#### Expanding the Lexicon

## The Dynamics of Wordplay

## Edited by Esme Winter-Froemel

Editorial Board Salvatore Attardo, Dirk Delabastita, Dirk Geeraerts, Raymond W. Gibbs, Alain Rabatel, Monika Schmitz-Emans and Deirdre Wilson

## Volume 5

# **Expanding the Lexicon**

Linguistic Innovation, Morphological Productivity, and Ludicity

Edited by
Sabine Arndt-Lappe, Angelika Braun, Claudine Moulin
and Esme Winter-Froemel

**DE GRUYTER** 

This book series was established in connection with the project "The Dynamics of Wordplay" funded by the German Research Foundation (DFG). The present volume is published in cooperation with the "Forum Sprache und Kommunikation Trier". Additional funding was provided by the University of Trier.

Cette collection a été créée dans le cadre du projet « La dynamique du jeu de mots », financé par la Deutsche Forschungsgemeinschaft (DFG). Le présent volume est publié également en collaboration avec le « Forum Sprache und Kommunikation Trier », l'Université de Trèves fournissant un financement complémentaire.



ISBN 978-3-11-050084-4 e-ISBN [PDF] 978-3-11-050193-3 e-ISBN [EPUB] 978-3-11-049816-5



This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 License. For details go to http://creativecommons.org/licenses/by-nc-nd/4.0/.

#### Library of Congress Cataloging-in-Publication Data

A CIP catalog record for this book has been applied for at the Library of Congress.

#### Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

© 2018 Sabine Arndt-Lappe, Angelika Braun, Claudine Moulin and Esme Winter-Froemel, published by Walter de Gruyter GmbH, Berlin/Boston
Printing: CPI books GmbH, Leck

Printed on acid-free paper
Printed in Germany

www.degruyter.com

#### **Contents**

Sabine Arndt-Lappe, Angelika Braun, Claudine Moulin, and Esme Winter-Froemel

Expanding the Lexicon: At the crossroads of innovation, productivity, and ludicity —— 1

#### I Linguistic Innovation

Natalia Filatkina

Expanding the lexicon through formulaic patterns — 15

Anette Kremer and Stefanie Stricker

Complex words in the early medieval *Leges Barbarorum* and their contribution to expanding the Old High German lexicon —— 43

Sören Stumpf

Free usage of German unique components — 67

#### **II Morphological Productivity**

Ingo Plag and Sonia Ben Hedia

The phonetics of newly derived words: Testing the effect of morphological segmentability on affix duration —— 93

Marcel Schlechtweg

How stress reflects meaning — 117

Sabine Arndt-Lappe

Expanding the lexicon by truncation: Variability, recoverability, and productivity —— 141

#### **III Ludicity**

Angelika Braun

Approaching wordplay from the angle of phonology and phonetics – examples from German —— 173

Georgette Dal and Fiammetta Namer

Playful nonce-formations in French: Creativity and productivity — 203

Esme Winter-Froemel

Ludicity in lexical innovation (I) - French - 229

Claudine Moulin

Ludicity in lexical innovation (II) - German - 261

#### **Appendix**

List of Abstracts / Liste des résumés — 289

List of Contributors and Editors — 297

Index — 301

Sabine Arndt-Lappe, Angelika Braun, Claudine Moulin, and Esme Winter-Froemel

# Expanding the Lexicon: At the crossroads of innovation, productivity, and ludicity

#### 1 The dynamic lexicon

Traditionally, the creation of new lexical units and patterns – understood in a wide sense as not being necessarily limited to the word level – has been studied in different research frameworks. Whereas approaches focusing on morphological productivity are directed at system-internal ('grammatical') morphological processes, other approaches have aimed at identifying general types of lexical innovation and describing them in the larger context of lexical change, thus integrating system-external factors related to the historical background of the innovations and their diffusion. In this way, lexical change provides insights into general motives of language change and basic mechanisms of language processing.

The aim of this volume is to discuss fundamental aspects of dynamic processes in the lexicon, including recent and ongoing changes as well as historical processes of change, and to bring new evidence to bear on the traditional dividing line between approaches oriented towards system-internal and system-external aspects.

Current research in language change is marked by a renewed interest in the lexicon, as documented by recent international conferences and publications on structural, typological and cognitive approaches to the lexicon and on regularities of lexical change in the larger context of language change (see, among many others, Blank 1997; Ágel et al. 2002; Brinton and Traugott 2005; Haspelmath and Tadmor 2009; Libben et al. 2012; Zeschel 2012; Ostermann 2015). At the same time, within theoretical linguistics, recent years have seen an increase in more and more psycholinguistically informed work on morphological complexity and productivity, which explicitly relates issues of productivity and modularity in the lexicon to what we know about lexical processing (e.g. Hay 2003; Baayen et al. 2011; Pirelli et al., in press).

The strong interest in this topic was also documented by the high number of submissions we received for the call for papers for our international workshop Expanding the lexicon / Extensions du lexique / Erweiterungen des Lexikons – Linguistic Innovation, Morphological Productivity, and the Role of Discourse-Related

Factors / Innovation linguistique, productivité morphologique et le rôle de facteurs liés au discours / Sprachliche Innovation, morphologische Produktivität und die Rolle diskursbezogener Faktoren held at Trier University (17–18 November 2016). The workshop brought together participants with different theoretical backgrounds and permitted multilingual discussions and exchange on a wide variety of topics ranging from aspects of the lexicon in medieval times to current innovations in German, English and Romance.

The contributions in this volume go back to papers presented at the workshop as well as to papers presented at the newly created Forum Sprache und Kommunikation Trier (www.fsk.uni-trier.de), which aims to foster inter- and transdisciplinary linguistic exchange on a broad range of linguistic phenomena, taking into account the cultural, social and historical contexts in which they are embedded. At the workshop and in the discussions, three main aspects emerged as being of key interest: 1) lexical innovation and conventionalisation, 2) productivity in its interplay with speaker creativity, and 3) the role of ludicity in lexical innovation. These aspects are addressed from different perspectives by various papers in the volume, as will be shown below. It should be stressed that many of the papers touch upon several of the aspects mentioned, thus demonstrating how closely they are interwoven. The following discussion of the three aspects and the papers grouped in each of the main parts of this volume should therefore be interpreted as showing only some of the many links and common lines of investigation. The reader is invited to cross-read the volume and to discover further convergencies, complementary discussions and perspectives for further research.

