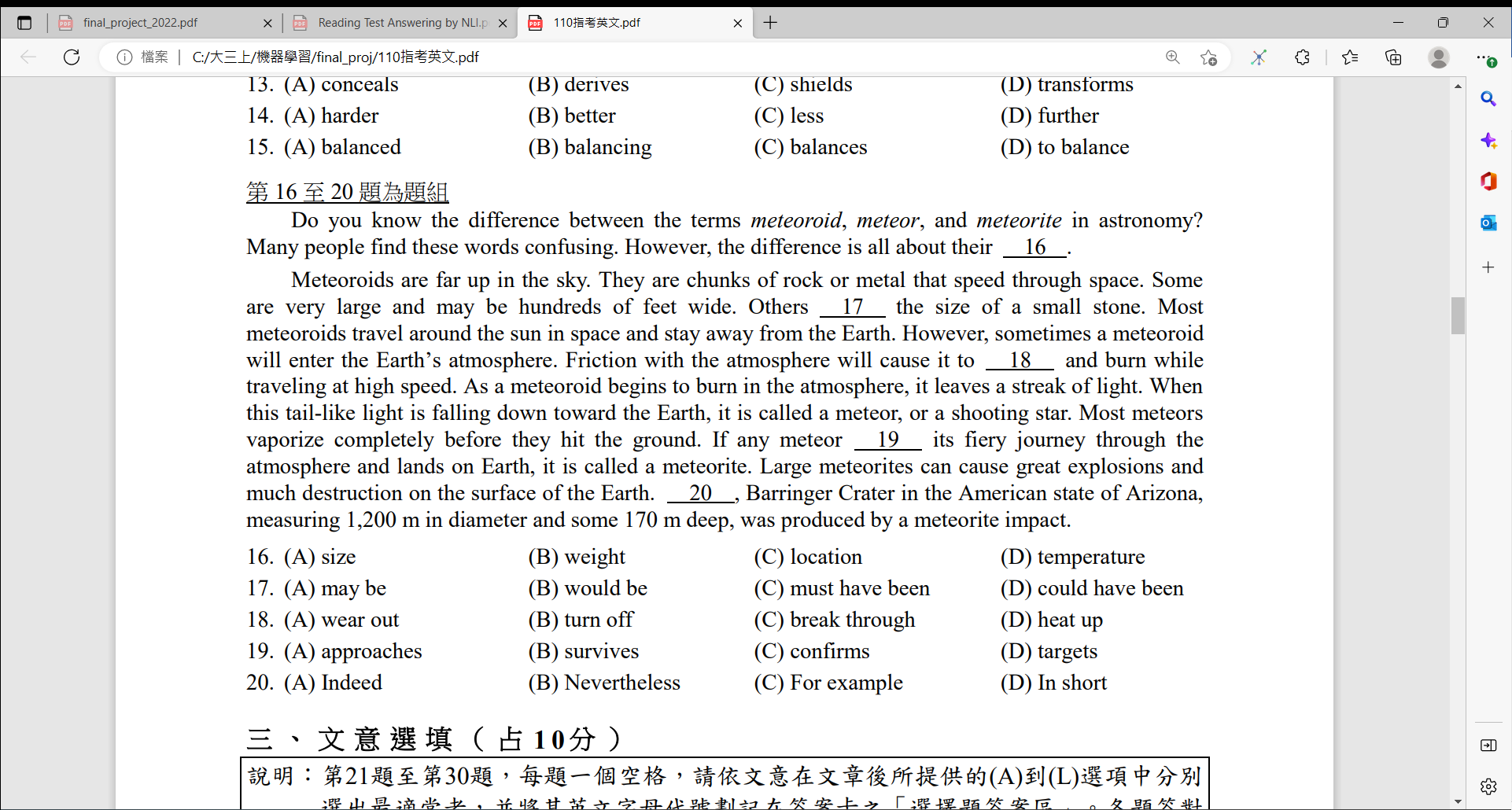
Result:

PREDICT = ['C', 'A', 'D', 'B', 'D', 'D', 'A', 'C', 'B', 'C', 'B', 'C', 'C', 'B', 'B', 'D', 'B', 'D', 'B', 'C']

