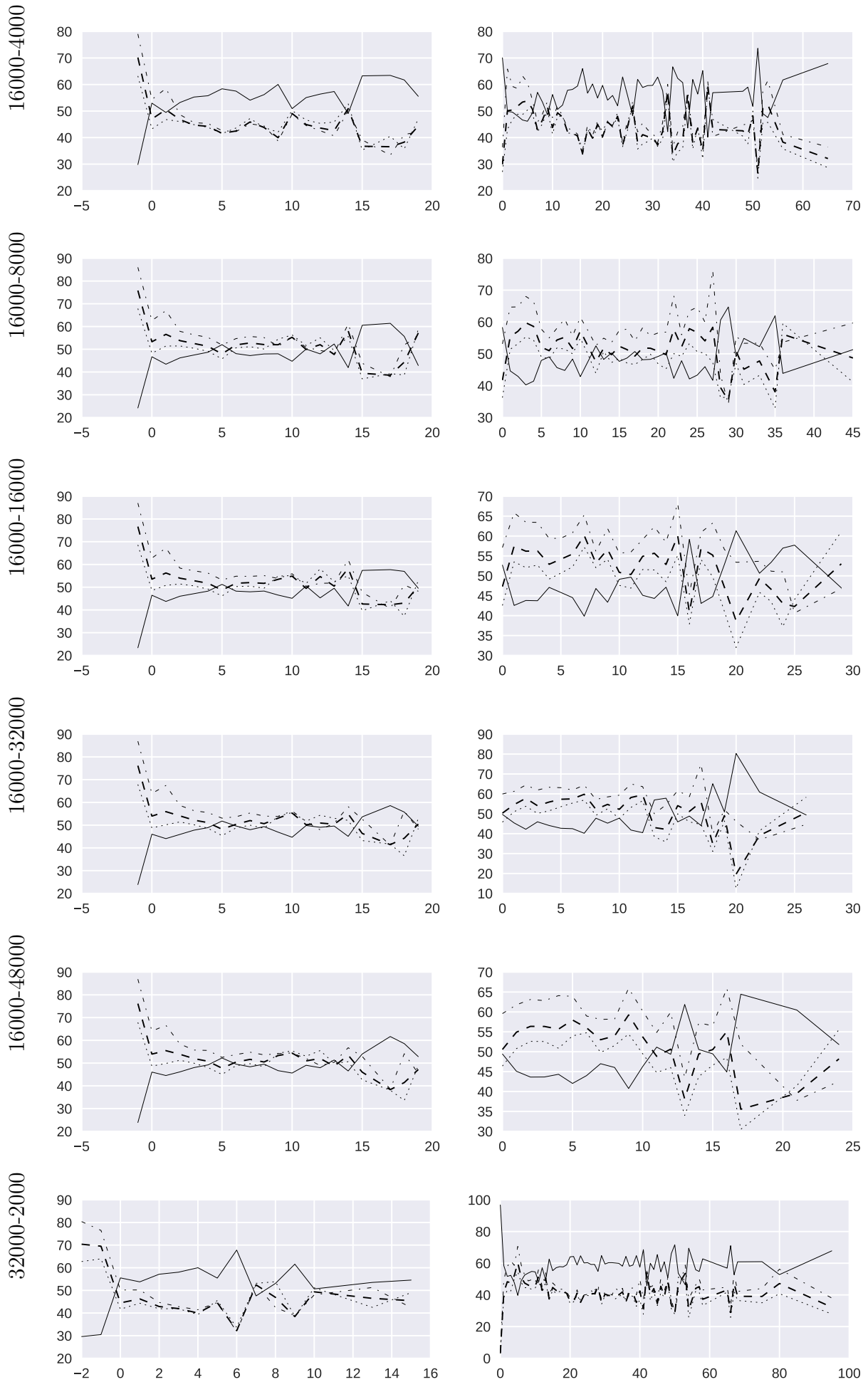


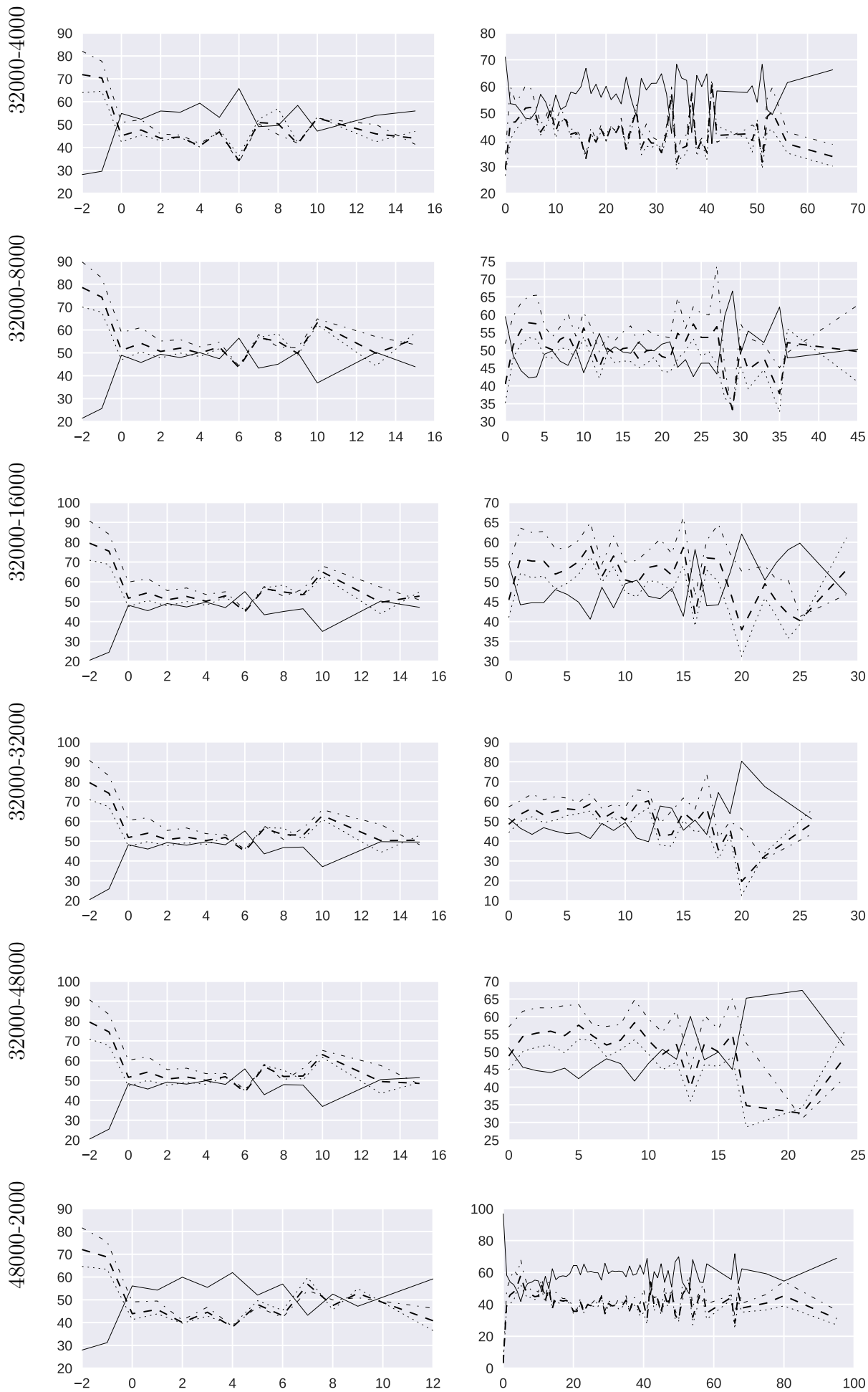
Figure E.147: BPE-based *Fastalign* for the direction English-Japanese: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.











