

INTER-ANNOTATOR AGREEMENT (I.E. ACROSS ANNOTATORS)

	ALL COMPARISONS			NO REF COMPARISONS		
	$P(A)$	$P(E)$	$\kappa$	$P(A)$	$P(E)$	$\kappa$
European languages, individual systems	0.601	0.362	0.375	0.561	0.355	0.320
European languages, system combinations	0.671	0.335	0.505	0.598	0.342	0.389
Haitian-English, individual systems	0.691	0.362	0.516	0.639	0.350	0.446
Haitian-English, system combinations	0.761	0.358	0.628	0.674	0.335	0.509
Tunable metrics task (Urdu-English)	0.692	0.337	0.535	0.641	0.363	0.437
WMT10 (European languages, all systems)	0.658	0.374	0.454	0.626	0.367	0.409

INTRA-ANNOTATOR AGREEMENT (I.E. SELF-CONSISTENCY)

	ALL COMPARISONS			NO REF COMPARISONS		
	$P(A)$	$P(E)$	$\kappa$	$P(A)$	$P(E)$	$\kappa$
European languages, individual systems	0.722	0.362	0.564	0.685	0.355	0.512
European languages, system combinations	0.787	0.335	0.680	0.717	0.342	0.571
Haitian-English, individual systems	0.763	0.362	0.628	0.700	0.350	0.539
Haitian-English, system combinations	0.882	0.358	0.816	0.784	0.335	0.675
Tunable metrics task (Urdu-English)	0.857	0.337	0.784	0.856	0.363	0.774
WMT10 (European languages, all systems)	0.755	0.374	0.609	0.734	0.367	0.580