The following analysis is based on a sample corpora(english as source, chinese as target and align). I set both vocabulary sizes as 10000.

1. the first figure is about whether the unk in target sentence has aligned words in source sentence.

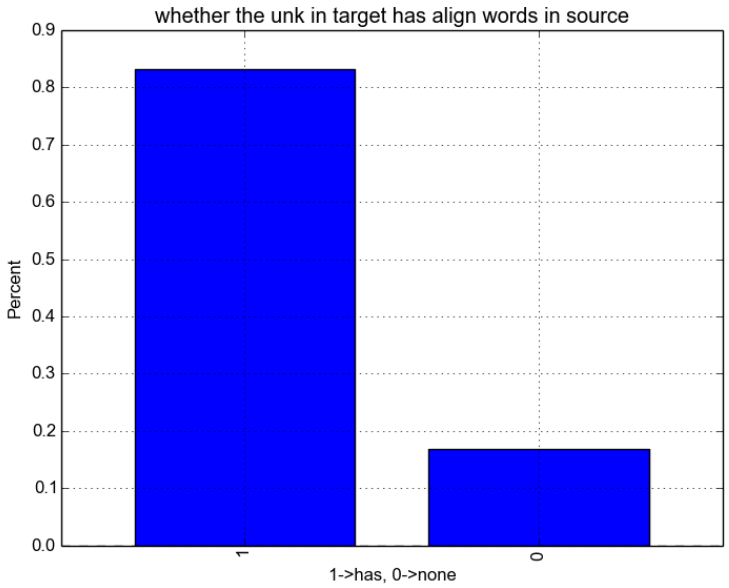

Figure 1: for over 80% cases, the unk has aligned words in source sentence

2. if the unk has aligned words in source sentence, how many words(unk or not) does the unk align to

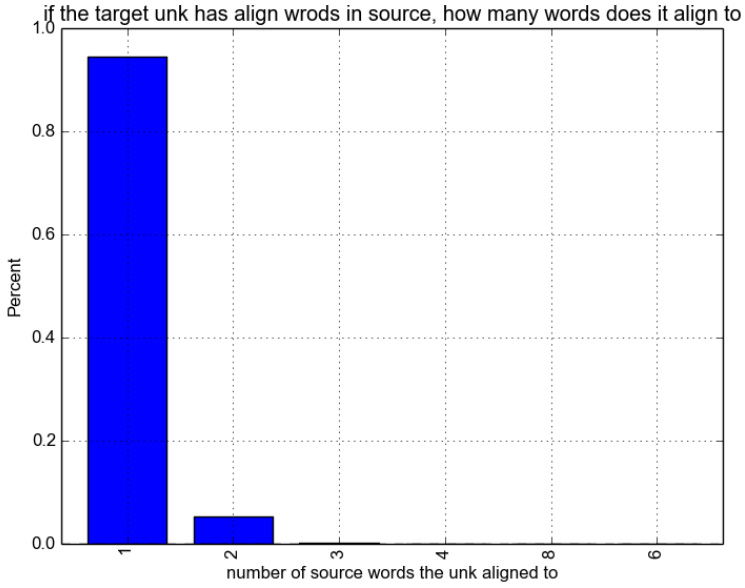

Figure 2: for over 90%cases, a unk aligns to only one word in source sentence

3. for those source words that aligned to the target unk, how many unk(source unk) do they contain

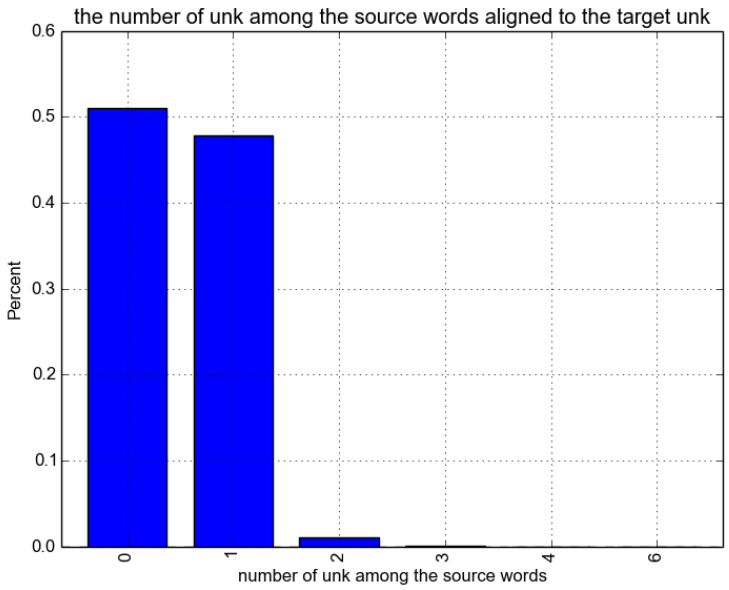

Figure 3: there is no unk in the aligned source words for almost 50% cases and only one unk for another 50% cases