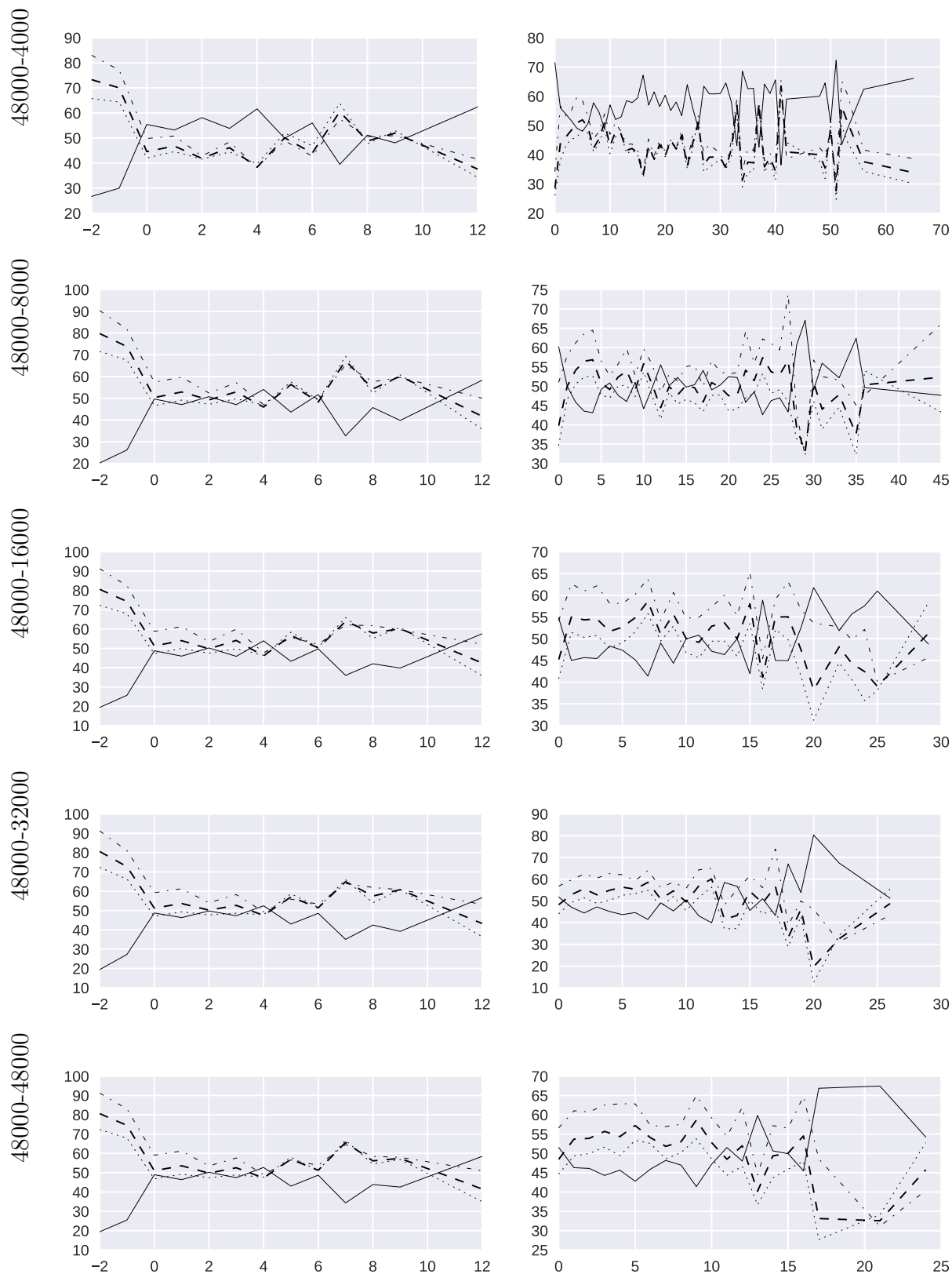
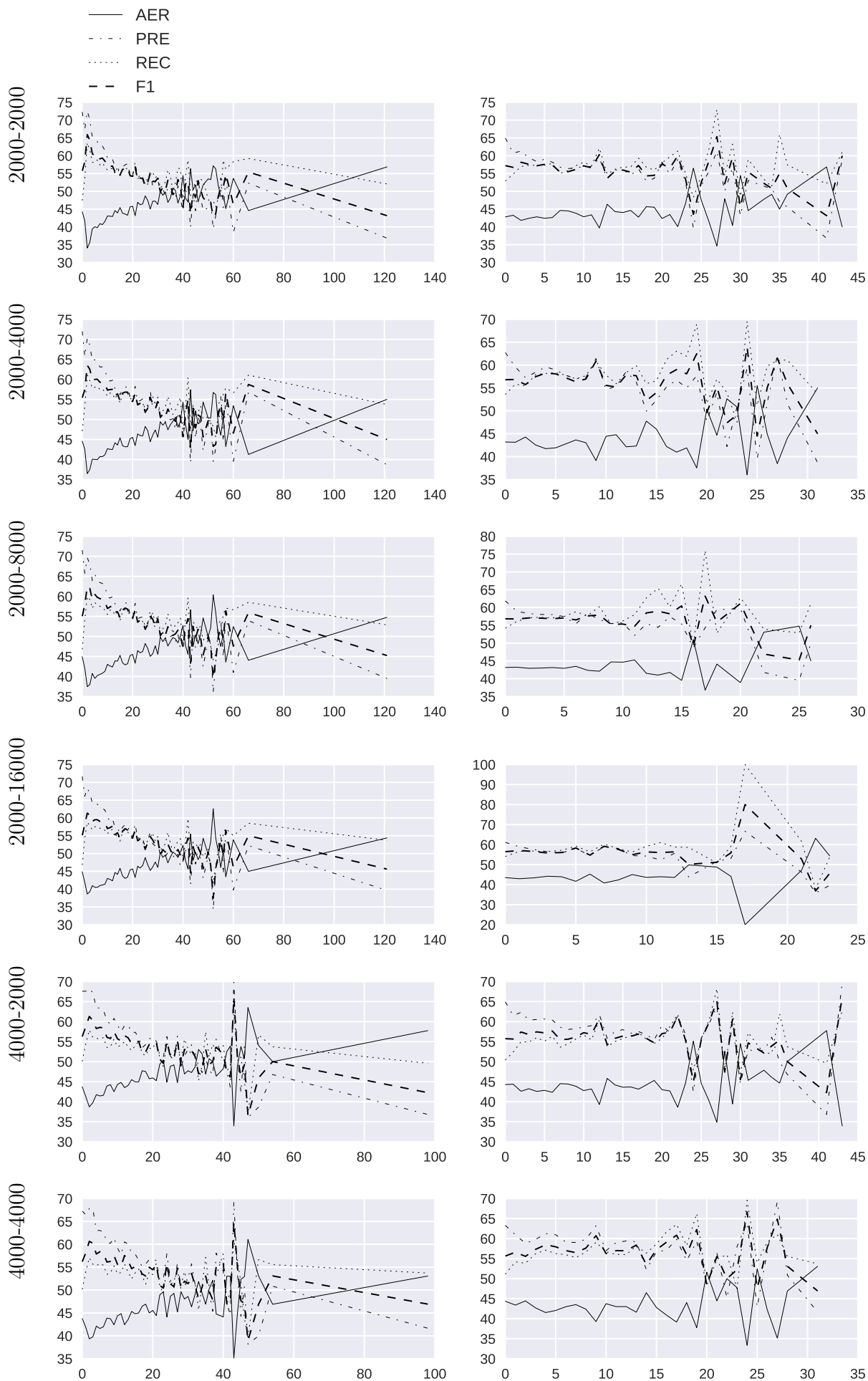
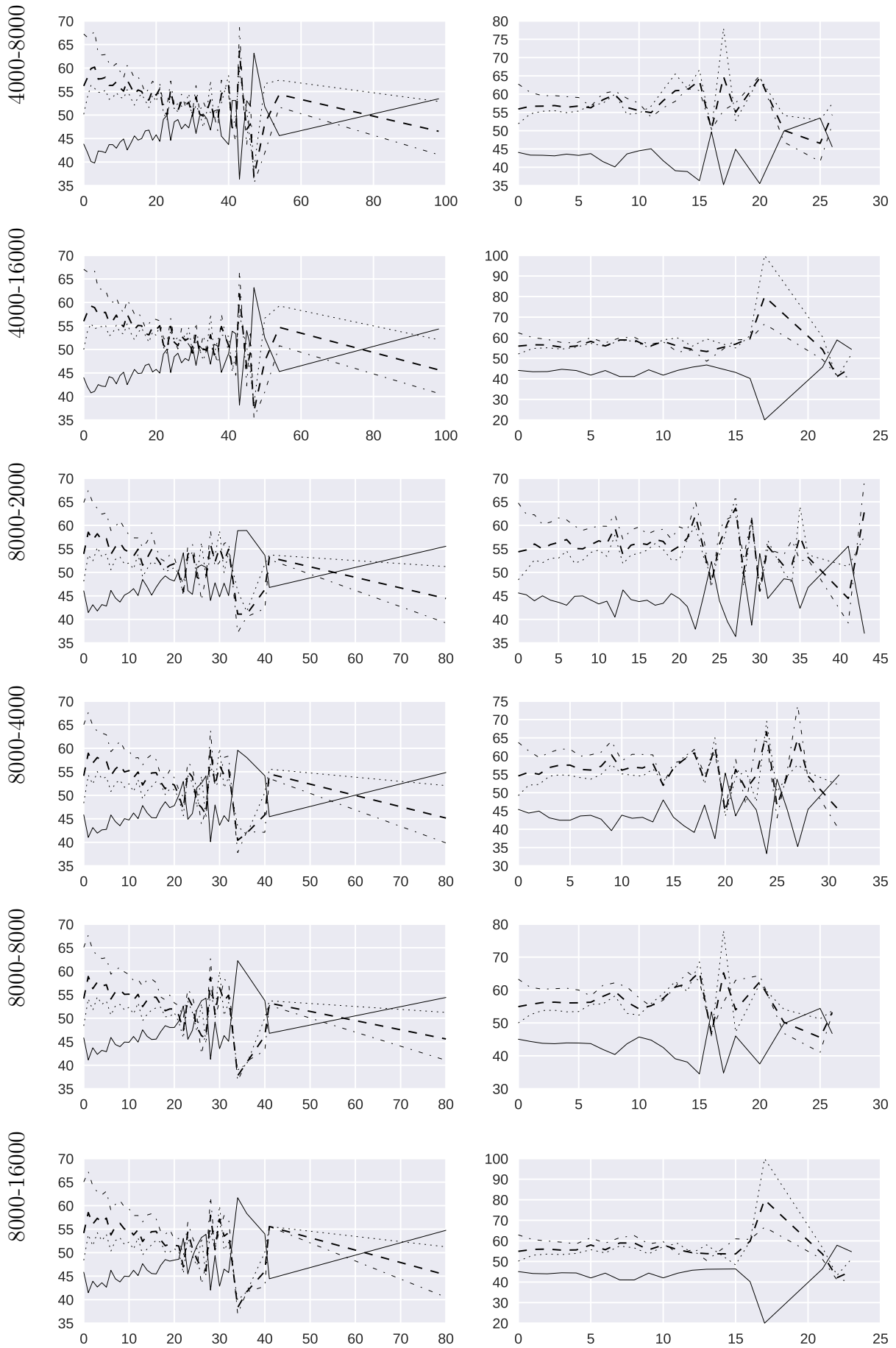


