Introduction

Aleksandar Trklja
Universität Innsbruck
Łukasz Grabowski
University of Opole

The notion of formulaicity has received increasing attention in disciplines and areas as diverse as linguistics, literary studies, art theory and art history. In recent years, linguistic studies of formulaicity have been flourishing (e.g. Wray 2002; 2008; 2009; Schmitt & Carter 2004; Wood 2010b,a; 2015; Kecskes 2016; Myles & Cordier 2017; Piirainen et al. 2020), and the very notion of formulaicity has been approached from various methodological and theoretical perspectives and with various purposes in mind, be it descriptive, exploratory or applied.

The object of investigation in linguistic studies are multiword expressions (MWE) but individual approaches and models differ in how MWE are defined and identified in language. For these reasons, it would be wrong to claim that all linguistic studies of formulaicity constitute a uniform field of research. There is no such a thing as 'formulaicity linguistics'. Linguistic formulaicity has become a superordinate term for the view that a large proportion of natural language consists of repetitive lexical units. This makes MWE somehow special with respect to alternative linguistic units of analysis that have theoretical foundations in formal syntactic, semantic or lexical structures. Such structures can be and are often included in the study of linguistic formulaicity but they do not provide the minimum necessary conditions against which MWE are set as linguistic units. In fact, there are authors who proposed new approaches or models that deny the existence of such structures. The minimum assumption shared by all studies of linguistic formulaicity is that a MWE is considered a unit because it is a linguistic expression that has been repeatedly reused. The very fact that a linguistic expression is re-used across different situations and by different language users constitutes a good ground to treat it as a unit of analysis. It is therefore no wonder that the main focus in the study of linguistic formulaicity is on the investigation

Aleksandar Trklja & Łukasz Grabowski. 2021. Introduction. In Aleksandar Trklja & Łukasz Grabowski (eds.), *Formulaic language: Theories and methods*, iii–viii. Berlin: Language Science Press. DOI: 10.5281/zenodo.4727659

of the effect repetition has on various language issues such as idiomaticity, language acquisition, formation of social discourses, translation-related issues etc. As one can see, the novelty of these studies does not lie in the introduction of new issues they address but rather in a new treatment of established issues.

Linguists of various schools have studied linguistic formulaicity using different approaches and research perspectives, and with different purposes in mind. In an attempt to provide a useful generalization and conceptual clarification, Gałkowski (2006: 163–164) argues that it is possible to distinguish between three major approaches to linguistic formulaicity, namely a linguistic, psycholinguistic and sociolinguistic one. The focus of the purely linguistic approach is on the investigation of formulaicity in terms of lexical and grammatical categories identified primarily using formal grammatical or functional lexical criteria. The psycholinguistic approach is primarily concerned with the study of how linguistic data is stored, processed as well as retrieved from the mental lexicon. Finally, the sociolinguistic approach explores situational and cultural aspects tied to the use of formulaic language (Gałkowski, ibid.). In reality, most studies combine these approaches as illustrated in Schmitt & Carter (2004); Wood (2010b,a; 2015); Wray (2002) or Underwood et al. (2004); Piirainen et al. (2020), among others. Also, there has been a plethora of research conducted in recent years by specialists in corpus and computational linguistics, who study formulaic language with primarily applied purposes in mind, such as development of natural language processing tools (NLP) or machine translation tools, fine-tuning textual classification methods etc. (cf. Forsyth & Grabowski 2015; Pęzik 2018). Given such a proliferation of research perspectives, it is no surprise that formulaic language has been defined, labelled and operationalized in many different ways (cf. Wray & Perkins 2000; Wray 2002; 2009), and each approach brings new insights into this interesting, yet at the same time, not fully and comprehensively explored phenomenon. This observation provided the main rationale for the present volume. We invited specialists that cover the whole spectrum of relevant issues and thus showcase their state-of-the-art research.

Thus, we present a selection of studies into formulaic language arranged into complementary sections. The first section with three chapters presents new theoretical and methodological insights as well as their practical application in the development of custom-designed software tools for identification and exploration of formulaic language in texts. The second section with two chapters presents examples of innovative research into formulaic language in language learning contexts. Finally, the third section with three chapters showcases research on formulaic language conducted primarily from corpus linguistic, discourse studies and translation studies perspectives.