#### 2 Innovation and conventionalisation

Studying processes of lexical expansion, the notion of lexical innovation and the diachronic evolution of lexical innovations becoming conventionalised and possibly reused in new ways, represent first topics to be dealt with. These issues are addressed from a theoretical perspective in Filatkina's contribution, which is complemented by Kremer and Stricker's investigation of lexical innovation in Old High German and Stumpf's analysis of innovative free usage of unique components in contemporary German. Moreover, the contributions which will be discussed in sections 3 and 4 below also touch upon synchronic and diachronic aspects of specific subtypes of lexical innovations and their subsequent diachronic evolution.

Natalia Filatkina's contribution, *Expanding the lexicon through formulaic pat*terns: the emergence of formulaicity in language history and modern language use, approaches the topic of innovation from the perspective of formulaic language. As word-formations, formulaic patterns are considered an important means of lexicon expansion and innovation. Filatkina uncovers substantial differences and characteristics in the way formulaic patterns contribute to lexicon expansion. The differences are particularly clear if studied from a (diachronic) perspective of the emergence of formulaic patterns and against the background of theories of language change. The argument is made that the usual "driving forces" of language change such as regularity / irregularity, codification / normatisation, cultural and contextual / discourse traditions and frequency do not apply to formulaic patterns in the same way as they do, for example, to sound change, grammatical or even lexical change. The emergence of formulaic patterns can best be understood as a process of integration of sometimes controversial aspects, among which frequency and regularity seem to be important accompanying factors but not always driving forces. Irregular, idiosyncratic paths based on conflicts and violation of norms shape the development of formulaicity as well if they are sufficiently supported by the speakers' / hearers' communicative needs and / or embedded into discourse and cultural traditions.

A special dimension of the investigation of lexical expansion and innovation is tackled in the paper by Anette Kremer and Stefanie Stricker (Selected Complex Words in the Early Medieval Leges Barbarorum and their Contribution to Expanding the Old High German Lexicon), namely the challenges encountered by the exploration of the topic in historical stages of languages for which our textual records provide only a very limited inventory of texts and a very small literary vocabulary. This is the case with Old High German (AD 700-1050) where the exploration of the lexicon is especially complicated due to the fact that extensive monolingual sources are not available on a large scale over the relevant time axis. A larger quantity of complementary Old High German material can be found in vernacular glosses in Latin manuscripts and in the sources explored in the paper for this volume, namely vernacular lexical items present in Latin law codes of the Germanic peoples written in the Early Middle Ages, the so-called Leges Barbarorum.

In their paper, the authors analyse a selection of complex lexical items (compounds, derivatives) taken from the Upper German law codes (Lex Baiuuariorum, Lex Alamannorum, Leges Langobardorum), as these form a relatively homogeneous tradition. The investigation is carried out with the database of the LegIT project and analyses the formation and use of relevant lexical items in the selected corpus, depicting pathways of expansion of these items in the lexicon of Old High German. Furthermore, the paper focuses on the dynamics of word for-

mation in Old High German, with special attention to complex words not documented outside of the Leges tradition. In this context, specific relations between their first and second elements can be traced and related to the specific text genre where they occur. The analysis of derivation cases draws special attention to lexical items resulting from morphological word formation processes that can be considered typical for the law texts, but are no longer productive, and for which we have hardly any evidence in other Old High German sources. Overall the results of the study show the manifold potential of investigation on the lexical level offered by the *Leges* sources for the medieval vernaculars. For further research, the analysis of these sources not only opens a specific reservoir of lexical domains not recorded elsewhere, but will also enable crosslinked analysis with findings in the textual and glossographic domain in order to trace general pathways of lexical development through time.

Sören Stumpf's paper, Free usage of German unique components: Corpus linguistics, psycholinguistics and lexicographical approaches, investigates how unique components in phrasemes can be (re-)used outside their original phraseological context and thus contribute to linguistic innovation and expansion of the lexicon. Normally, such unique components can only occur within set phrasemes (e.g. German ins Fettnäpfchen treten; an example from English would be happy as a sandboy), but as the author shows, they can be reactivated in language use and once usualised, eventually find their way into dictionaries. Exploring this type of lexical innovation through unique components has not yet been approached in a comprehensive way, and the author focuses in his study on findings from corpus studies on the German language and particularly the underlying debonding processes (Norde 2009). Furthermore, he addresses psycholinguistic issues exploring how phrasemes with unique components are processed in the mental lexicon, how their debonding can be grasped and how the motivation of the unique components plays a central role in this process. The author's findings point to the importance of further diachronic investigation of unique components as a source for lexical innovation and open methodological paths for crosslinguistic research. Furthermore, the topic investigated shows close links to aspects of productivity and creativity as well as ludicity in the expansion process, domains that are the subject of the following sections of the volume.

<sup>1</sup> For more examples see Dobrovol'skij (1988) or the "List of English Bound Words": https://www.english-linguistics.de/codii/codiibw/en/list-complete.xhtml (accessed 13 September 2017).

#### 3 Productivity

The discussion in section 2 has already indicated that one key means of lexical expansion which languages have at their disposal are productive word-formation processes. Such processes are traditionally defined as regular morphological mechanisms, and determinants of as well as constraints on their productivity have usually been described in terms of the components of the language system: phonology, morphology, syntax, semantics (and, to some extent, pragmatics). The articles that were discussed in section 2 above already point to a well-known delimitation issue here, as we have seen that word-formation in this sense is only one of several mechanisms of lexical expansion that can be productive (compare e.g. the processes described in Kremer and Stricker's paper with the productivity of unique components studied in Stumpf's article). In the present section, however, we limit the discussion of productivity issues to those arising in the synchronic study of word formation processes in the traditional sense.