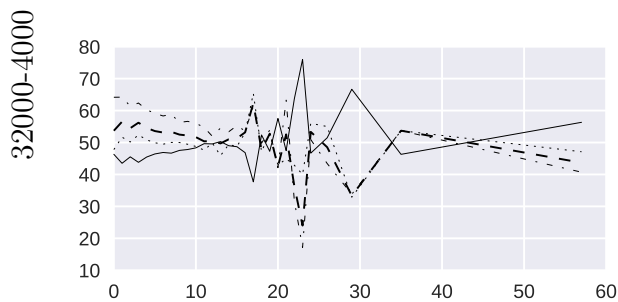
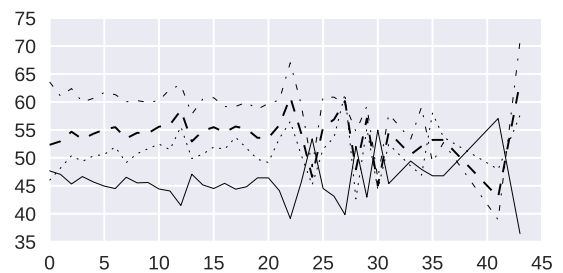
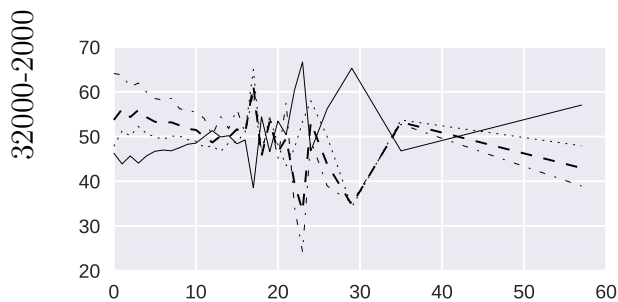
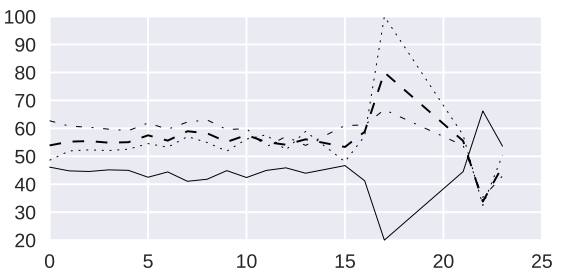
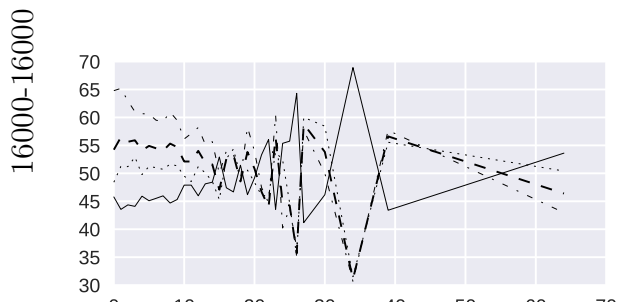
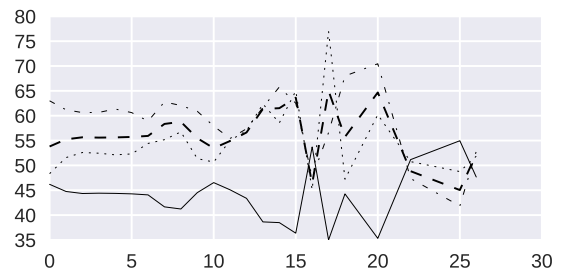
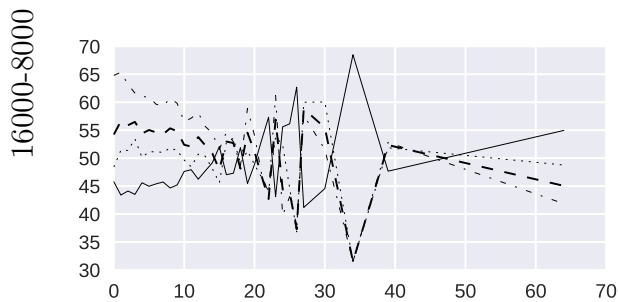
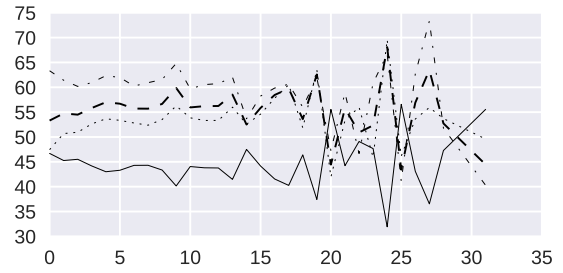
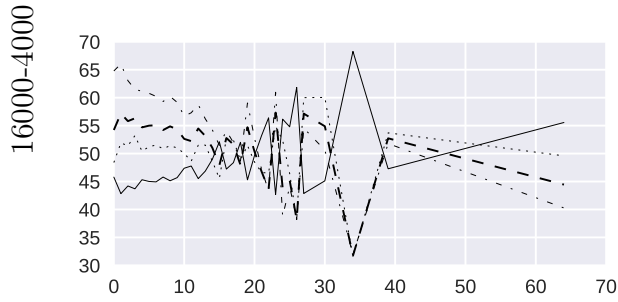
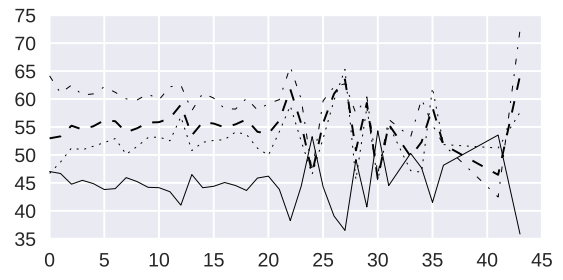
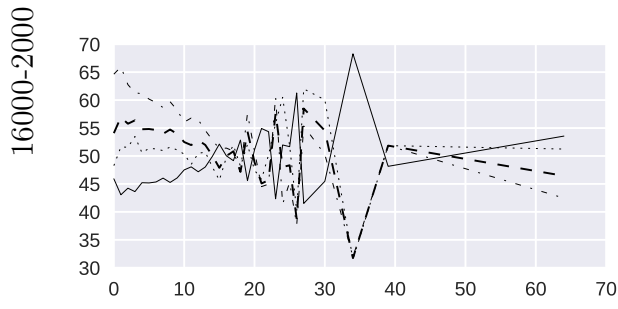
Figure E.148: BPE-based *Fastalign* for the direction Japanese-English: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.

Source

Target







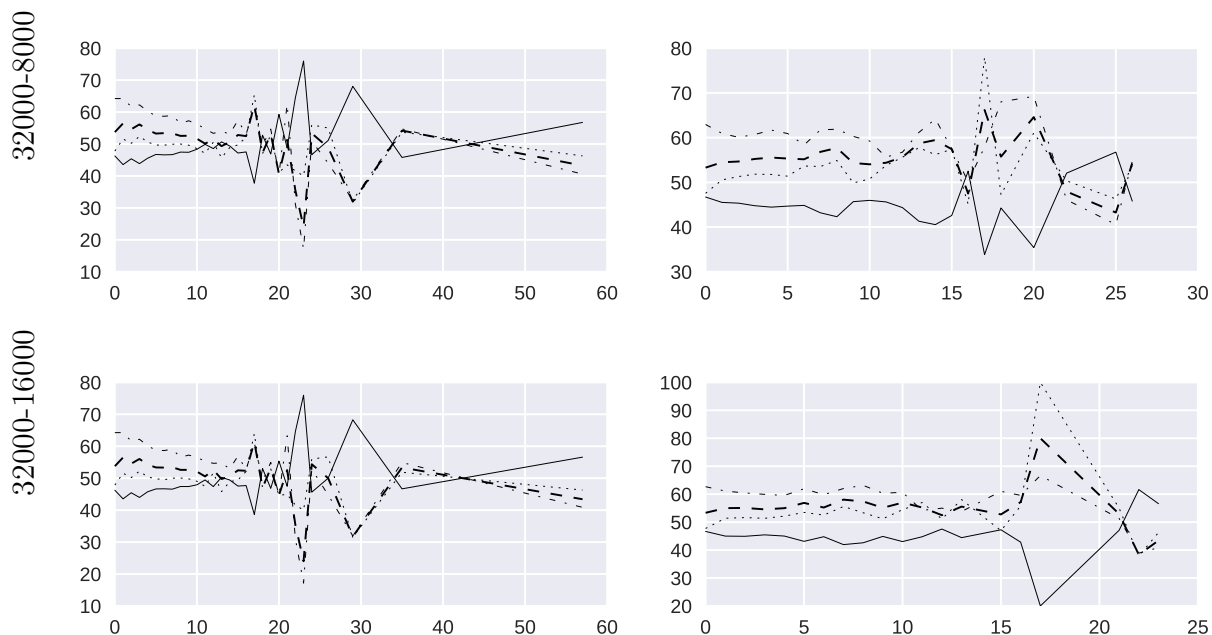
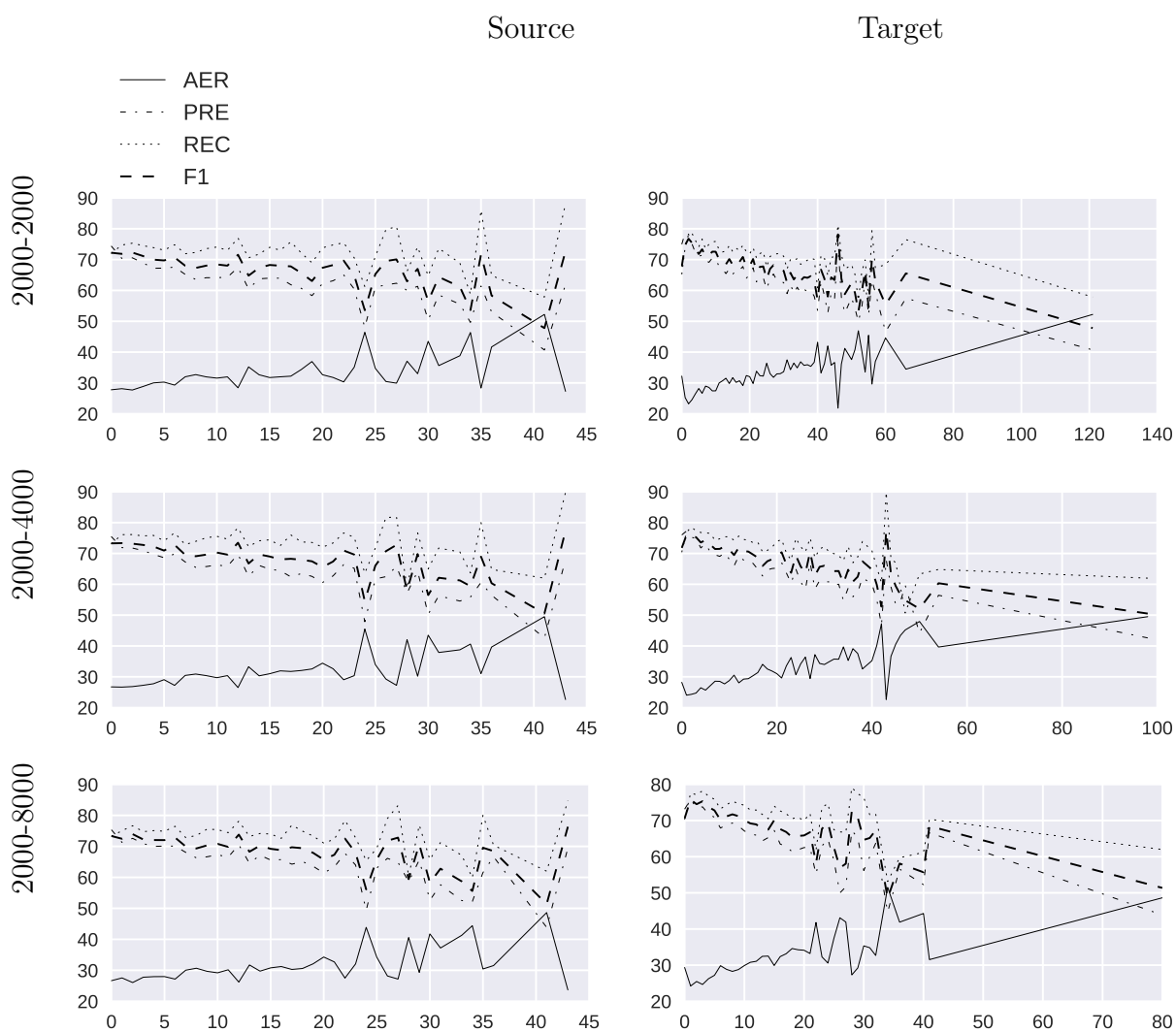
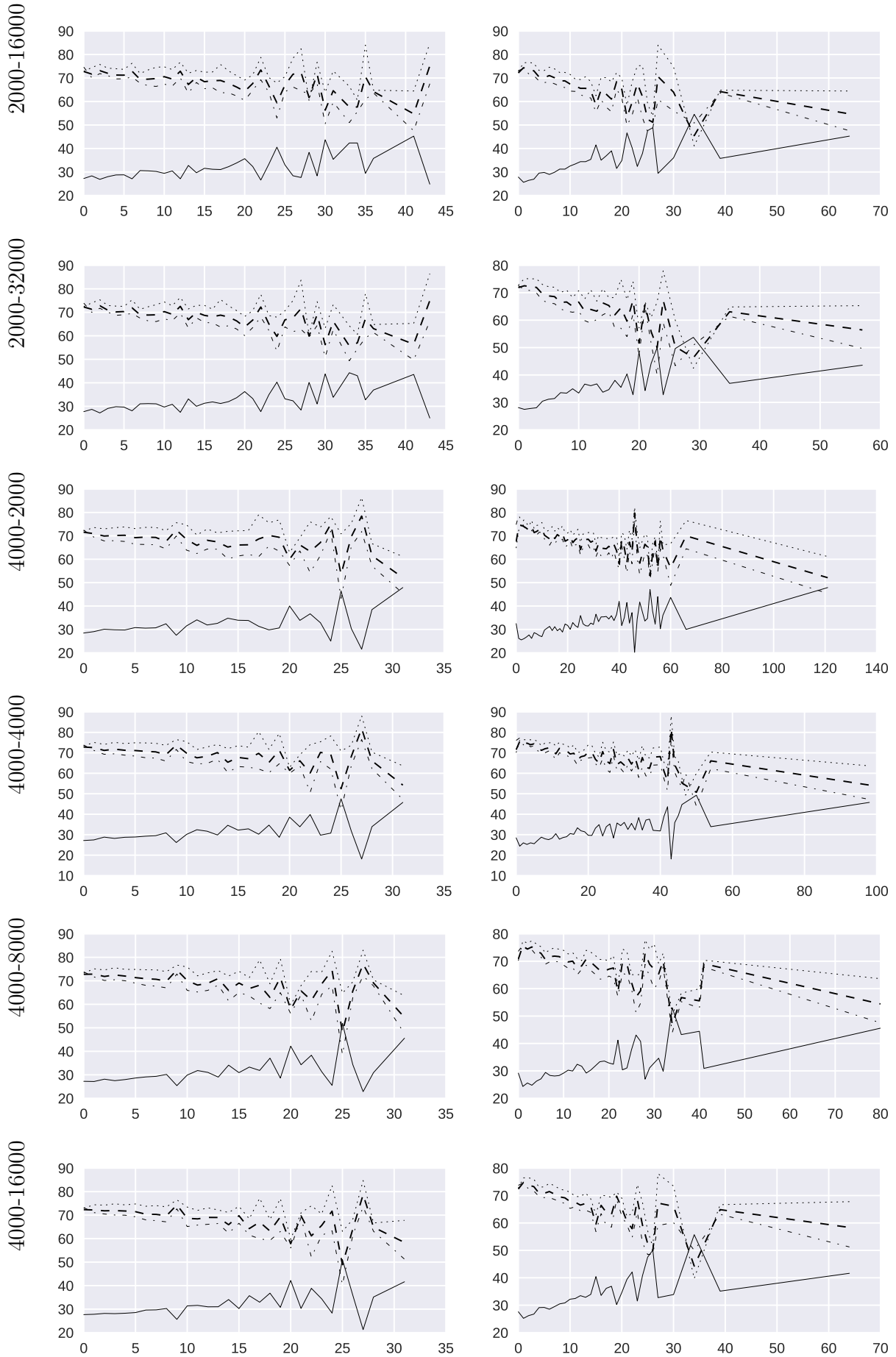
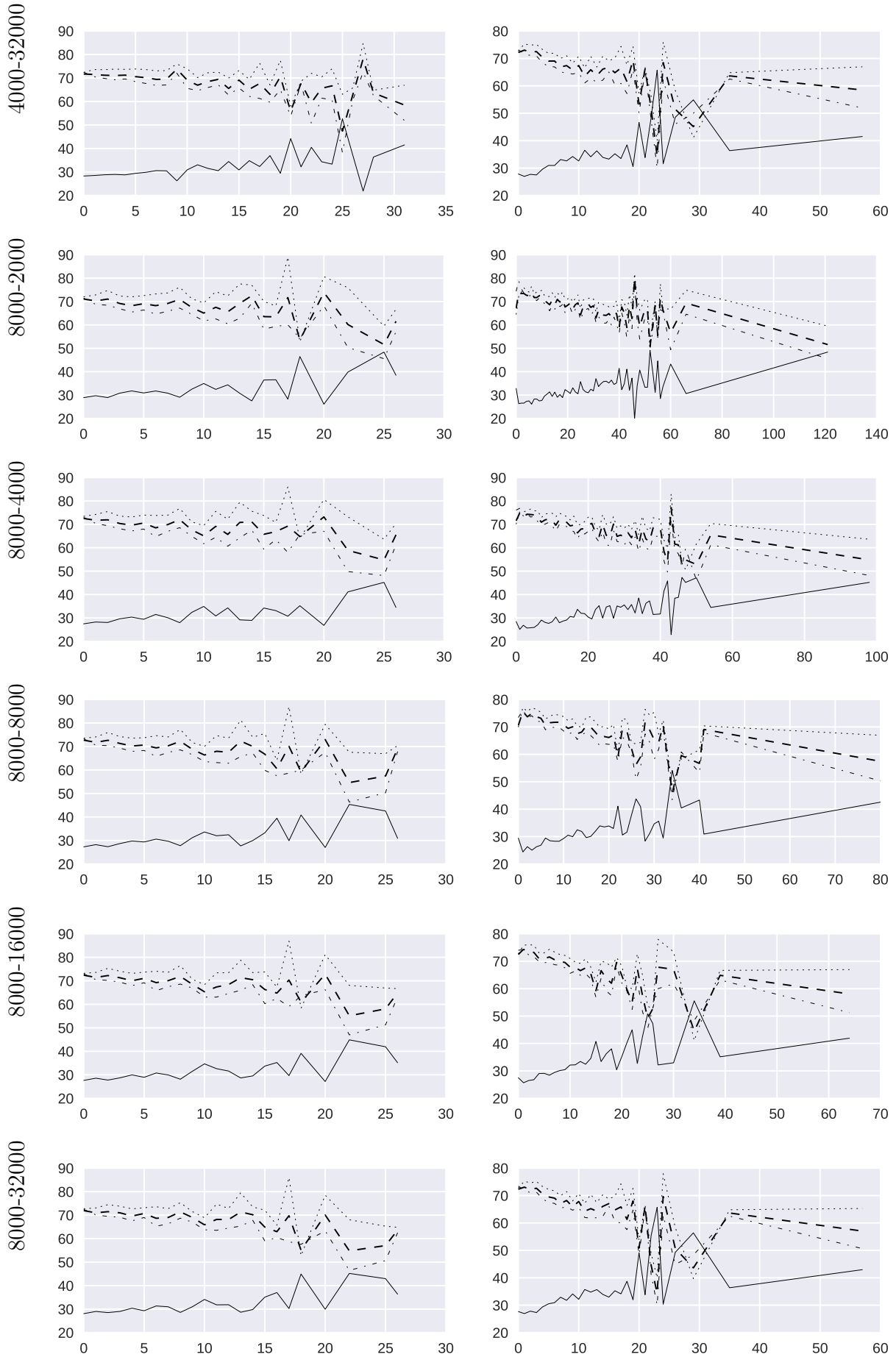