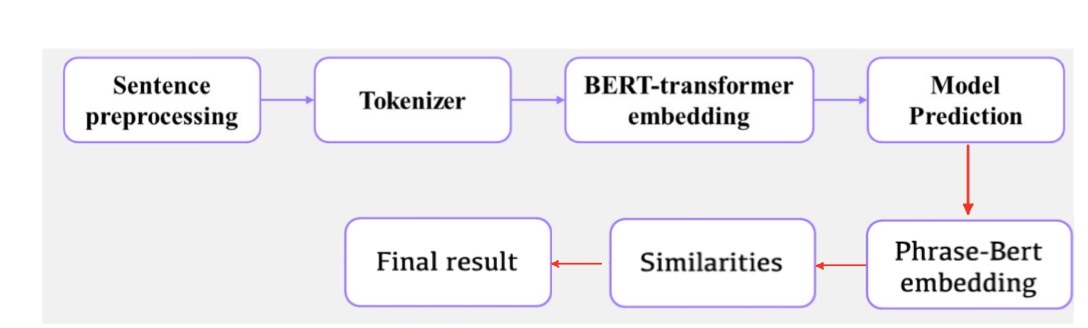
ANSWER = ['C', 'A', 'D', 'B', 'D', 'B', 'A', 'C', 'B', 'C', 'B', 'C', 'C', 'B', 'A', 'C', 'A', 'D', 'B', 'C']

Discussion:

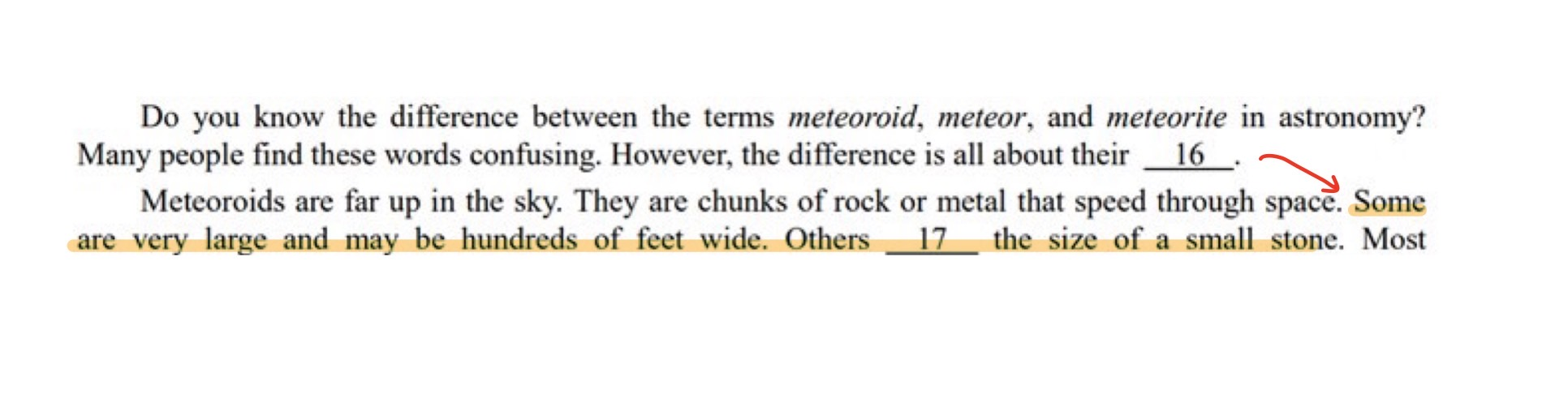
In section two, there’s a chance we get a wrong answer directly from the model if there is a multiple-word option, such as phrases.

