

## A Supplemental Material

We give here some examples of sentences mined in the BUCC corpora by our approach for language pairs which are not part of the official shared task.

0.1736	FR: A long terme, une société ouverte ne peut survivre que si les personnes qui y vivent croient en elle. DE: Auf lange Sicht kann eine offene Gesellschaft nur überleben, wenn die Menschen die darin leben auch an sie glauben.
0.1726	FR: Ce que toutes ces situations ont en commun est une importante diversité d'intérêts culturels, ethniques, et économiques. DE: Was alle diese Situationen gemeinsam haben, ist eine große Vielfalt kultureller, ethnischer und wirtschaftlicher Interessen.
0.4585	FR: Et dans ces régions, les prix de l'ivoire augmentaient vite, et conduisaient à la formation de groupes professionnels de chasseurs d'éléphants. DE: Auch in diesen Regionen stiegen die Preise für Elfenbein in raschem Tempo und führten zur Bildung von professionellen Elefantenjägergruppen.
0.4804	FR: En 1845, le chimiste britannique Charles Mansfield, travaillant sous la direction d'August Wilhelm von Hofmann, l'isole dans le goudron de houille. DE: 1845 isolierte der englische Chemiker Charles Mansfield während seiner Arbeit unter Leitung von August Wilhelm von Hofmann Benzol aus Steinkohlenteer.
0.3766	FR: Sept pays concourent pour le grand prix: l'Allemagne, la Belgique, la France, l'Italie, le Luxembourg, les Pays-Bas et la Suisse. DE: Beteiligt waren zwölf europäische Staaten: Baden, Belgien, Dänemark, Frankreich, Hessen, Italien, die Niederlande, Portugal, Preußen, die Schweiz, Spanien und Württemberg.

Table 1: BUCC corpus: examples of sentence pairs extracted in the French and German monolingual corpora. All but the last sentence pair are perfect mutual translations. This last example shows that our approach may wrongly align sentences which are mainly an enumeration of named entities. Both sentences enumerate several country names, but they don't match. Many of these erroneous alignments could be possibly excluded by some post-processing, e.g. comparing the number of named entities in each sentence.

0.2852	ZH: 政府应该改善我们的信息基础设施, 以便金融合同能够更好地描述经济风险的后果。 (The government should improve our information infrastructure so that financial contracts can better describe the consequences of economic risks.) RU: И правительствам следует улучшить нашу информационную инфраструктуру, с тем чтобы финансовые контракты мо (And governments should improve our information infrastructure so that financial contracts can better reflect the effects of economic risk.)
0.3799	ZH: 仅仅一年之后, 瑞士政府邀请所有欧洲国家以及美国、巴西和墨西哥等国的政府参加正式的外交会议。 (Just a year later, the Swiss government invited all European countries as well as the governments of the United States, Brazil and Mexico to participate in official diplomatic conferences.) RU: В следующем году швейцарское правительство пригласило правительства всех европейских стран, а также США, Бразилии и Мексики на официальную дипломатическую конференцию. (The following year, the Swiss government invited governments all European countries, as well as the US, Brazil and Mexico to the official diplomatic conference.)

Table 2: BUCC corpus: examples of sentences extracted in the Russian and Chinese monolingual corpora. These alignments seem to be perfect, according to a translation into English by an MT system.