Figure E.149: BPE-based **Fastalign** for the direction English-Vietnamese: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.







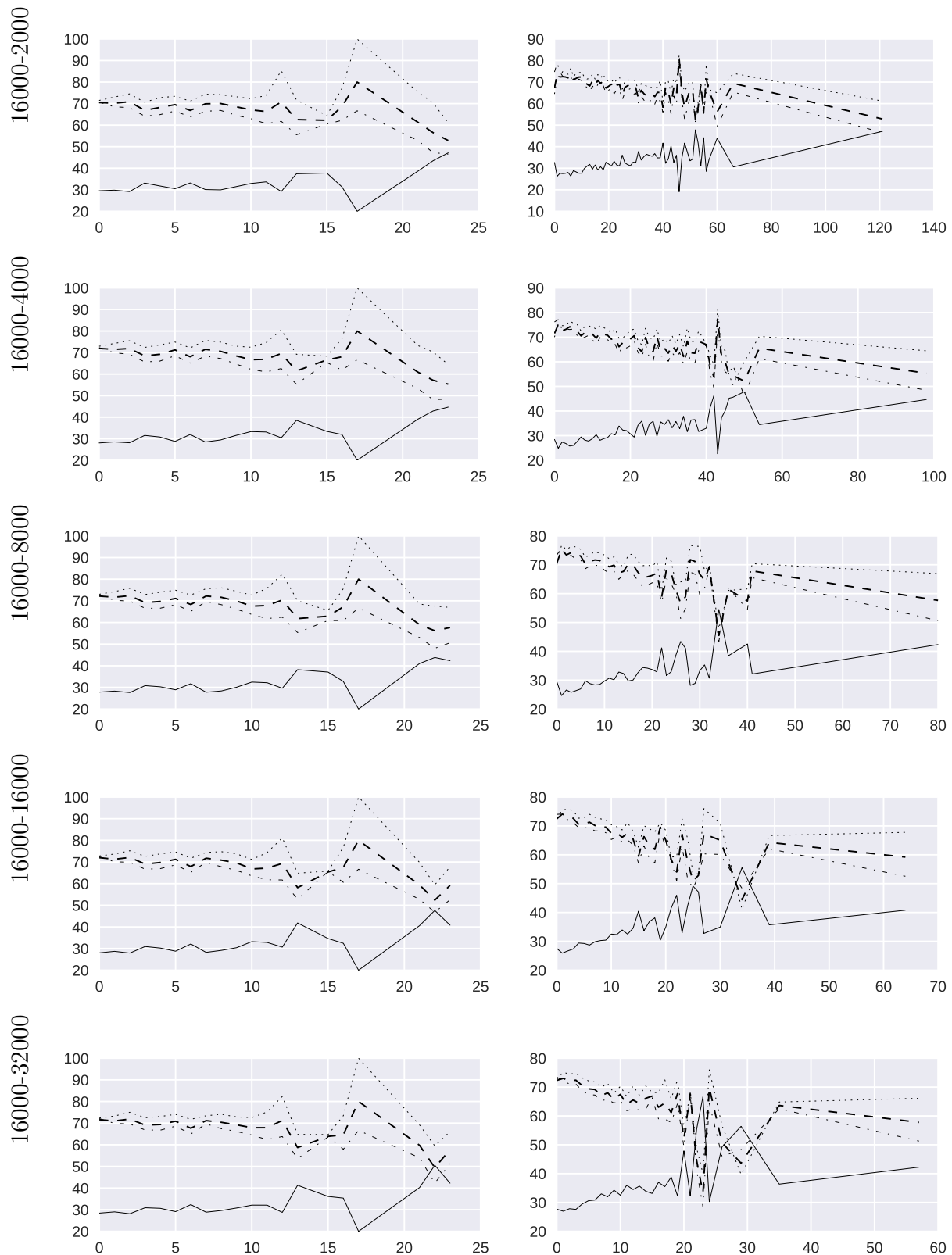


Figure E.150: BPE-based **Fastalign** for the direction Vietnamese-English: Alignment error rate (AER), F-score (F1), precision (PRE) and recall (REC) as a function of length difference. To obtain a length differences, we subtract a word-based sentence length from a BPE-based sentence length.