The first chapter by Joan Bybee and Ricardo Napoleão de Souza focuses on the relation between frequency effects typical of linguistic prefabrication and phonetic effects. By exploring a sample of adjective-noun sequences extracted from a conversational corpus, Bybee and de Souza show that certain phonetic effects, such as vowel duration, correspond to conventionalized structures found in prefabricated expressions. They also argue that phonetic effects are promising in view of future studies focusing on the notion of conventionality of prefabricated expressions. The authors demonstrate that prefabricated expressions constitute the conventional means of referring to these entities or concepts of some cultural importance despite being semantically compositional. In addition, they show that prefabs form clusters of semantically related word sequences and that they can contribute to creativity in language use.

Richard Forsyth looks into formulaic language from a corpus-driven perspective and proposes a set of computational procedures to quantify the degree of formulaic language in individual texts and language corpora. Forsyth implements his approach into a custom-designed freely available software written in Python and shows - using an additional criterion of coverage - how n-grams of various lengths emerge from the data and facilitate determination of the degree to which texts are permeated with recurrent sequences of words.

The chapter by Piotr Pęzik focuses on the identification of prefabricated expressions in dependency-annotated corpora. More precisely, he investigates restrictions on the valency of binary collocations and their tendency to be regularly subsumed by larger collocational chains. Specific examples from Polish and English are followed by a presentation of Treelets software, where the Author's approach has been implemented, which illustrates in practical terms how recurrent multi-word items may be systematically explored using dependency-based methods.

The second section opens with a contribution by Stephen Cutler, who deals with an important problem of how new formulaic language is acquired and stored by L2 learners of English. In these studies, two different learning paths are contrasted: fusion (operationalized through a focus on the sequence's elements and structure) versus holistic acquisition (operationalized through a focus on the spoken sound form of the sequence as a whole). Cutler argues that the findings provide further support to the claim that regular retrieval and simple corrective feedback help consolidate recall of the sequences learnt by L2 learners.

Ying Wang undertakes a successful attempt at a comparison of ideational functions of formulaic language in native student and expert academic writing. The chapter presents unique features of formulaic sequences identified in each text variety and shows that native student writing is more characteristic of everyday

and highly idiomatic formulaic sequences, among others, while expert academic writing abounds in formulaic language associated with research and scientific argumentation. In conclusion, Ying Wang presents an informative discussion on how the research findings translate into formal instruction.

The last section of the volume starts with a chapter by Andreas Buerki, who shows how changes in social discourse are reflected in phraseology. Taking the 2016 referendum on the United Kingdom's membership in the European Union as reflected in a large, tailor-made corpus of media texts, Buerki identifies various discursive strategies reflected in recurrent phraseologies and compares their use across time and specific topics. The results obtained in the study demonstrate how phraseological units reflect specific ideological positions. In addition, the present data indicates that formulaicity plays an important role in the Brexit discourse because it is more formulaic than the comparable discourse. Finally, the chapter casts new methodological insights into how phraseology, and formulaicity in general, can be used in discourse analytical research.

Łukasz Grabowski and Nicholas Groom undertake an attempt at employing the concept of grammar patterns in descriptive research on formulaic language in English-to-Polish translation. Their aim is to verify whether the Polish equivalents are realized with the same level of regularity. The detailed findings show that grammar patterns can be useful as a unit of analysis and a starting point for exploration of formulaicity in translation, and that they may cast more light onto some more general differences between semantics and pragmatics in source texts and translations.

Finally, the last chapter in the volume, by Mikhail Mikhailov, takes under scrutiny the concept of syntactic idioms and explores through a corpus linguistic analysis the structure, meaning and use of the Russian construction N-*s*-N and its English and Finnish matches. These counterparts are identified in parallel corpora. Mikhailov argues that the Construction Grammar approach used in his study helps make syntactic idioms more explicit for descriptive purposes, also when explored with the use of parallel and comparable corpora.

We believe that such a selection of original studies collected in this book will provide more insights into a fascinating phenomenon of formulaicity in language explored from both a systemic and textual angle. We sincerely hope that the volume will therefore come in useful for anyone interested in formulaic language, from both a theoretical and practical perspective.

