

English/Serbian Aligned Corpus (SrpEngKor)

This corpus consists of English source texts aligned with their translations into Serbian, and Serbian source texts aligned with their English translations, as well as of several aligned English and Serbian translations of literary texts originally in French. The texts belong to various domains: fiction, general news, scientific journals, web journalism, health, law, education, and movie sub-titles. There are also several Serbian translations of texts from the ‘Acquis communautaire’ corpus and from the ‘Intera’ corpus aligned with their originals. The alignment was performed on the subsentential level. The texts were segmented and aligned automatically and then manually checked. In most cases the alignment is one-to-one. The size of the corpus is 4,420,711 words (2,330,742 in the English part, 2,089,969 in the Serbian part).

Content: 31-01-2013

Size of the corpus

No.	Text	Type	English size	Serbian Size	Source language	
1	Balkan news	Gen	388933	381681	en/sr	
2	Svetislav Basara "Fama o biciklistima"	Lit	68080	52634	sr	
3	Nine stories by Hemingway	Lit	15573	12890	en	
4	Infoteka/Infotheca	Sci	195606	176261	sr/en	
5	Jane Austen "Northanger Abbey"	Lit	77996	75714	en	
6	Jane Austen "Mansfield Park"	Lit	161381	157488	en	
7	Edgar Allan Poe "The Fall of the House of Usher "	Lit	7214	6359	en	
8	Intera	Law/Pol	1039883	901893	sr/en	
9	George Orwel "1984"	Lit	105210	89950	en	
10	Verne "Around the World in 80 days"	Lit	65588	58735	fr	
11	Slate, 04-Sep-2011, Anne Applebaum: The Worst Mistake America Made After 9/11	Gen	1035	1023	en	
12	Slate, 05-Sep-2011, Christopher Hitchens: Simply Evil	Gen	1546	1442	en	
13	Slate, 21-Oct-2011, Christopher Hitchens: The New Libya's First Mistake	Gen	1005	927	en	
14	Slate, 21-Nov-2011, Christopher Hitchens: In God They Trust	Gen	1106	952	en	

15	Slate, 07-Feb-2012, Anne Applebaum: The Games Authoritarians Play	Gen	720	710	en	
16	http://www.signandsight.com, 21-Feb-2012, Thierry Chervel: A kiss for the whole world	Gen	2852	2496	en	
17	Peščanik, 24-Feb-2012, Svetlana Lukić: Fusnote	Gen	1360	1602	sr	
18	Peščanik, 29-Feb-2012, Saša Ilić: Pad Zoranove republike	Gen	974	1209	sr	
19	Slate, 06-Mar-2012, Anne Applebaum: Putin's True Victory	Gen	738	656	en	
20	Peščanik, 18-Mar-2012, Saša Ilić: Paradoks poraza Dobrice Ćosića	Gen	1325	1044	sr	
21	Peščanik, 2-Apr-2012, Nepotpisan tekst: Tadićeve izbore	Gen	610	493	sr	
22	Peščanik, 4-Apr-2012, Miloš Ćirić: Marš na Odu radosti	Gen	854	642	sr	
23	Peščanik, 16-Apr-2012, Saša Ilić: Stubovi unutrašnje politike	Gen	1557	1346	sr	
24	Polityka, Warsaw, 19-Apr-2012, Aleksandra Lipczak: Indignado generation finds happiness abroad	Gen	829	774	en	
25	Spiegel On-Line, 20-Apr-2012, Carsten Volkery: A Vote of No Confidence in Europe	Gen	725	656	en	
26	The New York Times, 26-Apr-2012, Paul Krugman: Death of a Fairy Tale	Gen	791	699	en	
27	Peščanik, 1-May-2012, Srđa Popović: Vozi majstore, svi su ušli	Gen	1462	1161	sr	
28	Peščanik, 3-May-2012, Vesna Pešić: Organizovana pljačka izbora	Gen	1409	1163	sr	
29	Peščanik, 5-May-2012, Wolfgang Klotz: Serbia's Foreign Policy	Gen	4550	3791	en	
30	Peščanik, 9-May-2012, Trajan Zec: Čedomir vs LDP	Gen	857	735	sr	
31	The New York Times, 17-May-2012, Paul Krugman: Apocalypse Fairly Soon	Gen	819	763	en	
32	Peščanik, 21-May-2012, Žarko Korać: Vreme kohabitacije	Gen	1588	1303	sr	
33	Peščanik, 22-May-2012, Srđa Popović: Trenutak istine	Gen	2107	1868	sr	
34	The New York Review of Books, 24-May-2012, Paul Krugman: How to End This Depression	Gen	2236	1942	en	
35	Peščanik, 28-May-2012, Dubravka Stojanović: Premijer	Gen	1131	973	sr	
36	Peščanik, 1-Jun-2012, Svetlana Lukić: Gde je stadoh, ja produžih	Gen	1656	1395	sr	
37	The New York Times, 10-Jun-2012, Paul Krugman: Another Bank Bailout	Gen	782	714	en	
38	Peščanik, 11-Jun-2012, Nadežda Milenković: Kritika eklektičkog uma	Gen	476	444	sr	
39	Peščanik, 11-Jun-2012, Saša Ćirić: Lice Srbije	Gen	1491	1318	sr	
40	Peščanik, 20-Jun-2012, Snježana Milivojević: Prvi i poslednji izveštaj o korupciji i medijima	Gen	2900	2559	sr	
41	The New York Times, 24-Jun-2012, Paul Krugman: The Great Abdication	Gen	788	716	en	
42	The New York Review of Books Blog, 03-Jul-2012, Charles Simic: My Fourth of July	Gen	1037	880	en	
43	New Statesman, 03-Jul-2012, Ryan Gallagher: America, Sweden and Assange	Gen	1058	907	en	