E.11 AER as a function of sentence length difference

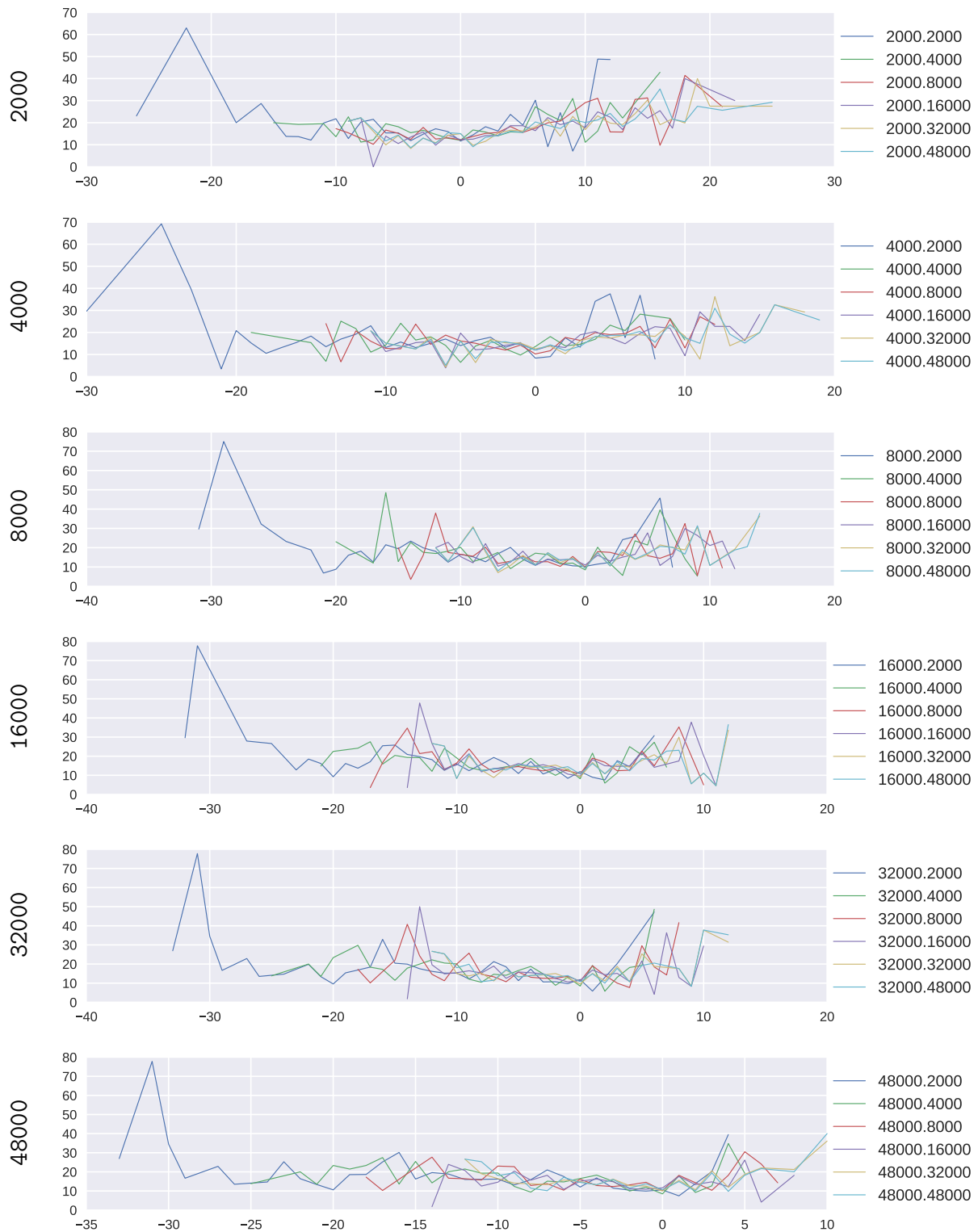


Figure E.151: The direction English-French: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.

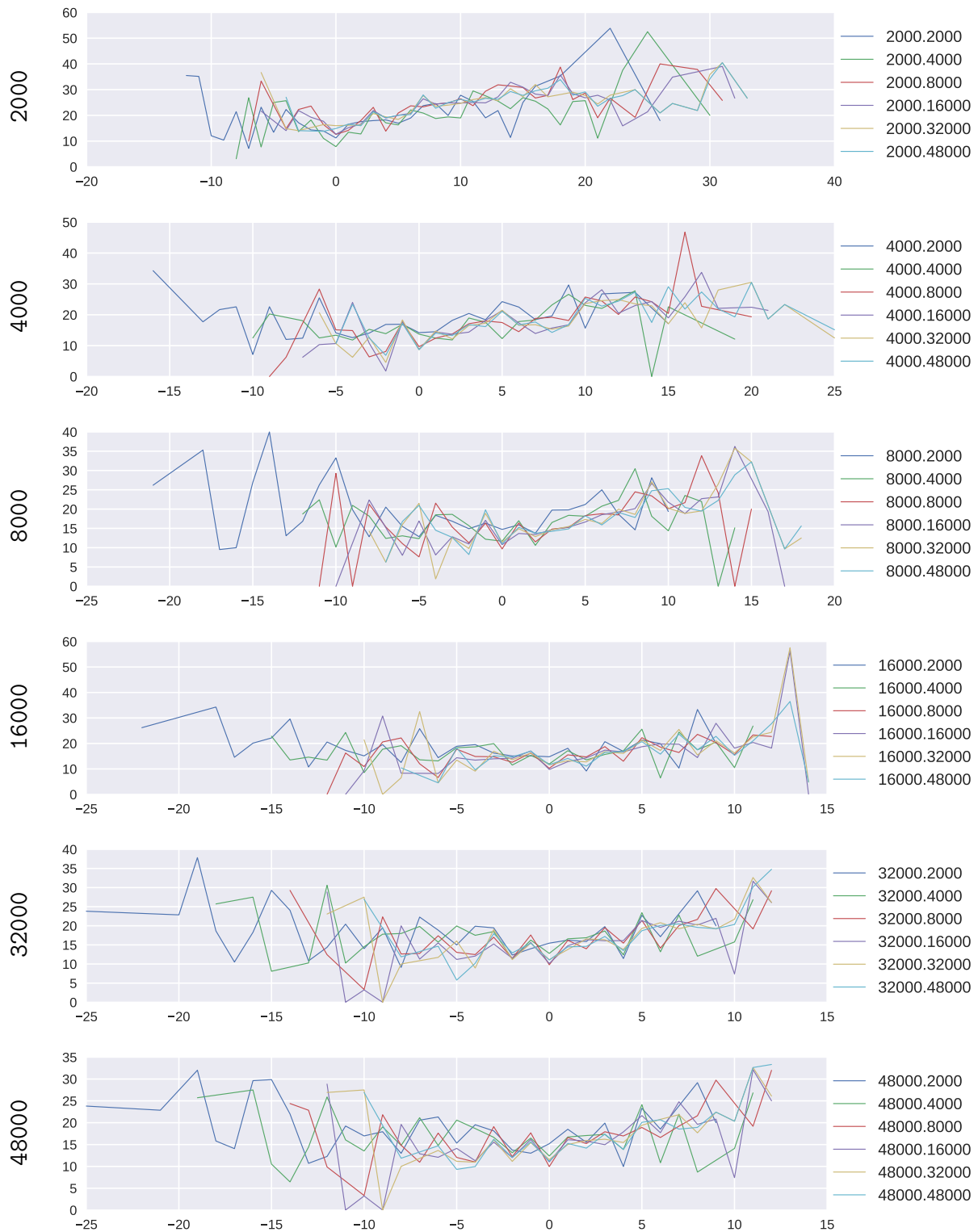


Figure E.152: The direction French-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.



Figure E.153: The direction English-German: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.

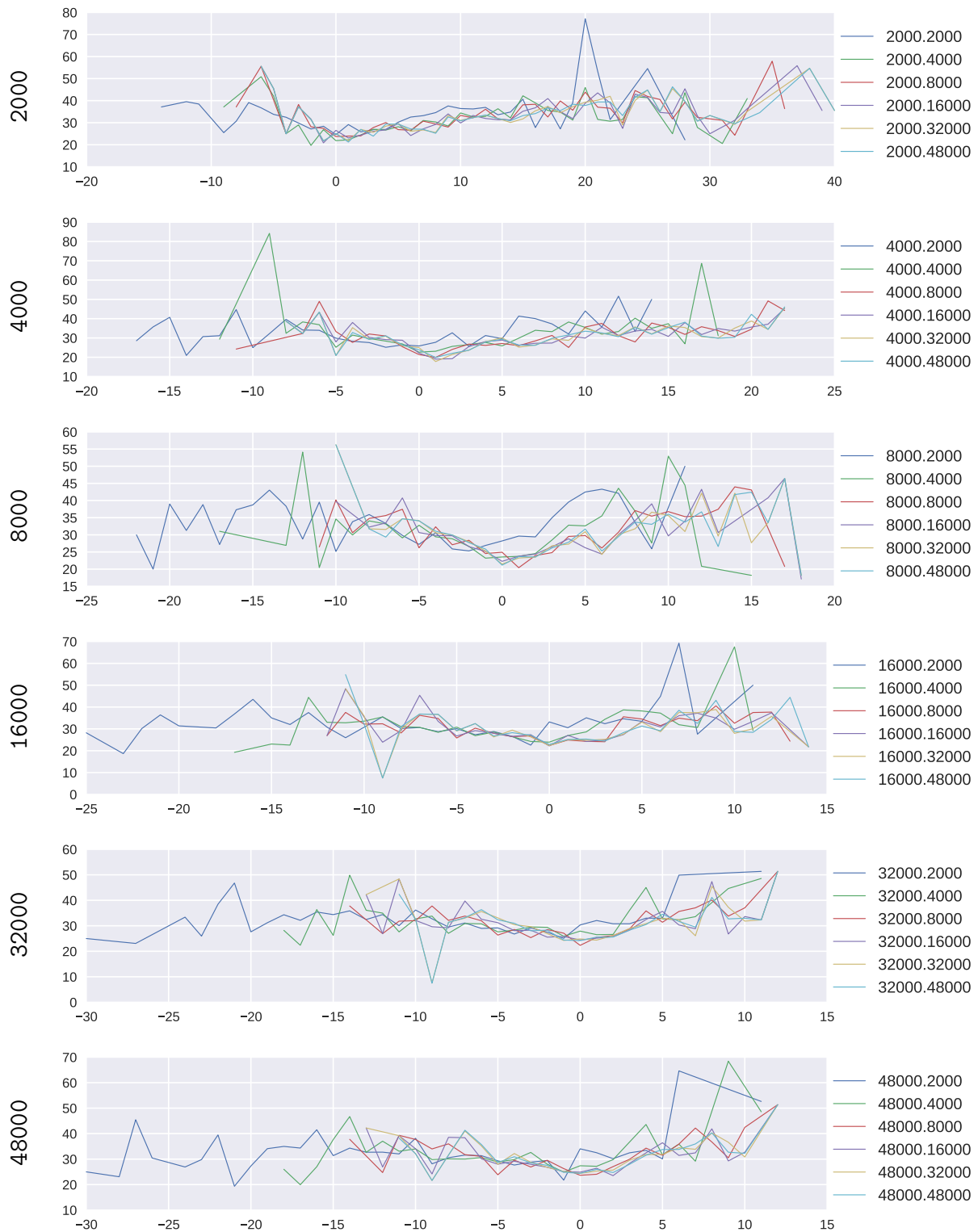


Figure E.154: The direction German-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.

