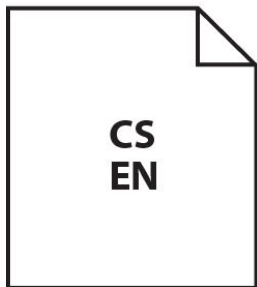


# Getting Balanced Vocabulary

Parent corpus



Child  
corpus



Balanced  
vocabulary



the\_  
že\_  
ying\_  
staying\_  
pra  
pracovat\_  
...