

<b>Code-mixing Table</b>					<b>Code-mixing Types</b>				
<b>Language</b>	<b>Total Words</b>	<b>Unique Words</b>	<b>Percentage</b>			<b>English-Sanskrit</b>	<b>Sanskrit-English</b>	<b>English-Bengali</b>	<b>Bengali-English</b>
<b>English</b>	500136	6312	83.6		<b>Inter-Sentential</b>	2356	2366	339	339
<b>Bengali</b>	46933	3907	7.84		<b>Intra-Sentential</b>	2338	851	124	0
<b>Sanskrit</b>	51246	7202	8.56						
<b>Total</b>	598315	17421	100						
<b>Transcription/Translation</b>									
<b>Languages</b>	<b>#types</b>	<b>#tokens</b>	<b>Types per line</b>	<b>Tokens per line</b>	<b>Avg. token length</b>				
<b>English[100%]</b>	40,324	601889	10.58	11.27	4.92				
<b>French (France)</b>	50510	645651	11.38	12.09	5.08				
<b>German[100%]</b>	50748	584575	10.44	10.95	5.57				
<b>Gujarati[100%]</b>	41959	584989	10.37	10.95	4.46				
<b>Hindi[100%]</b>	29744	716800	12.36	13.42	3.74				
<b>Hungarian[100%]</b>	84872	506608	9.13	9.49	5.89				
<b>Indonesian[100%]</b>	39365	653374	11.54	12.23	6.14				
<b>Italian[100%]</b>	52372	512061	9.23	9.59	5.37				
<b>Latvian[100%]</b>	70040	477106	8.69	8.93	5.72				
<b>Lithuanian[100%]</b>	75222	491558	8.92	9.2	6.04				
<b>Nepali[100%]</b>	52630	570268	10.03	10.68	4.88				
<b>Persian (Farsi)[100%]</b>	51722	598096	10.61	11.2	4.1				
<b>Polish[100%]</b>	71662	494263	8.99	9.25	5.86				
<b>Portuguese (Brazil)[100%]</b>	50087	608432	10.8	11.39	5.12				
<b>Russian[100%]</b>	72162	490908	8.96	9.19	5.79				
<b>Slovak[100%]</b>	73789	520465	9.39	9.75	5.37				
<b>Slovenian[100%]</b>	68619	516649	9.35	9.67	5.3				
<b>Spanish[100%]</b>	49806	608868	10.75	11.4	5.07				
<b>Swedish[100%]</b>	48233	581751	10.31	10.89	5				
<b>Tamil[100%]</b>	84183	460678	8.37	8.63	7.65				
<b>Telugu[100%]</b>	72006	464665	8.34	8.7	6.56				

<b>Turkish[100%]</b>	78957	453521	8.27	8.49	6.35				
<b>Bulgarian[100%]</b>	60712	564150	10.1	10.56	5.24				
<b>Croatian[100%]</b>	73075	531326	9.58	9.95	5.28				
<b>Danish[100%]</b>	50170	587253	10.4	11	4.98				
<b>Dutch[100%]</b>	42716	595464	10.52	11.15	5.05				