

Table 3.1. Contrastive analysis of coreference annotation schemata.

Reference	Language	Size	Underlying syntactic annotation	Mention scope					Head Markup	Relations	Inter-annotator Agreement
				Zero Subject	Nested NGs	Split NGs	Attributive NGs	Disc. Deixis			
Osenova and Simov (2004)	Bulgarian	11K sentences	✓	✓	✓	?	?	✓		identity, part-of, subset-of	unknown
Hinrichs et al. (2005)	German	22–85K sentences	✓		✓		✓		synt.	identity, anaphora, cataphora, bound rel., split antecedent, instance, expletive	unknown
LDC (2006)	Spanish	N/A			✓	✓	delimited, predicative-linked		synt.	identity, apposition	N/A
Magnini et al. (2006)	Italian	182K words			✓	✓	delimited, predicative-linked		synt.		unknown
Pradhan et al. (2007, 2012)	English, Chinese, Arabic	1.6M/950K /300K words	✓	✓	✓				synt.	identity, apposition	0.65–0.89 (MUC)
Iida et al. (2007)	Japanese	38K sentences		✓						identity, predicate-argument	$R = 0.893$ $P = 0.831$
Poesio and Artstein (2008)	English	94K words					✓	✓		identity, predicative link, discourse deixis, bridging relations	$\alpha = 0.6–0.7$

Table 3.1. Contrastive analysis of coreference annotation schemata (cont.).

Reference	Language	Size	Underlying syntactic annotation	Mention scope					Head Markup	Relations	Inter-Annotator Agreement
				Zero Subject	Nested NGs	Split NGs	Attributive NGs	Disc. Deixis			
Hendrickx et al. (2008)	Dutch	200K words				✓	✓	✓	synt.	identity, bound rel., bridging rel. predicative link	0.76 (MUC)
Nedoluzhko et al. (2009); Nedoluzhko (2010)	Czech	58K sentences	✓		partly	✓	partly	✓	synt., sem.?	identity, predic. link, discourse deixis with many sentences, bridging relations	$F_1 = 0.39-0.80$
Recasens and Martí (2010)	Spanish, Catalan	400K words	✓	✓	✓	✓	delimited, predicative-linked	✓		identity, predicative link, discourse deixis	$\alpha = 0.85-0.89$
Korzen and Buch-Kromann (2011)	Danish, English, German, Italian, Spanish	$5 \times 100K$ words	✓					✓		identity, discourse deixis bridging relations	25 – 100%
Muzerelle et al. (2013)	French	453K words			✓	✓	delimited, never linked			identity, associative coref.	unknown
Our approach	Polish	503K words		✓	✓	✓	delimited, never linked		sem.	identity, quasi-identity	$\kappa = 0.7424$