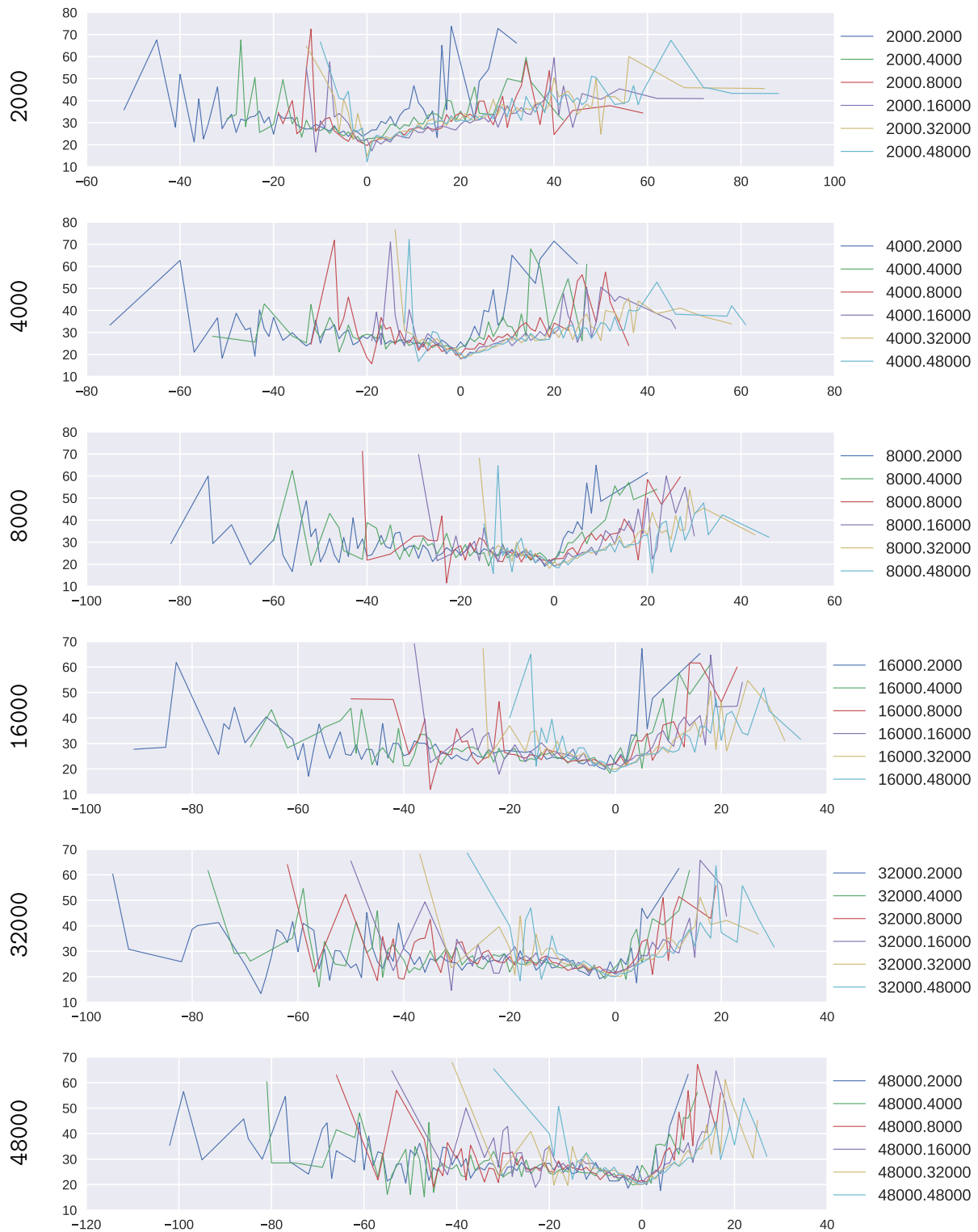


Figure E.155: The direction English-Czech: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.

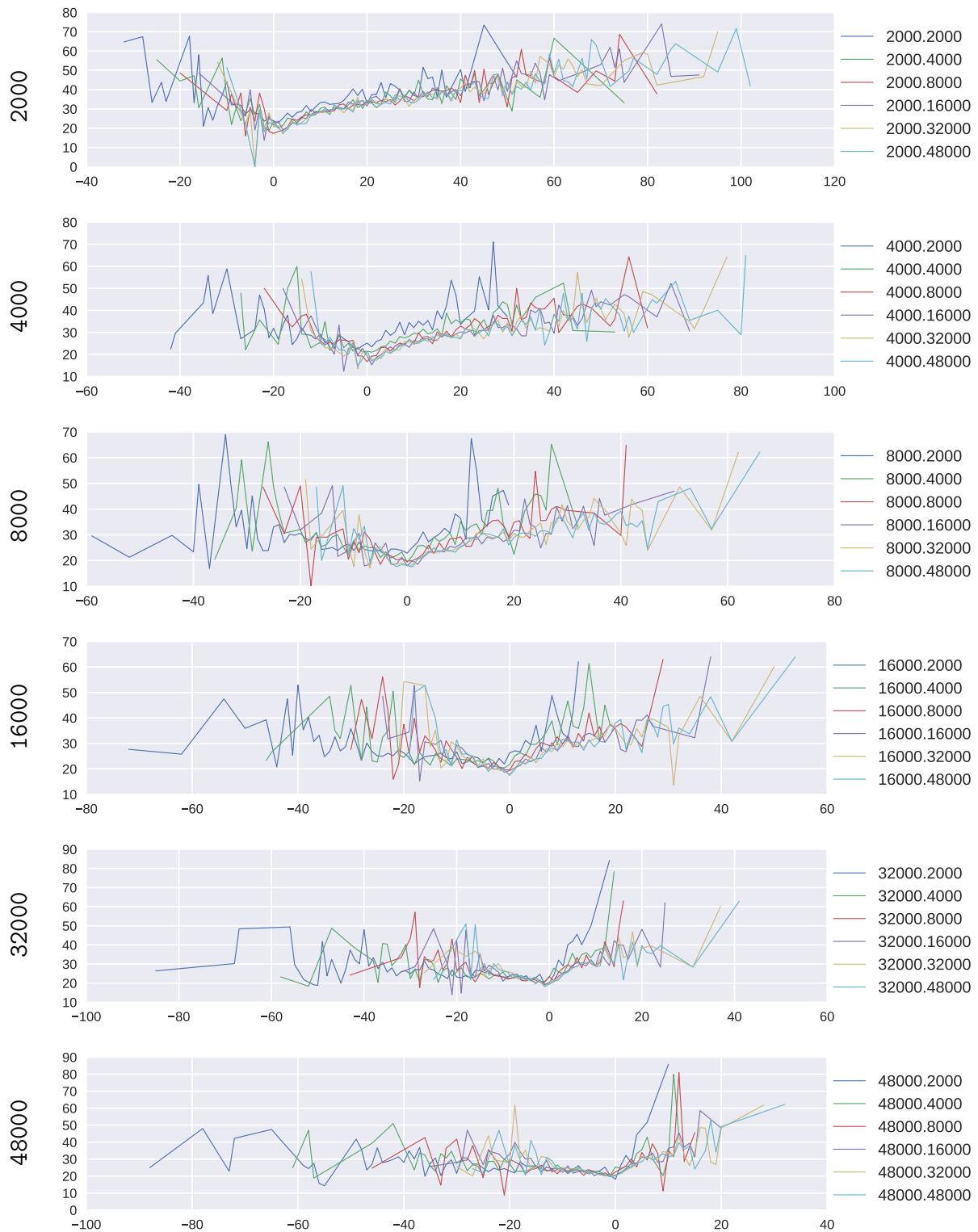


Figure E.156: The direction Czech-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.

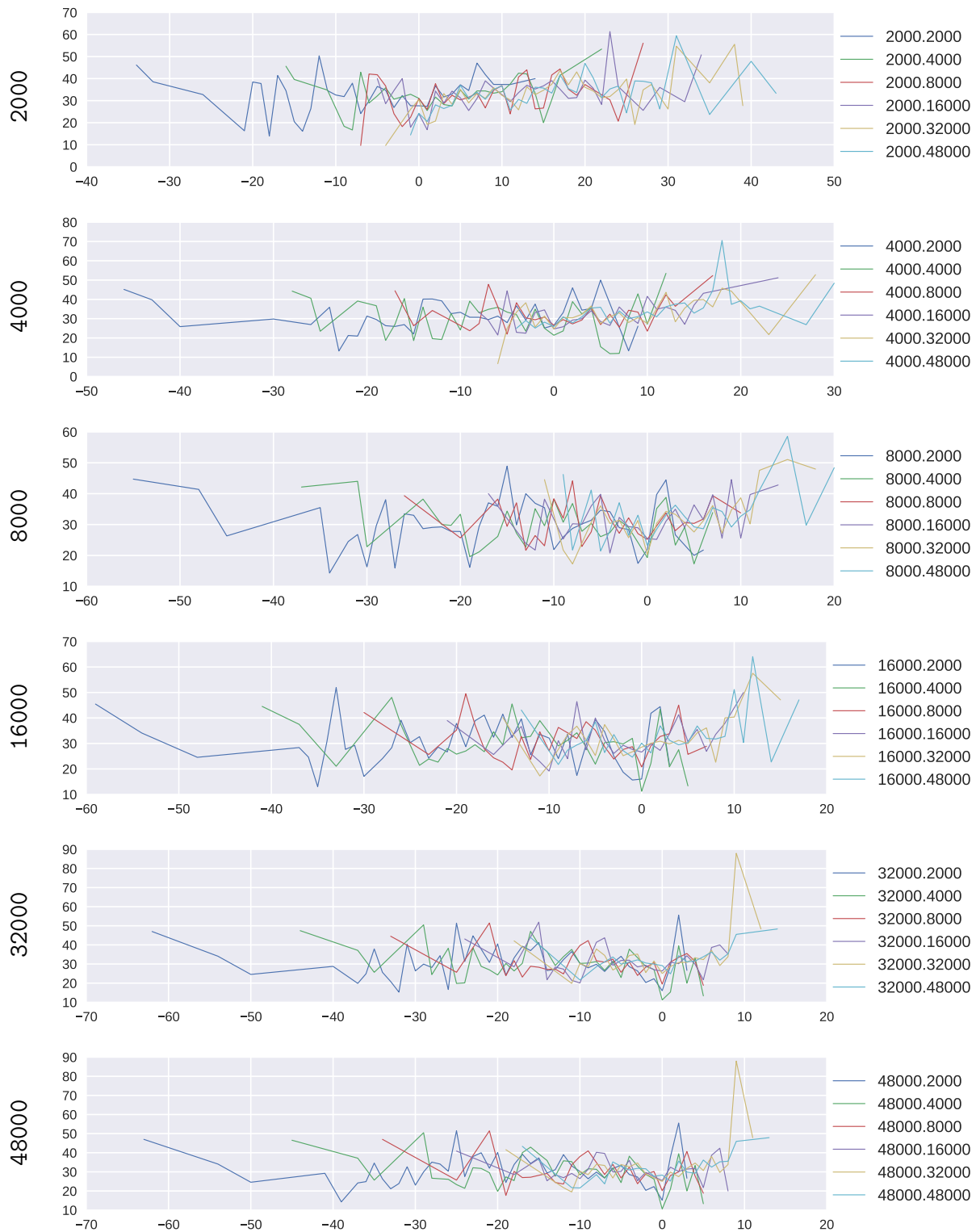


Figure E.157: The direction English-Romanian: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.