With respect to traditional notions of productivity, the articles in this volume provide interesting insights in mainly two ways: One concerns the question of the level of description needed to characterize productive processes. There are two articles in this volume, one by Ingo Plag and Sonia Ben Hedia, and one by Marcel Schlechtweg, which essentially show that, if we look at how novel linguistic expressions are used in actual speech (albeit, in Schlechtweg's case, in an experimental setting), it is necessary to take into account more than the system-internal components that traditional analyses have studied. Plag and Ben Hedia's article, The Phonetics of Newly Derived Words: Some Case Studies, deals with how prefixed words are realised phonetically in a corpus of English natural speech. They find that the pronunciation of prefixed words reflects the segmentability of that word. Segmentability encompasses both measures of semantic transparency as well as frequency based measures of the competitive activation of morphologically complex words and their bases in language processing (cf. Hay 2003). The findings are highly relevant for the study of lexical innovation: A high degree of segmentability is a characteristic property of productive processes. Building on Plag and Ben Hedia's findings, we can thus expect newly derived words to be pronounced differently (i.e. with longer prefix durations) from older, more lexicalised, derived words. It is an open question whether this type of effect can be captured in terms of the level of granularity that can be formulated with the help of phonological feature systems. Also, Plag and Ben Hedia's findings suggest that the study of newly derived words benefits from integrating the perspective of the speaker and the speech event in the research paradigm. Segmentability and

productivity are properties of individual words, as processed by the individual speaker.

Marcel Schlechtweg's contribution, How stress reflects meaning – The interplay of prosodic prominence and semantic (non-)compositionality in non-lexicalized English adjective-noun combinations, is concerned with the function of prosodic prominence in novel English adjective-noun constructions. On the basis of acoustic data elicited in a small-scale experimental study, the paper presents evidence that prominence patterns are influenced by both the semantic compositionality of the construct itself and the immediate sentence context in which the adjective-noun construct occurs. Two types of context are tested in the experiment: In the first type, the construct is followed by a relative clause that not only paraphrases the non-compositional meaning but also uses a metalinguistic description to explicitly mark the paraphrase as a definition (which is called so because...). In the other type of context provided in the experiment, non-compositionality is merely implied. Unlike in constructs with a compositional semantics, where the noun tends to receive most prominence, in non-compositional constructs the adjective tends to be marked as more prominent. However, the difference between compositional and non-compositional items is only robust in sentence contexts in which the meaning relation between the adjective and the noun is not explicitly provided with the help of a paraphrase. Again, this has implications for the study of productive processes of lexical innovation, as it shows that system-external factors like context influence the formal realisation of newly coined morphological constructs.

A second aspect that characterises discussions of productivity in this volume is the question if and how productive morphological processes are to be delimited from other, specifically creative or playful processes. The article Expanding the lexicon by truncation: variability, recoverability, and productivity by Sabine Arndt-Lappe presents an analysis of truncation patterns (mainly patterns of name truncation as in nickname and hypocoristic formation) in three languages (Italian, German, and English), with a focus on two aspects that have traditionally been used as criteria to delimit productive morphology from other processes. One is structural variability: outputs of truncation are shown to provide evidence of the existence of alternative forms, such that different patterns of truncation can be distinguished. Crucially, variability is systematic and determined by both universal and language-specific morphological factors. The other aspect is semantic transparency: it is argued that, even though in truncatory patterns compositionality of meaning does not correspond to compositionality of form, outputs of truncation may still be transparent, in the sense that the regularities that determine the shape of truncatory patterns as well as the way truncatory patterns are used

in context are optimally geared towards ensuring that the base forms are recoverable, despite the loss of segmental material. The case of truncations thus challenges traditional assumptions that take the degree of productivity of a morphological process to be correlated with formal predictability and semantic compositionality. Instead, like other articles in the present volume, the truncatory data seem to point towards an approach to productivity that relates this notion in a more integrative way to mechanisms of language processing and contextual factors.

### 4 Ludicity

The interplay of productivity and the speakers' creativity touched upon in the papers discussed in the preceding section as well as the central role of individual acts of innovation stressed in usage-based approaches to language change (see also Filatkina's contribution discussed in section 2) point to the active role of the speakers in processes of lexical expansion. One type of lexical innovation in which the active role of the speaker is particularly evident are ludic innovations. Although ludicity is obviously an important dimension in lexical expansion, its role has not yet been studied systematically in previous research. This aspect is also linked to the general topic of the book series in which this volume is included and which is dedicated to the dynamics of wordplay, the latter notion being understood in a broad sense, in order, among other things, to precisely include transitions between ludic and "serious" innovation and to explore degrees of ludicity in lexical innovation. In this way, the present volume also presents strong links to the upcoming volume on wordplay and creativity edited by Bettina Full and Michelle Lecolle (in press).

Among the papers of the present volume, the ludic dimension is directly addressed by Braun, Dal and Namer, Winter-Froemel, and Moulin, focusing on different aspects of ludic usage and on different levels of linguistic description.

Angelika Braun's contribution, Approaching wordplay from the angle of phonology and phonetics - examples from German, aims to outline the benefits and insights to be gained from a phonetically informed approach to wordplay studies. She argues that various types of wordplay and potentially ludic processes of lexical expansion can be described in a more fine-grained way from a phonetic / phonological perspective. Distinguishing between wordplay which is based on existing lexical items and wordplay involving the creation of new items (most importantly, blending), she proposes a classification of various subtypes of word-

play depending on which part of the syllable is involved and which phonetic processes can be observed. In this way, a fine-grained classification of various subtypes of wordplay and ludic processes of lexical expansion is obtained. This classification is tested by analysing more than 200 items collected by the author from TV shows, newspapers, posted advertisements and previous research papers. All of the examples studied are intended for a German audience, but the material also includes English items, which testifies to the importance of language contact in the domain of wordplay. Moreover, the survey confirms the manageability of the taxonomy proposed and provides first insights into the importance of specific patterns of wordplay. Although the contribution is dedicated to the analysis of specific speech events, the findings thus also shed light on lexical innovation and productive patterns of lexical expansion.

The complex interplay between creativity and productivity is also addressed in Georgette Dal's and Fiammetta Namer's contribution on Playful nonce-formations in French: creativity and productivity. While nonce-formations have been in the focus of current research on English and German, there is still a lack of studies on French. In order to fill this gap, the authors draw on corpus data available to identify recurring patterns of the emergence of nonce-formations and distinguish between different subtypes of nonce-formations according to structural features as well as different ways in which the nonce-formations are embedded in the utterance context. Adopting an approach which is based on the speakers' and hearers' perspective on nonce-formations, they argue that nonce-formations represent a micro-system of its own. According to the authors, studying this micro-system requires a complete methodological reversal, focusing on the forms themselves and adopting other criteria of identifying nonce-formations than the standard tools used in morphological studies. In this way, their contribution also provides important general insights into the possibilities and challenges of approaching productivity, combining structural analyses with pragmatic reflections on issues related to the use of the items in individual communication events.

