PROJECT IDEAS

Proposal 1:

Research yields that chunking PRCs simplifies their processings. The least common role taken on by a PRC in our findings was object and the most common was subject. We would like to look into what effect nesting may have on this distribution. For this we would use a dataset. We are most interested in common non-formal speech and would again use twitter for data-collection. The structures we are focusing on are as follows:

(1) She who her playmate who was bit had accused was punished.
(2) They who the class which was failing had sought out arrived on time.

Reali, F., & Christiansen, M. H. (2007). Processing of relative clauses is made easier by frequency of occurrence. Journal of memory and language, 57(1), 1-23.

Reali, F., & Christiansen, M. H. (2007). Word chunk frequencies affect the processing of pronominal object-relative clauses. Quarterly Journal of Experimental Psychology, 60(2), 161-170.

Proposal 2:

We would investigate the influences which may yield speech-to-text transcription errors and which factors may aggravate or minimize them. A corpus method of analysis would be used to answer this question. We would focus on professional transcripts (medical, legal, etc.). The following examples of transcription will be taken into account:

(1) gonna → transcription: going to
(2) till → transcription: until
(3) heal → transcription: heel

Chiari, I. (2007). Transcribing speech: errors in corpora and experimental settings. Proceedings of Corpus Linguistics 2007, Birmingham, 1, 13.

Zayats, V., Tran, T., Wright, R.A., Mansfield, C., & Ostendorf, M. (2019). Disfluencies and Human Speech Transcription Errors. *INTERSPEECH*.