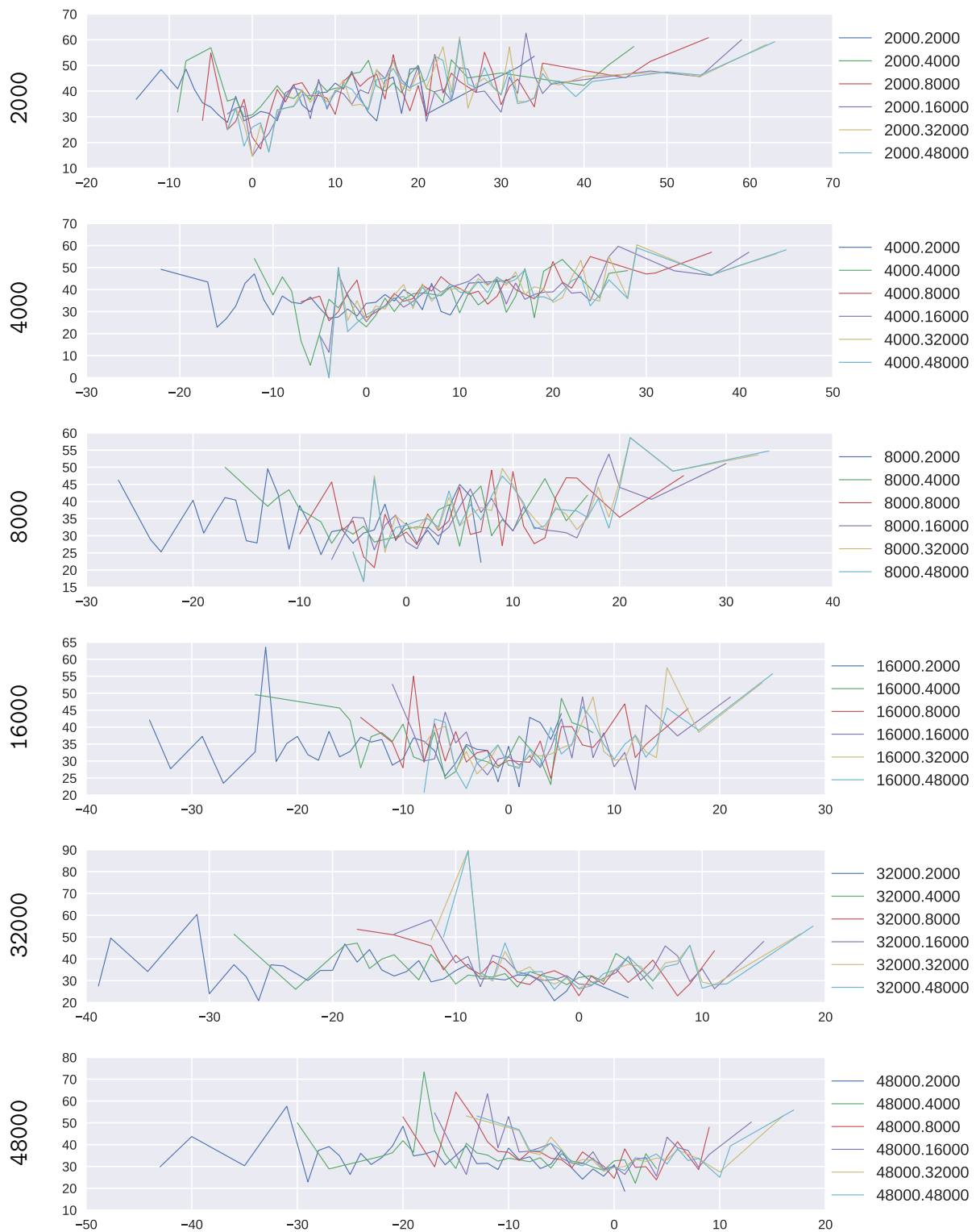


Figure E.158: The direction Romanian-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.



Figure E.159: The direction English-Japanese: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.

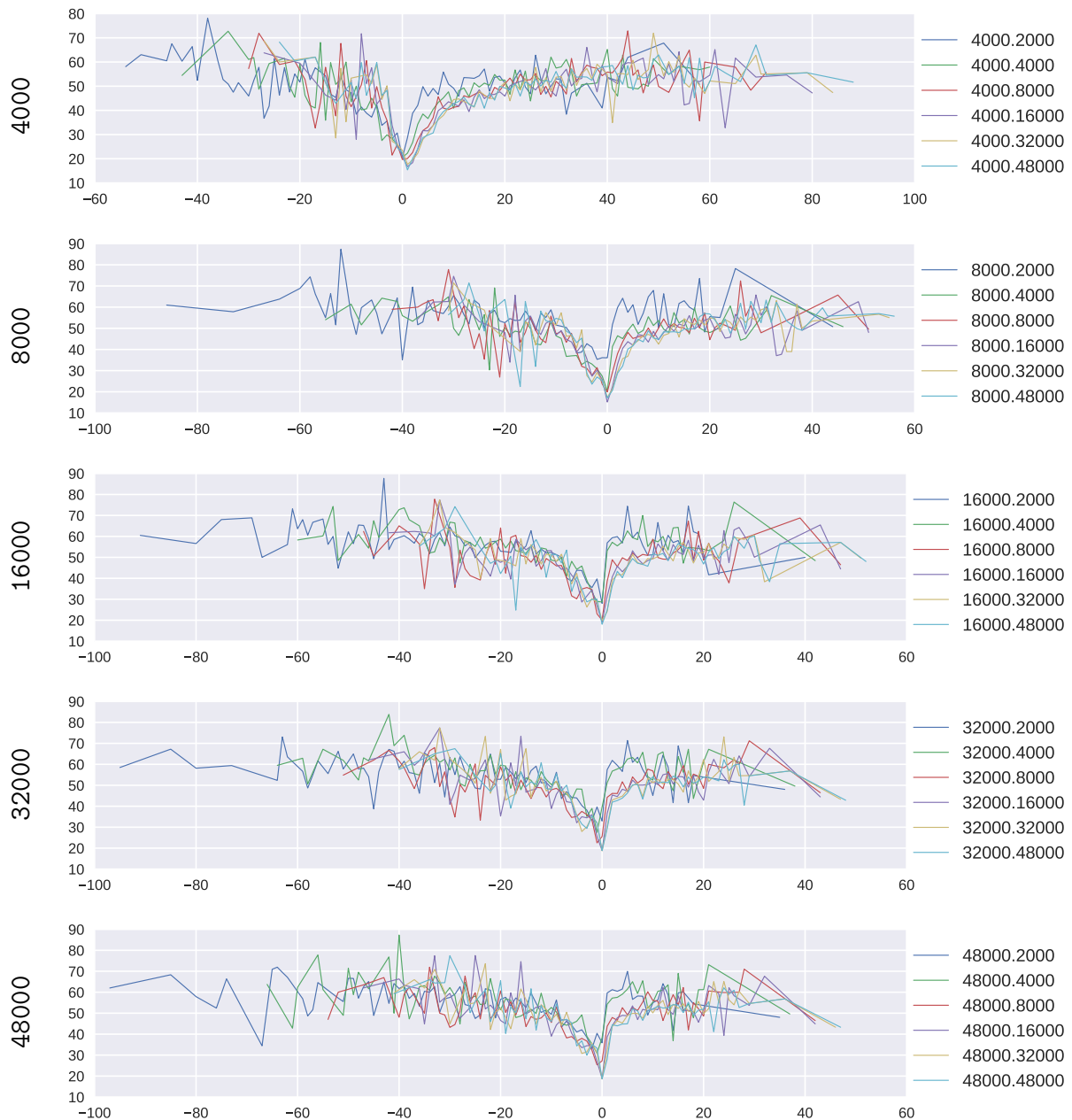


Figure E.160: The direction Japanese-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.

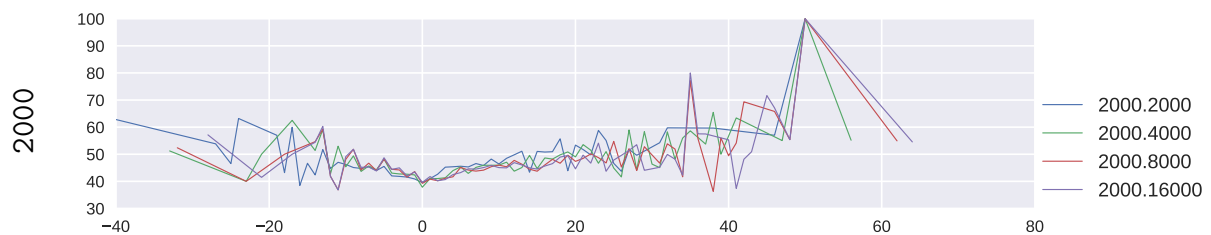




Figure E.161: The direction English-Vietnamese: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.



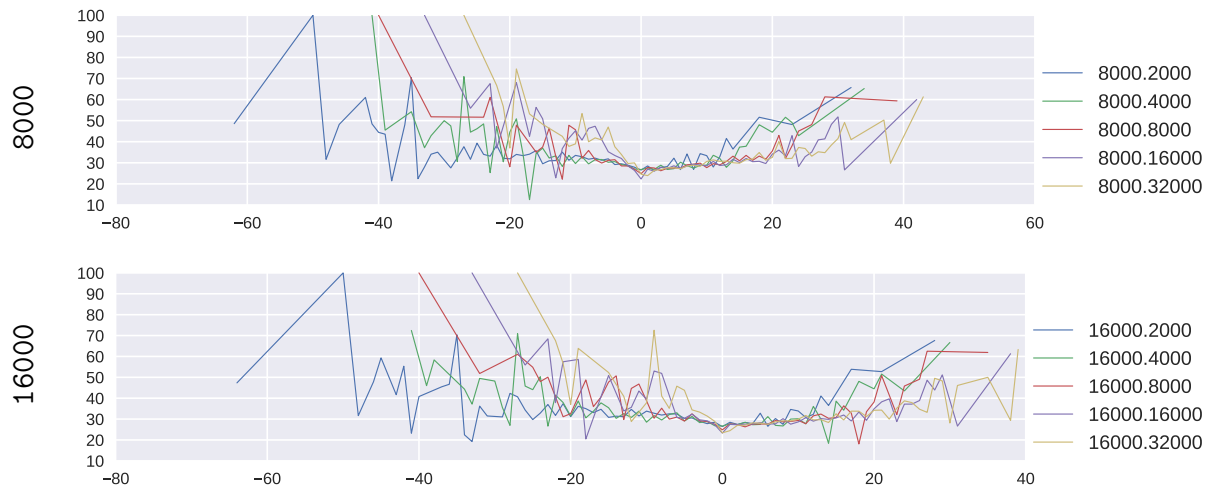


Figure E.162: The direction Vietnamese-English: AER score as a function of sentence length difference. The x-axis shows the sentence length difference. The y-axis represents the AER. The difference is computed by subtracting the length of the target sentence from the length of the source sentence.

