

The role of context in compound detection during reading of unspaced text

Friday, June 23, 2017 2:50 PM (20 minutes)

In Chinese, efficient segmentation strategies are crucial for fluent reading, as word boundaries are not signaled by obvious segmentation cues like spaces. In alphabetic languages like English, word boundaries are clearly indicated by interword spaces and presenting these languages in the same way as Chinese, that is without interword spacing, slows down reading to a great extent (Rayner et al, 1998). In the current eye movement experiment Finnish sentences were presented without spaces and we investigated what segmentation cues are used to determine word boundaries under these circumstances. It was found that one of the cues is the co-occurrence of letter bigrams, with low co-occurrence of letters (e.g., the rt-sequence in a phrase like 'itisclearthat') leading to swifter segmentation. Moreover, target word compounds (e.g., vuorileijona 'mountain lion') were processed more slowly when the first constituent (here mountain) fitted in with the preceding context as in 'hesawthemountainlionfromadistance'. Processing proceeded more smoothly when this was not the case (as in 'heheardthemountainlionfromadistance'). We argue that this finding is caused by more effortful segmentation in the first compared to the latter case.

Similar decisions have to be made in Chinese constantly, as in Chinese subsequent characters often but not always need to be unified to form compound words. To investigate this issue further, we conducted an eye movement experiment in Chinese in which subsequent character combinations were less or more likely to form compounds as a function of co-occurrence characteristics and preceding context. The data of this experiment are currently being analyzed.

Primary author: BERTRAM, Raymond (University of Turku)

Co-authors: HE, Liyuan (Tianjin Normal University); LIVERSEDGE, Simon (University of Southampton)

Presenter: BERTRAM, Raymond (University of Turku)

Session Classification: Contributed papers 5

Track Classification: Freely Contributed Paper