

*Oliver Mason (Birmingham, United Kingdom)*

## **Reconciling Phraseology and Grammar**

### **Abstract**

As shown in previous work (e.g. Mason 2008), multi-word units which have been automatically extracted from a corpus using a straightforward algorithm can be used to describe various features of sentences, such as their 'naturalness', or idiomaticity. For this, pre-calculated MWUs are matched with the actual words used in a sentence, and the degree of coverage is measured. Gaps or points of discontinuity in the match describe points of maximum variability or higher level unit boundaries.

However, this approach is at odds with traditional ways of describing the grammatical structure of sentences, as the MWUs do rarely correspond to established linguistic units, and the segmentation achieved through the MWU-mapping approach bears little resemblance to traditional phrase types.

There are two corpus-based approaches to grammar which come closest to the phraseological description: Sinclair and Mauranen's Linear Unit Grammar (2006) and Hunston and Francis' Pattern Grammar (2000). In this paper I will investigate how LUG, PG, and the description of grammar using MWUs can be used to bridge the gap between a purely bottom-up approach ignorant of all theoretical concepts and the more linguistically informed formalisms.

The connection with LUG is fairly straightforward, as it uses 'chunks' (which are defined intuitively only) as the basic unit of analysis, which could conceivably be seen as similar in form and function to automatically retrieved MWUs. Pattern grammar seems more closely related to valency grammar, and its use of traditional parts of speech (which are absent from both LUG and the MWU approach) and phrasal categories make it difficult to find any commonalities to the other two.

### **References**

- Hunston, S. / Francis, G. (2000), *Pattern Grammar*, Amsterdam: Benjamins.
- Mason, O (2008), 'Stringing together a sentence: linearity and the lexis-syntax interface', in Gerbig and Mason (eds.) "Language, People, Numbers", Amsterdam: Rodopi, p. 231-248.
- Sinclair, J. / Mauranen, A. (2006), *Linear Unit Grammar*, Amsterdam: Benjamins.