

# Sentence Alignment

Goal: Given a text in two languages, align sentences.

Assume: Sentences hardly ever reordered.

- Classical algorithm: Gale and Church (1993).
  - Based on similar character **length** of aligned sentences, no words examined.
  - Dynamic-programming search for the best alignment.
  - Allows 0 to 2 sentences in a group: 0-1, 1-0, 1-1, 2-1, 1-2, 2-2.
- Several algorithms for English-Czech evaluated by Rosen (2005).
  - Nearly perfect alignment possible by a combination of aligners.
- The “standard tool”: Hunalign (Varga et al., 2005).
- Another option: Gargantua (Braune and Fraser, 2010).

Illustration: MT Talk #7 ([https://youtu.be/\\_4lnyoC3mtQ](https://youtu.be/_4lnyoC3mtQ))