Abstract

Multi-word unit is a group of two or more orthographic words which carries one single meaning. Such linguistics feature is essential in conversational process, as it allows language user to convey meanings accurately and authentically. Lexical chunk is a type of multi-word unit that is commonly seen during language usage, yet often neglected in the linguistic research field.

This paper targets on Mandarin lexical chunks used in Singapore context. By using a corpus-driven approach, it is possible in identifying more forms of lexical chunks, other than those which are recognised earlier. The “Spoken Corpus of Singapore Mandarin” is made up of a total of 197,432 words. 114 lexical chunks are identified from this corpus.

First, a comparison will be made between lexical chunks and other types of multi-word units, such as idioms, clichés and discourse markers. It is important to differentiate them as they each have different characteristics in grammar, structures or phonology aspects of these multi-word units. Characteristics of lexical chunks are further summarized, such as fixed structure, fixed meaning and variable phonological syllables.
Next, an empirical investigation of the structural "patterns" represented by the lexical chunks will be studied. The lexical chunks are categorized according to their structure characteristics. Of which, 47.4% are lexical chunks with fixed slots, 24.5% are with variable slots, and 28.5% are repetitive lexical chunks.

Lastly, a frequency study will carry on the categorization of these lexical chunks, to identify the high, mid and low frequency lexical chunks. High frequency lexical chunks appear more than 40 times in the corpus, which is 12.3% of total lexical chunks. Mid frequency lexical chunks appear between 10 to 39 times, making up 25.4% of lexical chunks. 62.3% of the lexical chunks are low frequency, which appear less than 10 times within the corpus.

This study aims to bring more attention to lexical chunks in Singapore Mandarin speakers. By categorizing the lexical chunks, it is clearer of what lexical chunks can be focused and taught earlier. Mastering of lexical chunks will give Singaporeans better confidence in using Mandarin in daily conversation, as well as reducing the possibility of communication breakdowns due to influx spoken Mandarin.