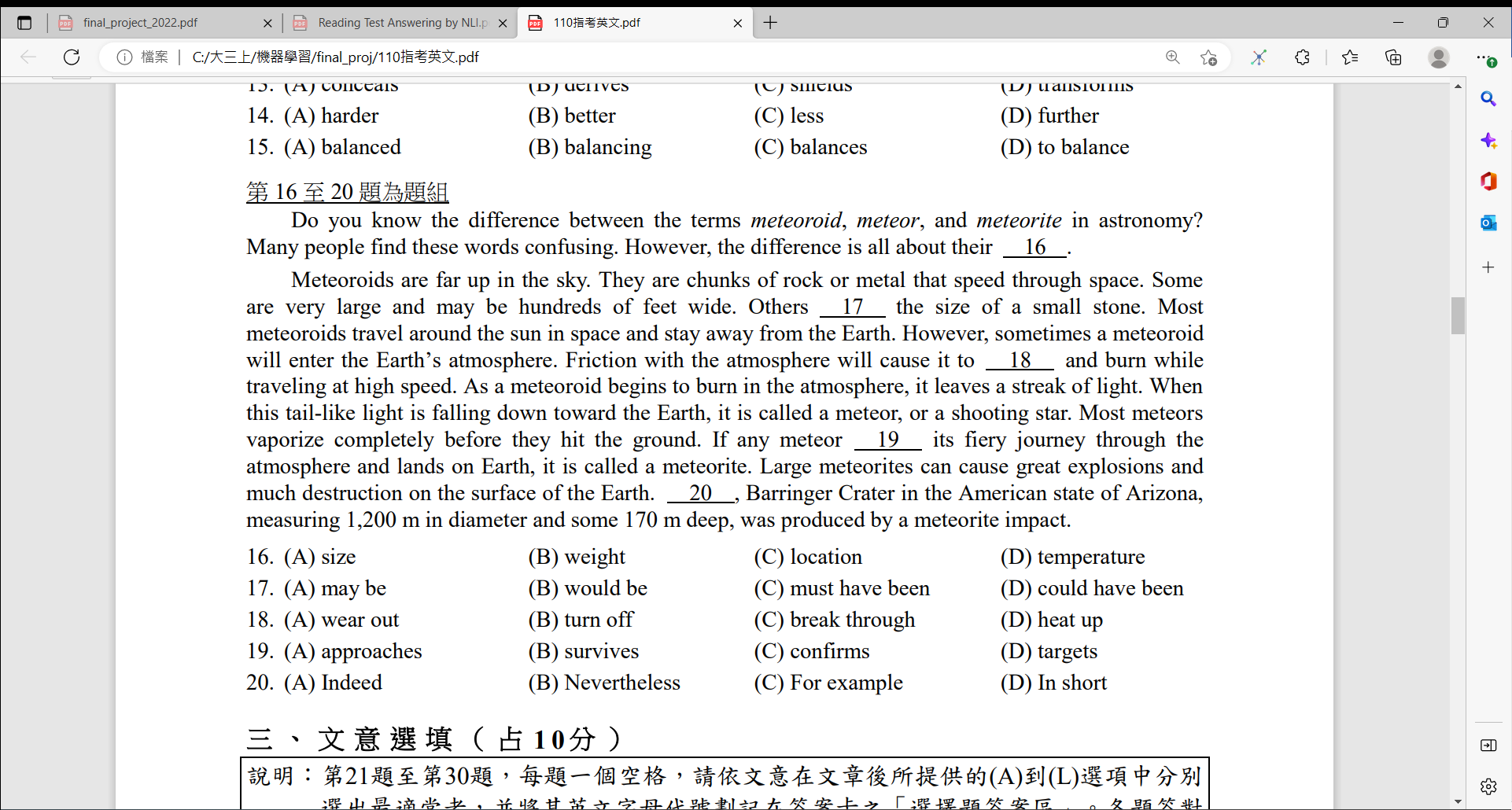


The original model is not able to fit a two-word token into the MASK (the multiple-word tokens will all get the same token ID, meaning it couldn’t handle phrases). In order to handle this problem, we applied another NLP model Phrase-Bert. This model is trained for producing powerful phrase embedding, which can be suitably useful for our problem. After we generate the best word vector for MASK from the original model, for each option, we simply calculate its similarity with the MASK based on the embedding results of Phrase-Bert and choose the one with the highest similarity as our final answer.

