Semantic Representation of Relativization: A Contrastive Study from the Natural Language Generation Perspective

Contrastive linguistics attempt to find out similarities and differences in both related and non-related languages. Semantic representation as an interlingua has been seen as an important resource for multilingual generation. We propose to carry out a contrastive analysis of Bangla, Telugu, Hindi, and English in order to develop a suitable language independent semantic representation for *relativization* which is linguistically realized very differently in the aforementioned languages. Interestingly, the languages under consideration belong to different language families. Bangla and Hindi are Indo-Aryan languages, English is Germanic language while Telugu belongs to the Dravidian family of languages.

Consider the English sentence- *That girl who is standing, is my friend* which has the possible translations into other three mentioned languages as presented in the following table-.

Bangla:

(i) ōi mēẏēṭā yē   dāṁṛiẏē āchē   sē āmāra bandhu.

   That girl-cl REL stand-pft   is   she my   friend

(ii) dāṁṛiẏē   thākā   mēẏēṭā āmāra bandhu.

    Stand-pft stay-participial girl-cl   my   friend

Hindi:

vo laṛaki jo   khaṛi   hai (vo) merī dosta hai

That girl-cl REL stand-pft is   she my friend is

Telugu

(i) ā   nilabaḍ-ina   am'māyi nā snēhiturālu.

 that stand.pres.prog-REL girl.nom   my friend.acc

(ii) ā   nilabaḍi unna   am'māyi nā snēhiturālu.

    that stand.pres.prog stay.REL girl.nom my friend.acc

We observe that English relative clauses can be translated both as a relative clause or as a participial clause in Bangla; while Hindi prefers the relative clause construction and Telugu only participial construction. We propose in this paper one semantic representation from which all the above structures can be generated. In this proposal, the relative clause or participial construction both are treated as the modifier of a noun. AMR (Banarescu et. al. 2019) and MRS (Copestake et. al 2005) both represent the relative clause information as modifier to the head noun. AMR uses inverse relation between the head noun and the verb that heads the

relative clause to express modification. MRS, on the other hand, equates the label value of the head noun and the head verb of the relative clause. Inspired from Paninian grammar, we name such modification *samaanaadhikaraNa* 'same locus'. The novelty of our representation lies in identifying aspectual information that the modifier clause/participial construction denotes. We postulate two relations representing perfective and imperfective modification of the head noun. The advantage of encoding this information is apparent for the language generation task. We have observed that the participial construction is the only mechanism for expressing relativization information in Telugu while the preferred construction in English and Hindi is relative clause. Bangla applies both mechanisms. This paper examines different kind of clausal modification of noun to determine whether one single semantic representation with aspectual nuances would be optimal to express all kinds of clausal modification of nouns.

This study is part of a larger ongoing project on multi- layered universal semantic representation for generation purposes in the light of Panini and Indian grammatical tradition.

References

Copestake, A. et. al 2005. Minimal Recursion Semantics An Introduction
https://www.cl.cam.ac.uk/~aac10/papers/mrs.pdf

Banarescu et. al. 2019. Abstract Meaning Representation (AMR) 1.2.6 Specification
https://github.com/amrisi/amr-guidelines/blob/master/amr.md [last accessed on 03-01-2023]

Srivastav, V.1991. *The Syntax and Semantics of Correlatives.* Natural Language and Linguistic Theory, 9:4 (1991: Nov.) p.637-687.

Subbarao, K.V., 2012. South Asian Languages: A Syntactic Typology
https://www.academia.edu/6666513/South_Asian_Languages_A_Syntactic_Typology   [last accessed on 10-01-2022]