Finally, the contributions by Esme Winter-Froemel and Claudine Moulin, Ludicity in lexical innovation (I / II) – French / German, are dedicated to ludicity in the lexicon, taking into account ludic usage and lexicalised items that can convey ludic effects. Lexicographic sources, including contemporary dictionaries as well as historical dictionaries of both languages, are explored to investigate the importance of ludicity across different types of innovations, languages, periods, and contexts of use. Complementing each other, the two contributions argue that ludicity should be recognised as a basic aspect motivating lexical innovation alongside other factors of lexical expansion. At the same time, the authors show that the current lexicographic practice of marking ludic items is still in part unsatisfactory, as labelling of pertinent items is still only unsystematic and not exhaustive.

Moreover, Esme Winter-Froemel's paper focuses on the question of how the lexicographic data can be reinterpreted from a usage-based perspective. These reflections point to basic methodological challenges that need to be dealt with when studying ludicity in the lexicon. In addition, she analyses how the speakers and hearers produce and perceive ludic items, taking into account structural, semantic and pragmatic patterns that emerge from the data provided by the Petit Robert 2016 as well as historical dictionaries from the ARTFL database. From the basic features of ludicity identified, markedness emerges as a common denominator that enables speakers and hearers to use the items as a joint action, where both interlocutors demonstrate their linguistic mastery and engage in a game of complicity. A diachronic survey based on the historical dictionaries of French, most importantly different editions of the Dictionnaire de l'Académie française, reveals basic patterns of evolution, including the emergence of ludic items from citational uses and from a reinterpretation of obsolete items, patterns of relative stability as well as wearout effects by which the lexical items are retained, but lose their ludic dimension. In this way, ludic items are identified as a highly dynamic domain of the lexicon.

These findings are equally confirmed by Claudine Moulin's paper. Before studying ludic innovations in German, the author presents general methodological reflections on the difficulties of tracing ludic items in lexicographic sources across the history of German, and argues that sources of metalinguistic reflection provide helpful additional information on the ways ludic items are used and perceived in different historical contexts. Particularly interesting in this context are the extensive reflections on wordplay and related phenomena during the Baroque period in linguistic societies such as the Fruchtbringende Gesellschaft, with the main actors Justus Georg Schottelius, Georg Philipp Harsdörffer, Philipp von Zesen, and Kaspar Stieler. Historical dictionaries (Kramer, Adelung) and contemporary reference works (most importantly *Duden online* 2017) are analysed with respect to the ways in which ludic items are described and to diachronic patterns that can be observed in the creation and subsequent evolution of ludic items. The author shows that nominal compounds and diminutives play a predominant role in this context. Finally, certain pathways for the evolution of ludic items from the 18th century to current use are identified (+ludic > -ludic [+neutral]; +dialectal > -dialectal > +obsolete; -archaic -ludic > + archaic -ludic > +archaic +ludic). These pathways tie in with some of the pathways identified for French and confirm the strong dynamics that can be observed for ludic items in the lexicon.

In addition to the phenomena studied in the papers summarised here, certain effects of ludicity also appear in other domains, e.g. in the formulaic patterns studied by Natalia Filatkina, which also exhibit playful modifications. It can thus be argued that ludicity represents an important dimension of lexical expansion. At the same time, various contributions highlight the transitions between ludic and non-ludic usage and the necessity to assume a continuum between creative usage and conventionalised items of the lexicon conveying certain stylistic or pragmatic effects. This can be seen as an additional justification for a deliberately broad understanding of wordplay and ludicity, which also takes into account what could be labelled "borderline cases" of wordplay and ludicity. Studying these "marginal" phenomena thus also allows us to gain general insights into the dynamics of the lexicon.

#### 5 Acknowledgements

The workshop *Expanding the lexicon / Extensions du lexique / Erweiterungen des* Lexikons and the preparation of this volume would not have been possible without the financial support which we received from the University of Trier and the German National Research Foundation (Deutsche Forschungsgemeinschaft, DFG). Moreover, we received considerable practical help, and our thanks go to the staff at Trier University - Birgit Imade in particular and also Gabriele Jacobs - as well as to our student assistants Helin Baglar, Simon Eultgen, Samira Jung, Valérie Keppenne, Alistair Plum, Florian Prasch, Sarah Repplinger, Armin Rotzler, and Marie Winter – they all contributed to the success of the workshop with impressive dedication and efficiency.

We would also like to thank the reviewers who evaluated the papers submitted and gave constructive feedback for the contributions to this volume. Moreover, we would like to thank the members of the Editorial Board of the book series The Dynamics of Wordplay for accompanying the preparation and publication of this volume in their usual most constructive and efficient way.

We very much enjoyed the discussions at the workshop, which would not have been possible without the contributors' openness to multilingual and interdisciplinary exchange, bringing together different perspectives and approaches to processes of lexical expansion. This also became manifest during the preparation of this volume, as every contribution was commented on by several reviewers with different theoretical backgrounds in order to ensure the interdisciplinary accessibility of the papers.

In addition, we would like to thank our student assistants Sophia Fünfgeld and Constanze Tress for helping us with the French translations of the abstracts and the formatting of the volume, and Martina Bross and Angela Oakeshott for assisting us in the linguistic and stylistic revision of the papers.

Finally, as usual, the editorial team at De Gruyter's has accompanied the various stages of the preparation of this volume most sensitively, providing extremely efficient support at the various stages of its preparation. We would therefore also like to thank Gabrielle Cornefert, Antje-Kristin Mayr, and Ulrike Krauß.

