Language modeling in the context of Natural Language Processing (NLP) refers to the development of probabilistic models that can predict the likelihood of a sequence of words appearing together in a sentence or document. Essentially, it involves understanding, generating, or completing texts based on the patterns and structures inherent in a language. Language models are fundamental to various NLP tasks and applications, including but not limited to:

1. **Text Generation:** Generating coherent and contextually relevant text based on a given prompt or starting sequence. This is used in applications like chatbots, content creation tools, and automated storytelling.

2. **Speech Recognition:** Translating spoken language into text by predicting sequences of words that are likely to occur together in the spoken language.

3. **Machine Translation:** Translating text from one language to another by understanding the probabilistic relationships between words in both the source and target languages.

4. **Autocorrect and Predictive Text:** Suggesting corrections for misspelled words or predicting the next word in a sentence based on the words that precede it.

5. **Sentiment Analysis:** Determining the sentiment expressed in a piece of text, which relies on understanding the combinations of words that express positive, negative, or neutral sentiments.

**Language models can be broadly categorized into two types:**

1. **Statistical Language Models:** These models use statistical methods to predict the probability of a given sequence of words appearing together. They often rely on counting the frequency of sequences and applying techniques like n-grams, where the probability of a word occurring is conditioned on the occurrence of the previous n-1 words.

2. **Neural Language Models:** These models use neural networks to predict the likelihood of word sequences. They are capable of capturing more complex patterns in language use than statistical models, partly because they can consider a larger context and learn representations for words (word embeddings) and sentences. Popular architectures for neural language models include Recurrent Neural Networks (RNNs), Long Short-Term Memory networks (LSTMs), Transformers, and variations thereof.

The advent of large-scale neural language models, such as those based on the Transformer architecture (e.g., GPT series by OpenAI, BERT by Google), has significantly advanced the capabilities of NLP systems, enabling more sophisticated and contextually nuanced text generation, comprehension, and analysis.