

2a) Whole word level is fast in computation and learns relationships, grammar and alignments that exist at word level. Subword level is useful when vocabulary size is small and also in hybrid with whole word level to give better translation.

2b) We may not require higher dimensions to represent characters.

2c) Neural networks could learn from resource rich languages and apply these learnings to resource poor languages.

2d) i) may be small vocabulary issue.

ii) unable to differentiate between passive and active voices.

iii) unable to understand various parts of speech and their relations.

2f) iii) there can be multiple good translations which may receive poor blue scores.

iv) advantages - standard without variation from human to human, objective criteria.

Disadvantages - several good translations could not be present in references, translation quality may be better captured by humans than objective scores.