Appendix F

Appendix: Our models versus models from related works

Models	English-Foreign		Foreign-English		Bi-direction	
	AER	F1	AER	F1	AER	F1
HMM Giza++ [Liang et al., 2006]	11.5		11.5		7.0	
IBM-4 Giza++ [Liang et al., 2006]	8.9		9.7		6.9	
HMM, joint [Liang et al., 2006]	6.1		6.6		4.9	
$FFNN_{u+c}$ [Tamura et al., 2014]						91.4
f2c [Kamigaito et al., 2014]					7.10	78.4
HMM+sym [Kamigaito et al., 2016]					5.97	79.2
NNSA [Legrand et al., 2016]	11.6	71.6	11.6	72.7	9.7	71.4
EMBALIGN [Rios et al., 2018]	29.43					
Add + SGD [Zenkel et al., 2019]	23.8		20.5		10.0	
FConv SmoothGrad [Ding et al., 2019]	21.2		18.1		8.5	
Multi-task + GIZA++ supervised [Garg et al., 2019]					4.6	
Nagata et al. [2020] (Supervised learning)					33.3	66.6
HMM+NN	12.44	45.21	11.32	46.75	10.19	47.78
HMM+NNCtxCNN	10.56	45.84	11.88	46.9	10.22	48.7
HMM+NNCharTgt	10.07	46.73	9.84	48	8.91	48.6
HMM+NNCharWord	10.43	46.77	9.41	48.77	8.94	49.43
HMM+NNCharJT	8.8	47.02	8.31	48.36	7.66	48.84
HMM+NNCharJB	8.78	47.14	7.27	46.36	6.89	47.7

Table F.1: Alignment error rate (AER) and F-score (F1) for English-French. Note that the training corpus of these experiments is the dataset Hansards which is different from the training corpus used in our analysis. For our models, bi-direction results are generated by using grow-diag-final.

Models	English-Foreign		Foreign-English		Bi-direction	
	AER	F1	AER	F1	AER	F1
Efmara1 (Word) [Östling and Tiedemann, 2016]					23	77
Add + SGD Zenkel et al. [2019]	30.4		26.6		21.2	
Multi-task + GIZA++ supervised [Garg et al., 2019]	19.8		17.5		16.4	
FConv SmoothGrad [Ding et al., 2019]	27.3		33.0		22.3	
Simalign (Subword, XLM-R - Argmax) [Sabet et al., 2020]					19	81
Nagata et al. [2020] (Supervised learning)					14.4	85.6
HMM+NN	26.78	70.95	29.44	68.21	26.92	71
HMM+NNCtxCc	34.36	63.19	32.31	65.34	31.72	65.94
HMM+NNCtxCNN	30.64	67.18	30.34	67.41	29.38	68.56
HMM+NNCharTgt	26.04	71.57	28.11	69.48	26.26	71.54
HMM+NNCharWord	24.98	72.64	29.77	67.76	26.35	71.43
HMM+NNCharBoth	27.14	70.6	29.31	68.34	27.86	70.04
HMM+NNCharJT	23.79	73.15	25.21	71.85	21.76	75.27
HMM+NNCharJB	23.69	73.38	24.9	72.16	22.42	74.71

Table F.2: Alignment error rate (AER) and F-score (F1) for English-German. For our models, bi-direction results are generated by using grow-diag-final.

Models	English-Foreign		Foreign-English		Bi-direction	
	AER	F1	AER	F1	AER	F1
Add + SGD Zenkel et al. [2019]	34.8		32.3		27.6	
Multi-task + GIZA++ supervised [Garg et al., 2019]					23.1	
FConv SmoothGrad [Ding et al., 2019]	34.2		33.4		27.2	
Nagata et al. [2020] (Supervised learning)					12.2	86.7
HMM+NN	32.73	67.29	37.81	62.2	32.3	67.71
HMM+NNCtxCNN	31.34	68.68	32.24	67.78	29.18	70.84
HMM+NNCharTgt	27.52	72.5	29.61	70.41	26.98	73.04
HMM+NNCharWord	26.4	73.62	29.95	70.07	26.9	73.12
HMM+NNCharJT	28.39	71.63	28.25	71.77	26.93	73.09
HMM+NNCharJB	25.62	74.4	28.3	71.72	25.39	74.63

Table F.3: Alignment error rate (AER) and F-score (F1) for English-Romanian. Note that the training corpus of these experiments is the NAACL 2003 corpus [Mihalcea and Pedersen, 2003] which is different from the training corpus used in our analysis. For our models, bi-direction results are generated by using grow-diag-final.

Models	English-Foreign		Foreign-English		Bi-direction	
	AER	F1	AER	F1	AER	F1
HMT [Kondo et al., 2013]					37.5	
f2f [Kamigaito et al., 2014]					47.7	52.2
IBM-4+itg [Kamigaito et al., 2016]					43.8	56.1
Nagata et al. [2020] (Supervised learning)					22.4	77.6
HMM+NN	49.68	50.32	47.09	52.91	35.97	64.03
HMM+NNCtxCc	47.93	52.07	45.67	54.33	34.87	65.13
HMM+NNCtxCNN	46.82	53.18	44.79	55.21	32.06	67.94
HMM+NNCharTgt	43.92	56.08	43.6	56.4	31.55	68.45
HMM+NNCharWord	45.35	54.65	40.16	59.84	30.71	69.29
HMM+NNCharBoth	43.71	56.29	40.26	59.74	30.54	69.46
HMM+NNCharJT	39.69	60.31	39.22	60.78	24.92	75.08
HMM+NNCharJB	40.92	59.08	37.71	62.29	25.52	74.48

Table F.4: Alignment error rate (AER), F-score (F1) for English-Japanese. For our models, bi-direction results are generated by using grow-diag-final.

Bibliography

- Shuoyang Ding, Hainan Xu, and Philipp Koehn. Saliency-driven word alignment interpretation for neural machine translation. In Proceedings of the Fourth Conference on Machine Translation (Volume 1: Research Papers), pages 1–12, Florence, Italy, August 2019. Association for Computational Linguistics. doi: 10.18653/v1/W19-5201. URL <https://www.aclweb.org/anthology/W19-5201>.
- Sarthak Garg, Stephan Peitz, Udhyakumar Nallasamy, and Matthias Paulik. Jointly learning to align and translate with transformer models. In Proc. IJCNLP-EMNLP, Hong Kong, China, November 2019. doi: 10.18653/v1/D19-1453. URL <https://www.aclweb.org/anthology/D19-1453>.
- Hidetaka Kamigaito, Taro Watanabe, Hiroya Takamura, and Manabu Okumura. Unsupervised word alignment using frequency constraint in posterior regularized EM. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 153–158, Doha, Qatar, October 2014. Association for Computational Linguistics. doi: 10.3115/v1/D14-1017. URL <https://www.aclweb.org/anthology/D14-1017>.
- Hidetaka Kamigaito, Akihiro Tamura, Hiroya Takamura, Manabu Okumura, and Eiichiro Sumita. Unsupervised word alignment by agreement under ITG constraint. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pages 1998–2004, Austin, Texas, November 2016. Association for Computational Linguistics. doi: 10.18653/v1/D16-1210. URL <https://www.aclweb.org/anthology/D16-1210>.
- Shuhei Kondo, Kevin Duh, and Yuji Matsumoto. Hidden markov tree model for word alignment. In Proceedings of the Eighth Workshop on Statistical Machine Translation, Sofia, Bulgaria, August 2013. Association for Computational Linguistics.
- Joël Legrand, Michael Auli, and Ronan Collobert. Neural network-based word alignment through score aggregation. In Proceedings of the First Conference on Machine Translation, pages 66–73, Berlin, Germany, August 2016. Association for Computational Linguistics.
- Percy Liang, Ben Taskar, and Dan Klein. Alignment by agreement. In Proceedings of the Main Conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics, HLT-NAACL '06, pages 104–111, Stroudsburg, PA, USA, 2006. Association for Computational Linguistics.
- Rada Mihalcea and Ted Pedersen. An evaluation exercise for word alignment. In Proceedings of the HLT-NAACL 2003 Workshop on Building and Using Parallel Texts: Data Driven Machine Translation and Beyond - Volume 3, HLT-NAACL-PARALLEL '03, pages 1–10, Stroudsburg, PA, USA, 2003. Association for Computational Linguistics. doi: 10.3115/1118905.1118906.
- Masaaki Nagata, Chousa Katsuki, and Masaaki Nishino. A supervised word alignment method based on cross-language span prediction using multilingual BERT. CoRR, abs/2004.14516, 2020. URL <https://arxiv.org/abs/2004.14516>.

- Anh Khoa Ngo Ho. Generative Probabilistic Alignment Models for Words and Subwords: a Systematic Exploration of the Limits and Potentials of Neural Parametrizations. PhD thesis, 2021.
- Robert Östling and Jörg Tiedemann. Efficient word alignment with Markov Chain Monte Carlo. *Prague Bulletin of Mathematical Linguistics*, 106:125–146, October 2016. URL <http://ufal.mff.cuni.cz/pbml/106/art-ostling-tiedemann.pdf>.
- Miguel Rios, Wilker Aziz, and Khalil Sima'an. Deep generative model for joint alignment and word representation. In Marilyn A. Walker, Heng Ji, and Amanda Stent, editors, *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2018, New Orleans, Louisiana, USA, June 1-6, 2018, Volume 1 (Long Papers)*, pages 1011–1023. Association for Computational Linguistics, 2018. doi: 10.18653/v1/n18-1092. URL <https://doi.org/10.18653/v1/n18-1092>.
- Masoud Jalili Sabet, Philipp Dufter, and Hinrich Schütze. Simalign: High quality word alignments without parallel training data using static and contextualized embeddings. *CoRR*, abs/2004.08728, 2020. URL <https://arxiv.org/abs/2004.08728>.
- Akihiro Tamura, Taro Watanabe, and Eiichiro Sumita. Recurrent neural networks for word alignment model. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1470–1480, Baltimore, Maryland, June 2014. Association for Computational Linguistics.
- Thomas Zenkel, Joern Wuebker, and John DeNero. Adding interpretable attention to neural translation models improves word alignment. *CoRR*, abs/1901.11359, 2019. URL <http://arxiv.org/abs/1901.11359>.