# Natural Language Processing: Where it all begins…

Abelether Germa

*Teesside University*

[*abel\_germa@live.co.uk*](mailto:abel_germa@live.co.uk)

Jess Bogue

*Teesside University*

[*n3078300@live.tees.ac.uk*](mailto:n3078300@live.tees.ac.uk)

Adam Postgate

*Teesside University*

[*n2095821@live.tees.ac.uk*](mailto:n2095821@live.tees.ac.uk)

# *Abstract*

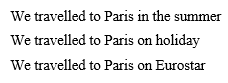
This

**Keywords:** Lexicon, Grammar, Semantics, Grammar rules

**1. Introduction**

This paper examines a series of sentences provided by Simon. Lynch, these sentences can be understood and deconstructed by building a lexicon and a grammar, although these sentences seem simple enough for humans, they are complex when it comes to machines. The paper will explore the decisions and approaches needed to create the system and will explain how it fits into a wider context of language processing. LKit will be the software used to develop this machine learning process.

The program will take in one of the three sentences from the third series of sentences, provided on the assignment brief and return a breakdown of the sentence types. A single sentence may derive from more than one type of sentence and the program will reflect this in its return value. The sentences to be considered are as follows;



**Figure 1: series of sentences**

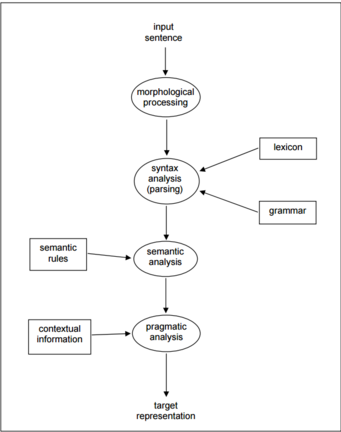
The sentences shown in figure 1 above, will be added to the lexicon along with others not provided to give a wider spectrum of variety, for sentence choices. For example; it could say;

We travelled to London on the Eurostar in Winter.

Each of these words would have to be included in the lexicon, in order for the sentence above to be applicable and parsable.

The programming involved in Natural Language Processing demonstrates how sentences are formed and how to react to that sentence in an appropriate manner. The machine therefore needs a solid understanding of common words and phrases, which is where the lexicon and the grammar are required. These provide word details and sentence rules that allow the language processor to determine a structure of a sentence.

The diagram shown in figure 2 below, illustrates how a sentence will be deconstructed through a series of stages.



**Figure 2: Data Flow Diagram**