#### 6 References

- Ágel, Vilmos, Andreas Gardt, Ulrike Haß-Zumkehr & Thorsten Roelcke (eds.). 2002. Das Wort Seine strukturelle und kulturelle Dimension. Festschrift für Oskar Reichmann zum 65. Geburtstag. Berlin & Boston: De Gruyter.
- Baayen, R. Harald, Petar Milin, Dusica Filipović Đurđević, Peter Hendrix & Marco Marelli. 2011. An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. Psychological Review 118. 438-482.
- Blank, Andreas. 1997. Prinzipien des lexikalischen Bedeutungswandels am Beispiel der romanischen Sprachen (Beihefte zur Zeitschrift für romanische Philologie 285). Berlin & Boston: De Gruyter.
- Brinton, Laurel J. & Elizabeth Closs Traugott. 2005. Lexicalization and language change (Research surveys in linguistics). Cambridge: Cambridge University Press.
- Dobrovol'skij, Dmitrij. 1988. Phraseologie als Objekt der Universalienlinguistik (Linguistische Studien). Leipzig: Verlag Enzyklopädie.
- Full, Bettina & Michelle Lecolle. In press. Jeux de mots et créativité. Langue(s), discours et littérature (The Dynamics of Wordplay 4). Berlin & Boston: De Gruyter.
- Haspelmath, Martin & Uri Tadmor (eds). 2009. Loanwords in the world's languages. A comparative handbook. Berlin & Boston: De Gruyter.
- Hay, Jennifer. 2003. Causes and consequences of word structure. New York & London: Routledge.
- LegIT. Der volkssprachige Wortschatz der Leges barbarorum (http://legit.ahd-portal.germling.uni-bamberg.de/, accessed 07 September 2017).
- Libben, Gary, Gonia Jarema & Chris Westbury (eds.). 2012. Methodological and analytic frontiers in lexical research (Benjamins Current Topics 47). Amsterdam & Philadelphia: John Benjamins.
- List of English Bound Words, https://www.english-linguistics.de/codii/codiibw/en/listcomplete.xhtml (accessed 05 September 2017).
- Norde, Muriel. 2009. Degrammaticalization. Oxford: Oxford University Press.
- Ostermann, Carolin. 2015. Cognitive lexicography. A new approach to lexicography making use of cognitive semantics (Lexicographica. Series Maior 149). Berlin & Boston: De Gruyter.
- Pirelli, Vito, Ingo Plag & Wolfgang U. Dressler (eds.). In press. Word knowledge and word usage: A cross-disciplinary guide to the mental lexicon. Berlin & New York: De Gruyter.

Zeschel, Arne. 2012. *Incipient productivity. A construction-based approach to linguistic creativity* (Cognitive Linguistics Research 49). Berlin & Boston: De Gruyter.

| Linguistic Innovation

#### Natalia Filatkina

# Expanding the lexicon through formulaic patterns

The emergence of formulaicity in language history and modern language use

**Abstract:** The article aims to study the role of formulaic patterns in the expansion of the lexicon. The notion of formulaic patterns is explained in section 1. It suggests that the formulaic character of human communication overarches single words, polylexical units, sentences and texts. As use of free word combination, formulaic patterns are a constitutive part of human interaction and, therefore, also of lexicon expansion. Section 2 provides a brief sketch of research findings (mostly based on data from standard German) concerning the interaction of formulaic patterns and word-formation products, which have up till now been considered the main tool of lexicon expansion. Here the argument is made that with regard to the new understanding of formulaic patterns, their role in the lexicon expansion process can be revised. Section 3 provides examples of the analysis of the emergence of formulaic patterns in language history and modern language use as an additional tool of lexicon expansion. In contrast to word formation, this has been subject to relatively little investigation so far. In section 3, the analysis is carried out against the background of language change theories. Such "driving forces" of language change as variation / creative modification, regularity / irregularity, codification / normatisation, the role of cultural and contextual / discourse traditions and frequency are applied to the emergence of formulaic patterns. As will be shown, the usual criteria with which we are familiar from existing language (change) theories do not apply to formulaic patterns in the same way as they do for example, to sound change, grammatical or even lexical change. The results of the study are summarized in the concluding section 4. \*

<sup>\*</sup> I would like to thank two anonymous reviewers, the editors of the volume, and Christian Pfeiffer for many insightful comments on an earlier version of this paper.

#### 1 The notion of formulaic patterns and their status in the lexicon

Speakers of any language generally enjoy considerable freedom in selecting lexical and grammatical items / tools of a given language in order to achieve their communicative goals most effectively. The success of a communicative act depends not only on the successful exploitation of a lexicon (good choice of individual words) and the correct application of grammatical rules, but also on an appropriate combination of words and rules with regard to the pragmatic and conventional aspects of a particular communicative situation. All forms of oral and written human interaction result from a large number of complex choices that Sinclair (1991: 109) described as "the open choice principle".

Nevertheless, Sinclair was also among the first scholars to empirically prove that although some word combinations, sentences and texts are the result of a complex choice based on linguistic freedom, others include "a large number of semi-preconstructed phrases that constitute single choices ("the idiom principle"), even though they might appear to be analysable into segments" (Sinclair 1991: 110). At the end of the 19th and the beginning of the 20th century, similar phenomena were recognized by Paul ([1880] 1995: 25), de Saussure ([1916] 1969: 177) and in Jespersen's concept of the "living grammar" (1968: 17–29). Corpus linguistics, usage-based approaches to language and cognitive sciences called attention to the fact that speakers' linguistic knowledge extends well beyond what can be described in terms of rules of compositional interpretation stated over combinations of single words. In the lexicon of a given language, preconstructed conventionalised items seem to be as productive as free word combinations. <sup>1</sup> This

<sup>1</sup> To my knowledge, much research remains to be undertaken as regards the quantification of this proportion in many languages. According to Sinclair, "the open choice principle" is even dominated by the "the idiom principle". For English and German, first figures have been provided in favour of this observation, cf. an overview in Filatkina (forthcoming: 44-48). With regard to a random sample of words starting with the letter f in a COBILD dictionary project, Stubbs (2001: 80-81) notes: "One phenomenon, by its sheer frequency, shows the strength of phraseological tendencies across the most frequent words in the language. Suppose we take all 47 wordforms which begin with f in the sample. In 41 cases, the following easily recognizable combinations account for the collocation of node and top collocate. [...] [NF: e.g.:] despite the fact that; faded away; fair enough [...]. In the remaining six cases, collocates further down the lists occur in recognizable phrases, such as: natural fabrics; animal feed, filing cabinet [...]. With many words, many more of the top 20 collocates are due to recognizable phrases. [...] I can think of no

idea has just started to find its way into linguistic analysis of modern languages. Depending on the research perspective, the terms *phraseme* or *Phraseologismus* (Burger 2015), lexical priming (Hove 2005), idiomatische Prägung (Feilke 1994), formelhafte Sprache (Stein 1995), formulaic language (Wray 2002), usuelle Wortverbindungen (Stever 2013), Sprachgebrauchsmuster (Bubenhofer 2009) or construction (Fillmore, Kay, and O'Connor 1988; Goldberg 1995) have been used in order to address this observation.<sup>2</sup>

