		System output		
		Exist		None
		Correct	Wrong	
Gold -standard	Exist	140	6	112
	None	-	38	704

Table 6: Result of the author mention detection

		System output		
		Exist		None
		Correct	Wrong	
Gold -standard	Exist	56	2	47
	None	-	23	872

Table 7: Result of the reader mention detection

are given by the syntactic parser KNP.[9]

7.2 Results of Author/Reader Mention Detection

We show the results of the author and reader mention detection in Table 6 and Table 7. In these tables, "exist" indicates numbers of documents in which the A/R mentions are manually annotated or our system estimated that some discourse entities are A/R mentions. From these results, the A/R mentions including "none" can be predicted to accuracies of approximately 80%. On the other hand, the recalls are not particularly high: the recall of author is 140/258 and the recall of reader is 56/105. This is because the documents in which the A/R do not appear are more than the ones in which the A/R appear and the system prefers to output "no author/reader mention" as the result of training.

7.3 Results of Zero Reference Resolution

We show the results of zero reference resolution in Table 8 and Table 9. The difference between the baseline and the proposed model is statistically significant ($p < 0.05$) from the McNemar's test. In Table 8, we evaluate only the zero endophora for comparison to the baseline model, which deals with only the zero endophora. "Proposed model (estimate)" shows the result of the proposed model which estimated the A/R mentions and "Proposed model (gold-standard)" shows the result of the proposed model which is given the A/R mentions of gold-standard from the corpus.

From Table 8, considering the zero exophora and the A/R mentions improves accuracy of zero endophora resolution as well as zero reference resolution including zero exophora.

From Table 8 and Table 9, the proposed model given the gold-standard A/R mentions achieves extraordinarily high accuracies. This result indicates that improvement of the A/R mention detection improves the accuracy of zero reference resolution in the proposed model.

[9] http://nlp.ist.i.kyoto-u.ac.jp/EN/index.php?KNP

	Recall	Precision	F1
Baseline	0.269	0.377	0.314
Proposed model (estimate)	0.282	0.448	0.346
Proposed model (gold-standard)	0.388	0.522	0.445

Table 8: Results of zero endophora resolution

	Recall	Precision	F1
Baseline	0.115	0.377	0.176
Proposed model (estimate)	0.317	0.411	0.358
Proposed model (gold-standard)	0.377	0.485	0.424

Table 9: Results of zero reference resolution

8 Conclusion

This paper presented a zero reference resolution model considering exophora and author/reader mentions. In the experiments, our proposed model achieves higher accuracy than the baseline model. As future work, we plan to improve the author/reader detection model to improve the zero reference resolution.

References

Masatsugu Hangyo, Daisuke Kawahara, and Sadao Kurohashi. 2012. Building a diverse document leads corpus annotated with semantic relations. In *Proceedings of the 26th Pacific Asia Conference on Language, Information, and Computation*, pages 535–544, Bali,Indonesia, November. Faculty of Computer Science, Universitas Indonesia.

Ralf Herbrich, Thore Graepel, Peter Bollmann-Sdorra, and Klaus Obermayer. 1998. Learning preference relations for information retrieval. In *ICML-98 Workshop: text categorization and machine learning*, pages 80–84.

Ryu Iida, Kentaro Inui, and Yuji Matsumoto. 2006. Exploiting syntactic patterns as clues in zero-anaphora