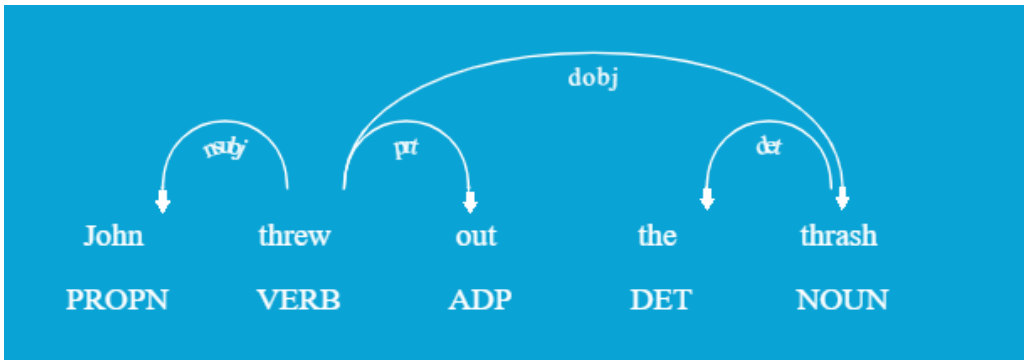

Fig. 1 sentence with a total dependency length of 6, calculated between each head and it's independent. Example, the direct object of the verb "threw" is the noun "thrash" with a DL of 3 and a dependency length of 1 between "threw" and "out".

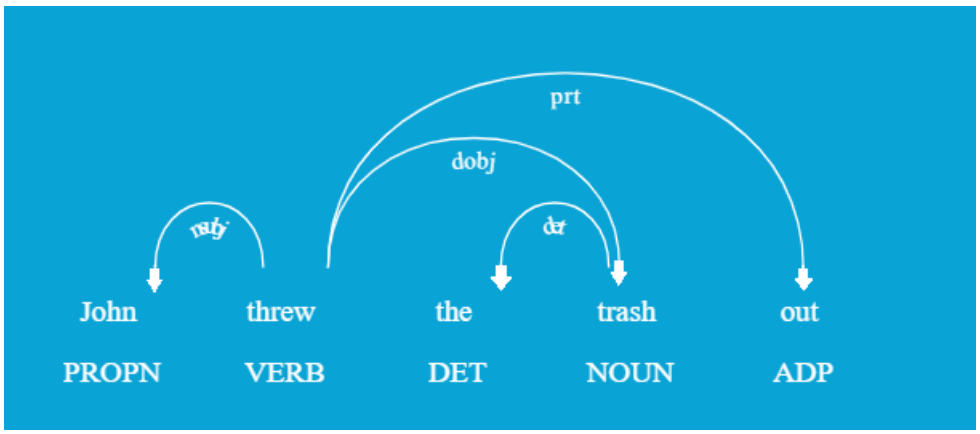

Fig. 2, sentence total has a dependency length of 7. While the distance between "threw" and "trash" decreased from 3 in Fig. 1 to 2 in Fig. 2, the distance between "threw" and "out" increased from 1 to 3.

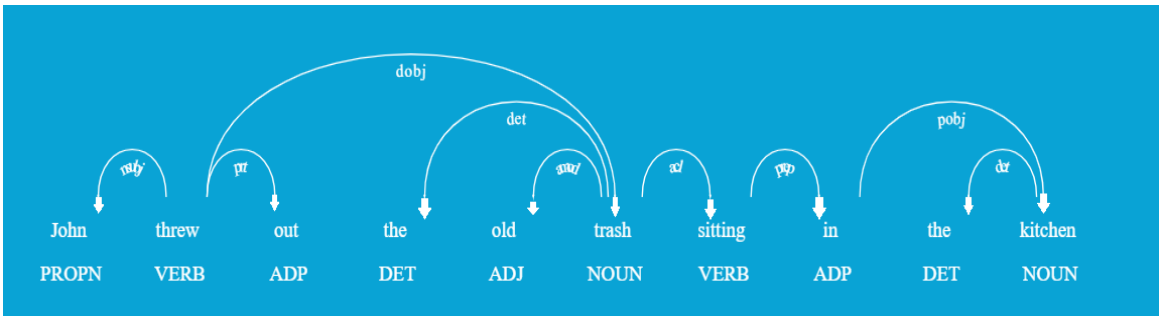

Fig. 3, sentence with a total dependency length of 14. Same sentence as the sentence in Fig. 1, but with added component such as the adjective "old" and the adverbial phrase "sitting in the kitchen".

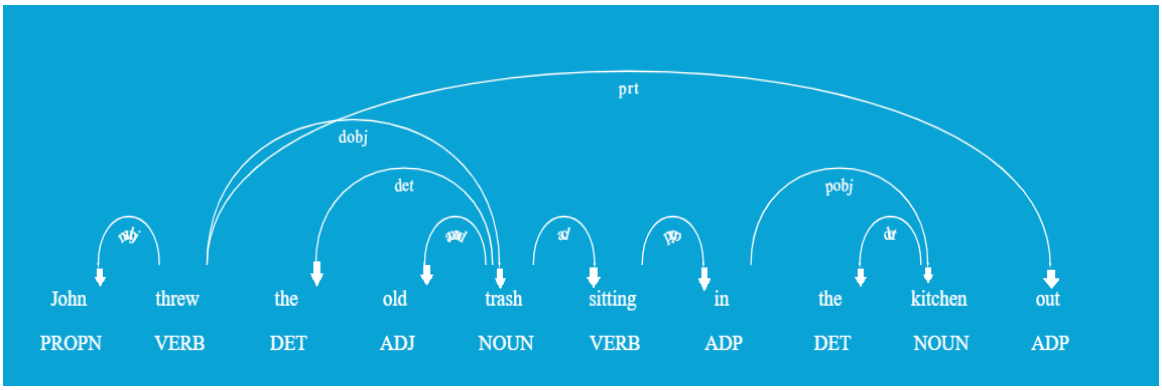

Fig. 4, sentence with total dependency length of 20. Although it has the same semantics and number of words as the sentence in Fig. 3, the total dependency length was inflated because the length between "threw" and "out" is now 8 instead of 1.

Assuming that the sentence in Fig. 1 is denoted as sentence A. We define sentence A as a sequence of n words, where each word is a set of other words with the uniqueness property that if $w \in x \in A$ and $w \in y \in A$ then $x = y$.

We can denote the k^{th} word in A as A_k. Sentence A total dependency can therefore be denoted

$$\sum_{k=1}^{n} \sum_{w \in A_k} \emptyset(w)$$

Where $\emptyset : A \rightarrow N$ satisfies $\emptyset(w) = d$, where $w = A_j$, the unique superset of w is A_k, and $|j - k| = d$. In short, sentence A total dependency can be denoted

$$\sum_{k=1}^{n} \sum_{w = A_j \in A_k} |j - k|$$