For any linguistic theory that is based on a view of language as a system of signs (Systemlinguistik) or a conglomerate of dynamic grammar rules recruiting a static lexicon into sentence generation (Generative Grammar) such items pose a problem because they cannot be clearly attributed to one particular linguistic domain within this system, e.g. to the lexicon. Even though these items are highly lexicalised and conventionalised signs, their function tends rather to be one between grammar, lexicon, syntax and discourse or, as Wray (2008) puts it, they push the boundaries between these domains. Consider example (1a):

- (1) a. to brush one's teeth
  - \*to wash one's teeth b.
  - \*to clean one's teeth C.
  - French: se laver les dents lit. 'to wash the teeth'
    - German: sich die Zähne putzen lit. 'to clean the teeth'
    - Italian: pulire i denti lit. 'to clean the teeth'
    - Russian: чистить зубы (čistit' zuby) lit. 'to clean teeth'

The pattern (1a) can be used without any semantic difficulties for addressing a daily morning and evening sanitary activity, but is rather idiosyncratic with regard to the verb constituent: Examples (1b) and (1c) are formed with regard to the (same) rules of English grammar as (1a) and would therefore have to be regarded as correct. Their meaning will also be understood, but it would be confusing for a native speaker of English to hear them being used to name the same sanitary activity as (1a). The meaning in (1b) and (1c) is different from the meaning of example (1a). The explanation for this confusion lies in the fact that the preferred structure of this word combination in English favours the verb to brush and does

reason why a sample of words beginning with f might be untypical of the whole 1,000-word sample. We therefore have initial evidence that all of the most frequent lexical words in the vocabulary have a strong tendency to occur in well-attested phraseological units."

<sup>2</sup> For a complete overview and the substantial differences between these approaches cf. Filatkina (forthcoming).

not allow for its substitution without a change of meaning. The preferred structure becomes particularly apparent if compared to other languages (1d) where the preferred structures include a different verb constituent.<sup>3</sup>

Other examples are not only stable in terms of their formal structure. With regard to their form, they are quite regular as they are formed according to the rules of German grammar. However, with regard to their meaning, they are irregular as their holistic meaning is not predictable from the literal meaning of their individual constituents, i.e. it is idiomatic, cf. the modern German example (2a). The substitution of any single constituent even by family-resembling lexemes as in (2b) would destroy the idiomatic meaning.

- (2) a. Perlen vor die Säue werfen lit. "to cast pearls before swine" 'to offer something valuable to someone who does not know its value' b. \*Diamanten vor die Schweine werfen
  - lit. "to cast diamonds before pigs"

In order to use (2a) according to the linguistic conventions of modern German, one needs to know that with the preferred structure of this idiom *Das ist / wäre* Perlen vor die Säue (geworfen / zu werfen) lit. "it is / would be pearls (cast) before swine" one can comment on any type of useless action that a person executes and another one does not appreciate, but only in colloquial speech. Within the framework of traditional approaches, formulaic patterns with semantic irregularity such as (2) were considered rare "exceptions" mostly satisfying stylistic or aesthetic, not essential communicative needs. Consequently, they were not a central focus of theoretical linguistic studies.

An extensive attempt to grasp the complex nature of such utterances was undertaken within the framework of phraseology. The complexity was already reflected in the defining criteria of phrasemes. According to Burger (2015), phrasemes are polylexical items that must consist of at least two constituents, have a more or less stable form in which they are frequently reproduced by speakers and can be idiomatic in meaning. Research traditionally focused mainly on one type of polylexical word combination, namely idioms such as in (2) or English spill the beans or break the ice, because they were considered to be at the centre of the phraseological system. But as usage-based approaches show, the formulaic

<sup>3</sup> Though language contact plays a role in lexicon expansion with the help of formulaic patterns, for reasons of space, it cannot be touched upon in this article. The methodological and theoretical importance of a contrastive perspective at such a core level as determining what is formulaic in a historical text is briefly pointed out in footnote 17.

character of human communication reaches far beyond the items that can meet the criteria of phrasemes. It extends beyond single word conventionalised structures such as routine formulae and?, congratulations!, truly (speaking), adverbial / prepositional constructions like *nonwithstanding* or text markers such as Middle High German firnim 'remember, memorize, pay attention' on the one hand and formulaic text genres such as contracts, business correspondence, newsletters, recipes, announcements etc. on the other. The texts are formulaic because they can be produced and understood correctly only if they follow the conventionalised traditions of their formulaic matrix. Further examples of frequently used patterns that have largely been excluded from the scope of research into phraseology are listed in (3):

- (3) a. German: allen Grund (haben), allen X zum Trotz, allen Ernstes, auch immer, nicht zuletzt
  - lit. "(to have) all the reason, in spite of all X, quite seriously, also always, not least" (Steyer 2013: 239-287)
  - English: you take, a little bit, one X after another, NP or something (Langacker 1987: 35–36)

Moreover, the criteria established for phrasemes on the basis of modern languages turn out to be static and therefore not applicable to the study of the diachronic dynamics of formulaic patterns. Polylexicality appears to be problematic from the outset because of the general lack of any (mandatory) spelling norms in the language history. As will be shown in section 3, stability is the exception rather than the rule in historical language use, frequency cannot be employed due to the fragmentary character of historical textual heritage (among other more substantial restraints), and idiomaticity often poses problems resulting from the temporal and cultural distance between today's researcher and the text under investigation.

This is why in Filatkina (forthcoming) typologically heterogeneous units (1–3), single words and whole texts are described as *formulaic patterns* in a wider sense. I will use this term in the following article although it is not yet well-established within linguistic research. Based on the analysis of an extensive data set from Old German, the following definition of formulaic patterns is proposed:

Formelhaft sind im weitesten Sinn:

- Einwortausdrücke, typologisch heterogene Kombinationen aus mehreren Konstituenten bzw. ganze Sätze und / oder Texte.
- die holistisch verstanden werden müssen,
- sich auf unterschiedlichen (auch noch nicht abgeschlossenen) Stadien der formalen, semantischen und funktionalen Konventionalisierung befinden können,

- aber eine stabile zugrundeliegende syntaktische und / oder kognitive Struktur aufweisen.
- auf Gebrauchskonventionen einer Sprachgemeinschaft beruhen, deren etablierte kulturelle (auch kommunikative) Erfahrungen und Wissensbestände sie tradieren,
- die sich durch eine starke Funktionalisierung im Kommunikationsprozess bzw. im e) Textaufbau auszeichnen können (Filatkina forthcoming: 2-3 and 151-156).