Obviously enough, this volume would not have been possible without many people involved in its preparation, compilation and production. First of all, we would like to thank the Authors of the chapters for accepting our invitation and for further smooth collaboration through the entire production process, from

the initial submission, review stage, revision stage to the very preparation of final versions of the chapters. We would also like to cordially thank our reviewers (Mikhail Kopotev, Stephen Jeaco, Francis Bond, Janusz Malak, Tadeusz Piotrowski, Stanisław Goźdź-Roszkowski, Łucja Biel, Laura Vilkaite, Jiří Milicka, Larisa Leisiö, Rita Jukneviciene, Magda Stroińska, Martin Hilpert, Cristiano Brocas, Alexander Rosen), who gave of their time for careful inspection and evaluation of all submitted chapters. Last but least, special thanks are extended to Editors of the series "Phraseology and Multiword Expressions" at Language Science Press, in particular to Michael Rosner, Manfred Sailer and Agata Savary, for giving us a green light to prepare and publish the volume, as well as to Sebastian Nordhoff and Felix Kopecky for their invaluable help in typesetting and technical matters. In particular, our sincere thanks are extended to Michael Rosner, who successfully and flexibly co-ordinated the entire volume preparation process despite difficult pandemic-related circumstances.

Aleksandar Trklja, Łukasz Grabowski (Volume editors)

References

Forsyth, Richard & Łukasz Grabowski. 2015. Is there a formula for formulaic language? *Poznań Studies in Contemporary Linguistics* 51(4). 511–549. DOI: 10.1515/psicl-2015-0019.

Gałkowski, Błażej. 2006. Kompetencja formuliczna a problem kultury i tożsamości w nauczaniu języków obcych. *Kwartalnik Pedagogiczny* 4. 163–180.

Kecskes, I. 2016. Deliberate creativity and formulaic language use. In K. Allan, A. Capone & I. Kecskes (eds.), *Pragmemes and theories of language use, perspectives in pragmatics, philosophy & psychology 9, pp*, 3–20. Cham, Switzerland: Springer International Publishing.

Myles, Florence & Caroline Cordier. 2017. Formulaic sequences (FS) cannot be an umbrella term in SLA. *Studies in Second Language Acquisition* 39. 3–28.

Piirainen, Elisabeth, Natalia Filatkina, Sören Stumpf & Christian Pfeiffer (eds.). 2020. *Formulaic language and new data: Theoretical and methodological implications*. De Gruyter. DOI: 10.1515/9783110669824. https://www.degruyter.com/document/doi/10.1515/9783110669824/html (1 March, 2021).

Schmitt, Norbert & Ronald Carter. 2004. Formulaic sequences in action: An introduction. In Norbert Schmitt (ed.), *Formulaic sequences: Acquisition, processing and use*, 1–22. Amsterdam, Netherlands: John Benjamins.

Underwood, Geoffrey, Norbert Schmitt & Adam Galpin. 2004. The eyes have it: An eye-movement study into the processing of formulaic sequences. In Norbert Schmitt (ed.), *Language learning & language teaching*, vol. 9, 153–172. Amsterdam: John Benjamins. DOI: 10.1075/lllt.9.09und. https://benjamins.com/catalog/lllt.9.09und (1 March, 2021).

Wood, David (ed.). 2010a. *Formulaic language and second language speech fluency: Background, evidence and classroom applications*. London: Continuum.

Wood, David (ed.). 2010b. *Perspectives on formulaic language: Acquisition and communication*. London: Continuum.

Wood, David. 2015. *Fundamentals of formulaic language*. London: Bloomsbury.

Wray, Alison. 2002. *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Wray, Alison. 2008. *Formulaic language: Pushing the boundaries*. Oxford: Oxford University Press.

Wray, Alison. 2009. Identifying formulaic language: Persistent challenges and new opportunities. In Roberta Corrigan, Edith Moravcsik, Hamid Ouali & Kathleen Wheatley (eds.), *Formulaic language. Vol. 1. Distribution and historical change*, 27–51. Amsterdam: John Benjamins.

Wray, Alison & Michael R. Perkins. 2000. The functions of formulaic language: An integrated model. *Language & Communication* 20(1). 1–28.