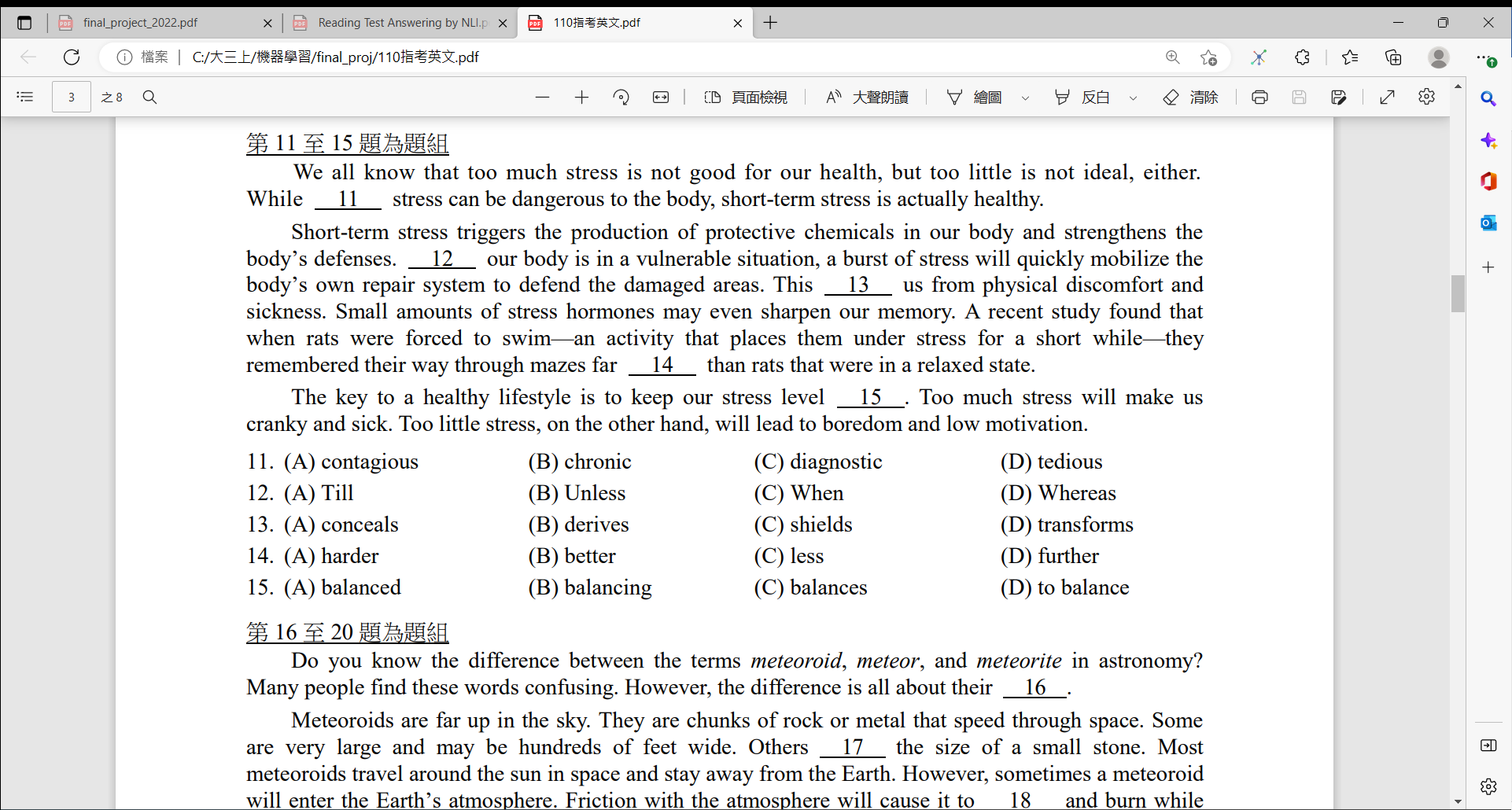


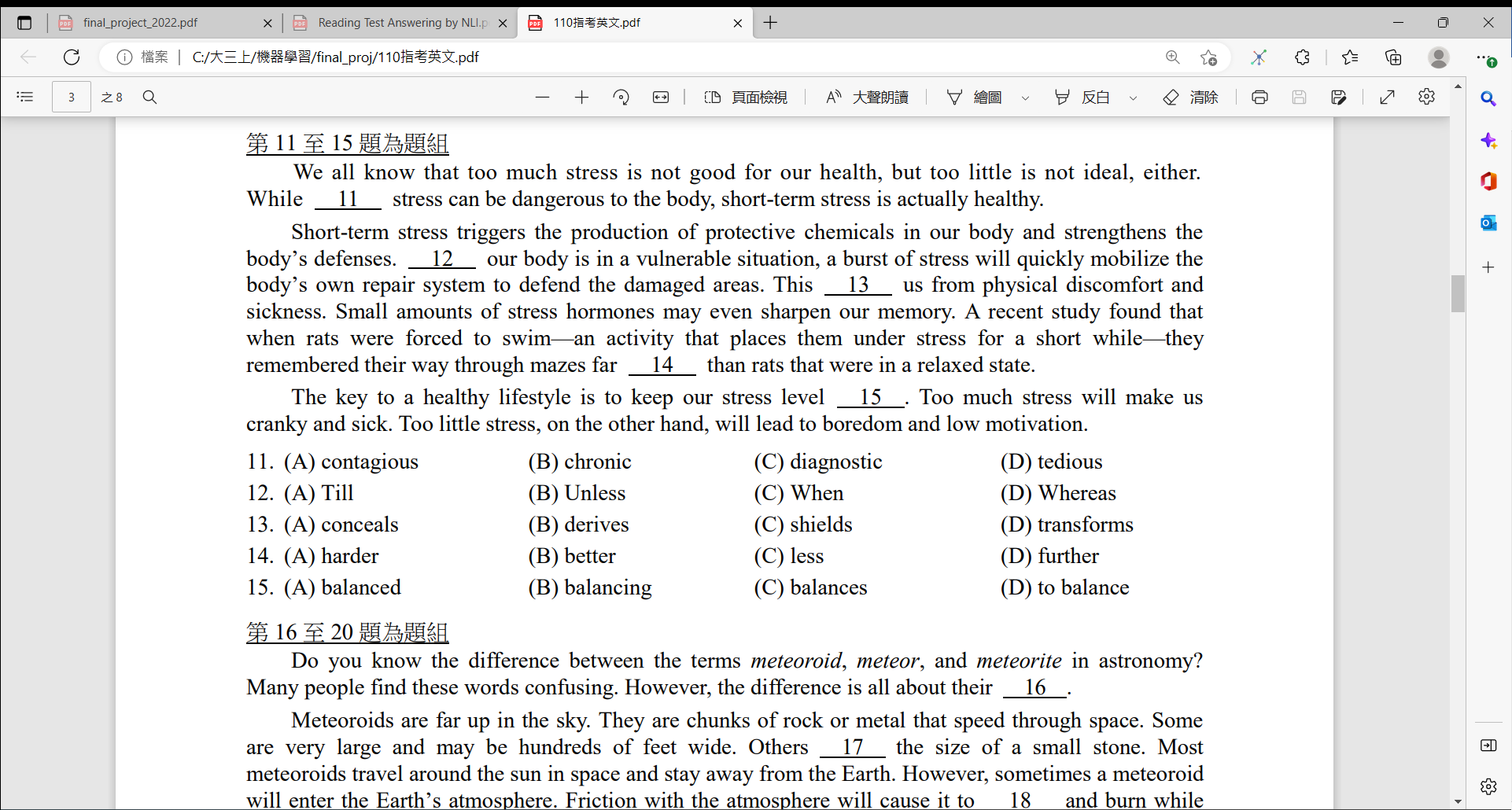
Although there are mostly similar types of questions in one section. Some of the questions cannot be answered correctly using the previous model.





For question 16, the answer actually depends on the sentences located in the next paragraph. As we know the calculated vector of the MASK should be mostly dependent on other tokens nearby, the model couldn’t generate the best answer based on the given input since the distance between the MASK position and the clue is too far away.





This type of question requires grammar checking. Since different tenses of a verb can share the same or similar vector values, the result obtained from the model could not be concrete. Therefore, another model for grammar checking should be applied in this case. Classifying these special types of questions could take much effort. Since we haven’t thought of an ideal way to detect this question type plus we mainly focus on the Reading Section for this project, we’ve decided to neglect the minor special cases in these sections.