[Formulaic patterns in the broadest sense are:

- single words, typologically heterogeneous combinations of words, sentences and / a) or texts
- that must be understood holistically, b)
- c) can show varying degrees of conventionalisation (ranging from high to low) with regard to their form, meaning and functions, but have a stable underlying syntactic and / or cognitive structure,
- d) are based on and reflect the cultural and communicative traditions of the society they are used in, and
- e) which can be characterised by a considerable degree of functionalisation in the production and reception of a particular act of oral communication, written text (genre) or discourse (translation: NF)].

Formulaic patterns provide evidence for the necessity of understanding language as a continuum of different linguistic and extra-linguistic domains that have to be described in their entirety. Current usage-based linguistic theories systematically develop the notion of a language as an entirety. Within the paradigm of Construction Grammar, for example, formulaic patterns have played a central role from the very beginning (Langacker 1987; Fillmore, Kay, and O'Connor 1988; Goldberg 1995). In fact, it was the inability of other (particularly formal) language theories to describe "exceptions", i.e. formulaic utterances as in (1-3), that led to the establishment of Construction Grammar. One of its major principles is the assumption that a human language consists of signs representing conventionalised form / meaning correspondences that are not strictly predictable from the properties of their component parts or from other constructions. The term *construction* is generally applied to generalisations over typologically very different language instances, regular and irregular, ranging from morphemes and compounds (door frame or lighthouse) to idioms (spill the beans) and degree modifiers (sort of / kind of) to abstract constructions such as caused-motion, ditransitive or resultative constructions. They differ with regard to their cognitive representations (from concrete utterances on the language surface to abstract cognitive schemas) but all tend to have a more or less restricted structure that has a certain meaning as

well as different lexical slots whose specification can vary depending on the context. All these extremely heterogeneous constructions stand on equal footing in building the basis for human communication and understanding processes, without being ascribed exclusively to core grammar or to the lexicon. The difference between the terms formulaic patterns and constructions is twofold: the former does not include morphemes but extends its scope to formulaic texts and discourse; the latter prototypically does not include texts (cf. a different approach in Östman 2005), but incorporates morphemes.

The usage-based perspective changes the status of formulaic patterns from peripheral (stylistic or aesthetic) "exceptions" to central means of human interaction. Consequently, it also sheds fresh light on their role as tools of lexicon expansion. Referring to features c), d) and e) from the above definition of formulaic patterns, this point will be made in section 3 and applied to the emergence of formulaic patterns in language history and modern language use.

### 2 Formulaic patterns, word formation, and lexicon expansion

With regard to their function as a means of lexicon expansion, polylexical word combinations were already studied in early research on phraseology. The term formulaic pattern was not used in this paradigm. As noted above, research traditionally focused mainly on idioms. Their contribution to the expansion of the lexicon was compared to that of word-formation products (Fleischer 1992; Barz 2005; Stein 2012). At least for German, there is a vast amount of literature dedicated to this topic.4 But with a focus on idioms, phraseology was treated as the rarest and least significant path (Barz 2005: 1673; Barz 2007: 30; Stein 2012: 228). Taking into consideration the pivotal role of formulaic patterns in the communication process (cf. section 1), such a conclusion cannot be sustained. The "old" field is opening up for new discussions guided by the assumption that artificial boundaries between single words and formulaic patterns might be a misleading perspective.<sup>5</sup>

<sup>4</sup> In addition to the above-mentioned work of W. Fleischer cf. Hartmann (1998), Barz (2005, 2007) and Stein (2012).

<sup>5</sup> In its turn, research on word-formation has traditionally pursued the idea that the development of new words is formulaic in nature as it generally functions according to specific patterns, e.g. certain productive types of derivation, composition and conversion that may differ in their productivity from language to language. For new insights cf. Arndt-Lappe (2015).

In the traditional research, attention was drawn to the many similarities or "the fuzziness" of the boundaries between compounds and idioms. These were explained by a number of facts. In addition to the shared "naming" function, both tools of lexicon expansion can be products of idiomatisation, e.g. (4):

(4) German: ein großes Tier
lit. "a big animal"
'an important and influential person'
German: Grünschnabel
lit. "green beak"
'a young, inexperienced but often cheeky person'

Consequently, compounds and idioms undergo similar lexicalisation processes with metaphorisation and metonymisation being the most productive. With regard to idiomatisation, compounds and idioms were proclaimed complex lexical signs whose meaning is not derivable from the meaning of their constituents.

It was also pointed out in previous research that sharing the referential function of naming means competition between phrasemes and word-formation products in some cases and complementarity in others (Barz 2007: 27–29). The cases of competition include the coexistence of a phraseme and a word-formation product that both use the same lexical constituents, e.g. idiom (4) ein großes Tier 'an important and influential person' versus compound Großtier 'a big animal'. Strictly speaking, such utterances do not compete as they differ semantically. Examples of semantically similar utterances can be found as well, cf. German stark wie ein Bär sein versus bärenstark, Schwarzer Markt versus Schwarzmarkt. However, they do not seem to be widespread. In cases of complementarity, a wordformation, e.g. Grünschnabel (4), does not have an immediate equivalent among phrasemes and vice versa. Due to the fact that the communicative needs of the speakers are met either by a word-formation product or by a phraseme, the simultaneous existence of both appears to be unnecessary. Again, the focus on idioms led previous research to the conclusion that polylexical utterances are particularly productive in negatively connotated target domains such as HUMAN MISBEHAVIOUR (deception), CHARACTER (stupidity), STATE (drunkenness) or INTERPER-SONAL RELATIONS (reprehension) (Fleischer 1992, 1996, 1997). Although this seems to be true for idioms, a different understanding of formulaic patterns sheds fresh light on this research question as well. Recent studies that employ the concept of Construction Grammar demonstrate that in the process of name creation lexicalised phrases, e.g. A + N phrases rote Karte 'red card', may function as names just as A + N compounds (Freikarte 'free ticket') do. The choice between these two forms is governed by the principle of analogy: It is largely dependent on the availability of similar constructions in the mental lexicon of the speakers (Schlücker

and Plag 2011: 1539). Lexicalised phrases and compounds are equally productive constructions that make distinctions between lexicon (compounds) and syntax (phrases) irrelevant for language users.

