Notes

The NER software used can be found here. The tagset includes persons (PERS), nationalities/religious or political groups (NORP), organisations (ORG), geopolitical entities (GPE) and locations (LOC).

Observations: Google Translation (Part 2)

Mistakes:

1. In one case, one whole clause was not translated. The sentence (1) was of the form "Q enn' P" where *enn'* is the quotative particle (equivalent to *ani* in Telugu), and only P was translated.

2. Two particularly long statements (6, 9, 13) were fragmented:

    (6) was converted to one grammatical sentence (without the complete meaning of the original) and two incomplete ones.

    (9) was almost correctly translated, but split into two sentences (the first of which is ungrammatical).

    (13) was converted to one complete and one incomplete sentence.

3. There were occasionally issues in translating names of persons, especially those from outside Europe or with initials (1, 2, 7, 13, 15).

4. Numbers written using Malayalam numerals were translated improperly or not at all (11, 13, 15, 19).

5. The subject of the translation was in the dative case of the original and the meaning was inverted (12).

6. The verb in a verbless sentence (15), of the same type as (2) [see point 1 below], was not inserted.

Correct translations:

1. A verb was correctly inserted in a verbless sentence. The sentence (2) was of the form "Victory for X, Y, Z" with no verb; the output was "X, Y, Z win".

2. Another case was of an adjective phrase being converted to an adverb phrase (4). "X, who is in a hosptal, is …" was converted to "X is … in a hospital".

3. A sentence fragment was fully and correctly translated. The fragment (8) was an adverb clause captioning a picture.

Observations: NER (Part 3)

1. A place was mistakenly tagged as an organisation (4, 18). The reason is not clear.

2. The exact distinction between the GPE and LOC tags is not fully clear. Sentences (5).

3. A person was mistakenly tagged as organisation (6).

4. "LNG" (liquefied natural gas) was mistakenly tagged as an organisation (9). It could be because it was being used to refer to a type of bus. Further, one bus route was tagged as a person and the other as an organisation in the same sentence.

5. "Mobile Vaccination Unit" was tagged as a person (16). The reason is unclear.
6. The word "Executive", actually part of "Executive Chairman" was tagged as part of the organisation name it immediately followed (19).
7. A person was mistakenly tagged as a geopolitical entity (20).

In general, the main problem appears to be with the identification of organisations. In many cases (1, 3, 4 and 6 above), the entity actually tagged as an organisation was in fact something else.
However, it must be noted that no non-named entities were mistaken as named entities.