Tf-idf and PPMI are the models of language implicitly used in the lecture slides Term-Document Matrix & Similarity Measures.

1.
For the Tf-idf model, the entities needed are included as: term frequency (tf) and inverse document frequency (idf). This means that we would need to have multiple documents and certain terms that occasionally show up in some of or all of the documents used for the model to be applicable.

For the PPMI model, the entity needed is actually a term-context matrix, which is used to calculate the PPMI for different words given target words, such that we would be able to know if two events, or to be more specific, if two words co-occur more than if they were independent. For example, the term-context matrix mentioned above tells the marginal probability of the occurrence of each term, occurrence of each context, and the joint probability of the occurrence of each term-context combination.

2.
The linguistic entity I heard of but is not presented in the model is speech recognition and classification. The two models focus mostly on plain text, rather than other forms of linguistic expressions.

3.
Broadly, for the Tf-idf model, we need nltk, numpy for data preprocessing, word2vec from genism for word embedding, and lastly, cosine_similarity from sklearn to calculate the cosine similarity scores between different word pairs or term pairs.
For the PPMI model, we need nltk and numpy for data preprocessing, especially matrix operations, and we would also need math package very likely to calculate the PPMI for word pairs or term pairs.

4.
I think these models could be helpful in the field of speech recognition, text classification and so on. However, they may not be considered as effective in subjects that have completely no relationship with speech or text.