4. this figure measures the same thing by portion

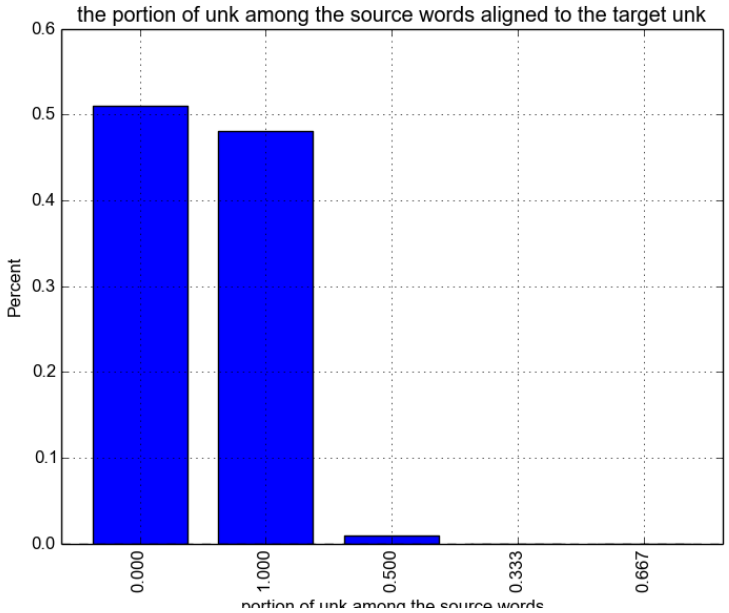

Figure 4: the result is the same as in the above figure

5. for those source words that aligned to the target unk, the relative distances between them and the target unk

