1.

In SLP3, topics related to the algorithm involving in natural language processing are covered more compared to SLP2. To be more specific, there are more topics related to machine learning, or say, deep learning in SLP3. For example, logistic regression, neural networks are both covered. SLP3 also explain how they work, by incorporating the probability theories involved in the book, such as naïve Bayes.

SLP3 also cover the idea of word embeddings, which is not in SLP2. This idea of converting various words to different vectors proves to be very efficient in the field of natural language processing lately, as well as the idea of vector space, and how vectors that are close to each other in a vector space are considered similar to each other. These are all topics that are barely covered in SLP2.

Lastly, SLP3 put more focus on parsing than SLP2 does. This is probably because as the natural language processing techniques develop, more complicated context could possibly be processed by the computer. Therefore, parsing techniques targeted for more complicated contexts become. Important, which explains why they are covered in SLP3.

In contrast, in SLP3, there is less focus on semantics. This is probably because there is not a lot to be explored in semantics involving NLP.

2.

After computing bigrams for both tables of content, I used the Grokking model to calculate the similarity scores of the bigrams computed for each table of content, resulting in two csv files (SEE ATTACHED FILES), in which the first column represents the word pairs evaluated and the second column represents the corresponding cosine similarity scores. Based on the two csv files, I found out that for SLP3, there are more ‘N/A’ for the scores, this potentially indicates that more vocabularies in SLP3 are no in the raw data that we used to train the Grokking model, which indirectly shows that the subjects of SLP3 may be more advanced compared to SLP2, such that the words that are used to name the subjects in SLP3 are not commonly, or at least not frequently used by people. One more thing I noticed is that the similarity scores are all pretty low. This is also probably due to the raw data that we used to train the model. If we were to use a book that is about NLP to train the model, then the similarity scores may go up by a large amount.