Another well investigated area of the "joint action" of phrasemes and wordformation as tools of lexicon expansion is the use of phrasemes as a basis for the creation of new words. In Germanic linguistics, the phenomenon has been addressed as dephrasemische / dephraseologische Wortbildung (Fleischer 1992; Stein 2012: 231–233). It is illustrated in (5a) by means of an example from modern German. Interestingly, even irregular constituents as in German Fettnäpfchen "little pot of fat" in (5b) take part in lexicon expansion. The constituent is irregular because it is obsolete and opaque with regard to the underlying cultural knowledge (an old custom in traditional farmhouses of placing a small pot to collect fat near the stove, cf. Röhrich 2004) for the majority of the native speakers of German. In dictionaries of modern German (duden.de; dwds.de), it is noted as bound to this idiom. However, according to the corpus analysis in Stumpf (2015a: 497), the actual boundness of the constituent to the idiom does not exceed 66%. This means that in the remaining 34% of all contexts studied in (Stumpf 2015a) Fettnäpfchen also occurs in isolation; its meaning, then, is the same as its correspondent meaning in the idiom. Thus, the possibility of re-motivating the compound synchronically without linking it to the underlying cultural knowledge opens up this irregular constituent for "free usage" in the lexicon.

- (5) a. Haare spalten > Haarspalterei lit. "to split hairs" > "hair splitting" 'to be excessively precise, pedantic'
  - bei jemandem ins Fettnäpfchen treten > Fettnäpfchen lit. "to step in in a little pot of fat" > "little pot of fat" 'to drop a clanger'

<sup>6</sup> More precisely, Schlücker and Plag (2011: 1539) note: "The larger the number of lexicalized compounds with the same adjective or noun, the higher the probability of the subjects choosing a compound. The larger the number of lexicalized phrases with the same adjective or noun, the higher the probability of the subjects choosing a phrase."

<sup>7</sup> For further examples see also the contribution by Stumpf 2017. The role of irregularity in the development of formulaic patterns will be studied in section 3.2.

### 3 The emergence of formulaic patterns and the principles of language change

An alternative approach to the comparison of word-formation products and formulaic patterns which can help to answer the question of the nature of lexicon expansion is the analysis of the dynamics of the emergence of formulaic patterns in language history and modern language use. In particular, studying diachronic processes of the emergence of what is considered formulaic in modern languages can provide the necessary insights. However, at the present stage of international research, for the majority of languages, the implementation of this approach faces methodological difficulties, a theoretical vacuum and most importantly the lack of empirical data (Filatkina 2012, 2013, forthcoming). Since its establishment in the 19th century, historical linguistics has focused strongly on the analysis of the "open choice principle" and on the description of various but single and isolated linguistic domains such as phonetics, grammar or the lexicon. The historical roots of the other basis of human communication, "the idiom principle", remain without exception a fundamental research question for all languages. The diachronic study of the emergence of formulaic patterns is often neglected entirely, even in publications claiming the status of reference works on language change (for a detailed overview cf. Filatkina, forthcoming). However, the research conducted for Old German (Filatkina 2009, 2012, forthcoming)<sup>8</sup> shows that analysing formulaic patterns can cast new light on the existing language (change) theories and the understanding of lexicon expansion. The main point is that the accepted criteria with which we are familiar from existing theories do not apply to formulaic patterns in the same way as, for example, to sound change, grammatical or even lexical change. Such criteria as variation / creative modification, regularity / irregularity, codification / normatisation as well as the role of cultural and textual / discourse traditions and frequency of use are the subject of discussion in the present section.

<sup>8</sup> One possible methodology to detect and extract novel formulaic patterns from modern oral and written texts is shown in Schreiber, Mahlow, and Juska-Bacher (2012).

#### 3.1 Formulaic patterns and the role of variation / creative modification

In any natural language, even pre-constructed formulaic patterns are never absolutely stable and unchangeable, cf. feature c) in the definition of formulaic patterns in section 1. This point has already been made by classical research on phraseology and has led to a shift of paradigms (Burger 2015). Although in the collocation to brush one's teeth verb substitution is not allowed, as shown in (1), different types of grammatical and lexical variation do not violate conventional usage: to brush my teeth, to brush and polish one's teeth, the teeth were brushed. to brush the front teeth. One of the major achievements of phraseological research in recent years is the understanding that even highly idiomatic units, such as German Perlen vor die Säue werfen (2), are not as fixed as has previously been thought. On the other hand, as was pointed out in section 1, computer linguistics, cognitive sciences and most recently Construction Grammar suggest that free entities of a language are not so free but rather pre-constructed. Thus, in any modern language, variation does not contradict but faithfully accompanies formulaicity.

The diachronic investigation of formulaic patterns also supports the view that such patterns are less characterised by syntactic fixedness than has often been assumed. At the historical stages of the language, we see that fixedness or stability can only be attributed to a basic structure underlying a formulaic pattern. As a whole, this pattern possesses a certain meaning, pragmatic function and structure, but both the filling of its lexical slots and grammatical elements are only in the process of being formed. The patterns that might be considered formulaic in a certain language at the current point in time are always products of a process of change, which is inherently enabled by variation – the most natural form of existence of any actively used language and the driving force of any change.9 As shown in Filatkina (2013), formulaic patterns undergo diachronic changes at all levels: structure, semantics, pragmatics, ways of syntactic contextualisation, distribution in texts, stylistic connotations, frequency of use, degree of familiarity, cultural image component and so on. The idiom *Perlen vor die Säue* werfen (2), for example, occurs 33 times in German texts from the 9th to 16th century (cf. the corpus description in Filatkina, forthcoming). Each time, however, it has a different structure and syntactic contextualization, and moreover it also reveals a semantic change from a very narrow meaning (which can only be found in reli-

<sup>9</sup> For English, cf. Corrigan, Moravcsik, Ouali, and Wheatley (2009: XVI).