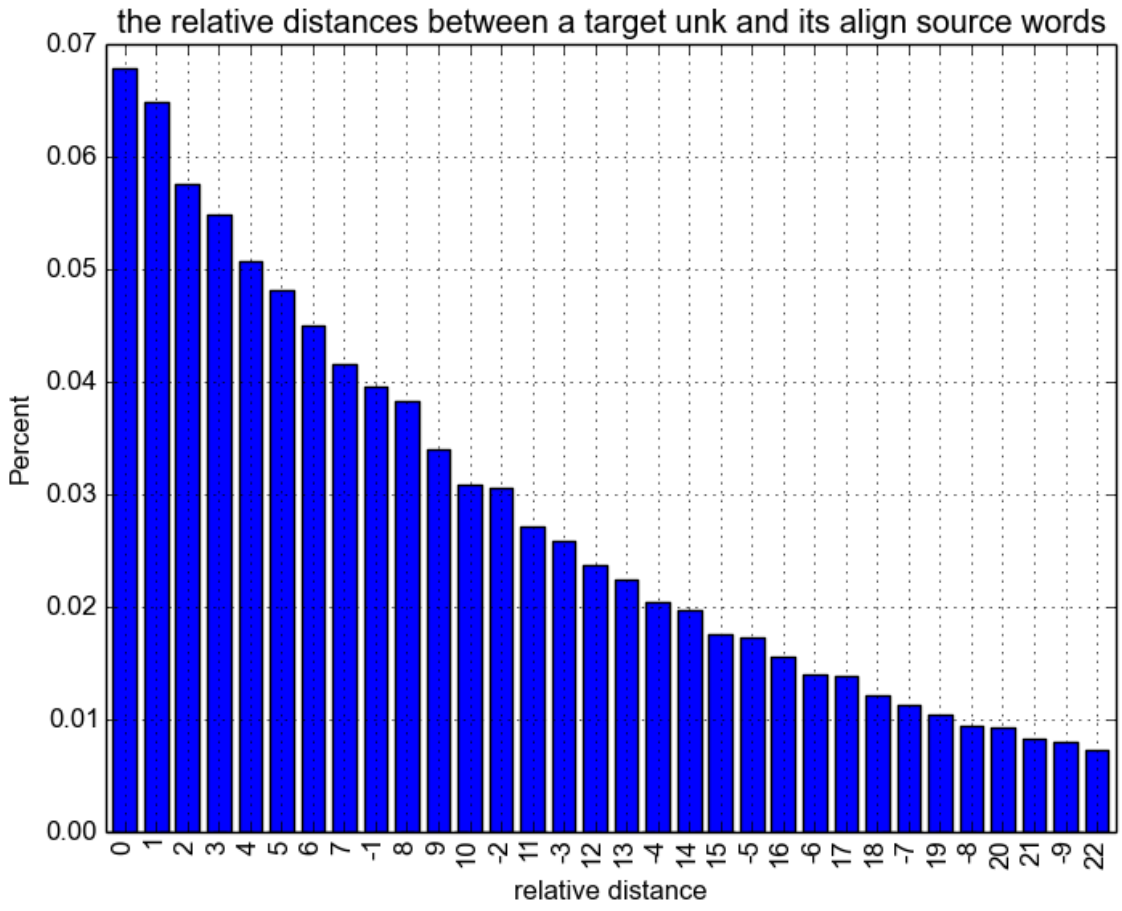

Figure 5: it looks like english word in a alignment pair is more likely to be on the right of chinese word

6. for those source words that aligned to the target unk, the relative distances pattern between them and the target unk

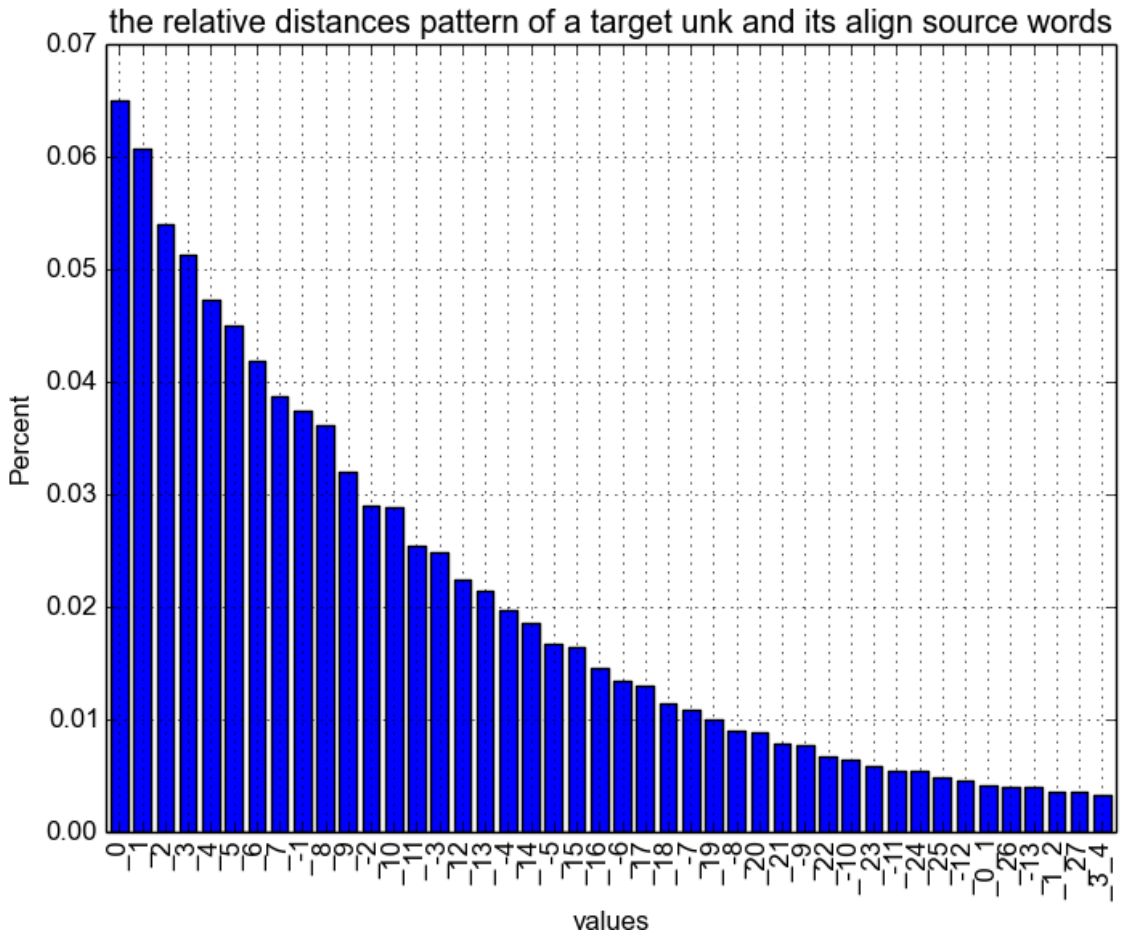

Figure 6: the '_' in the labels of x-axis emans a aligned word and the number following '_' means the relative distance. Most of the time, there is only one aligned word for a target unk and the distances distribution between them looks close to the above figure.